Embodied Pedagogical Agents: From Visual Impact to Pedagogical Implications

Haake, Magnus

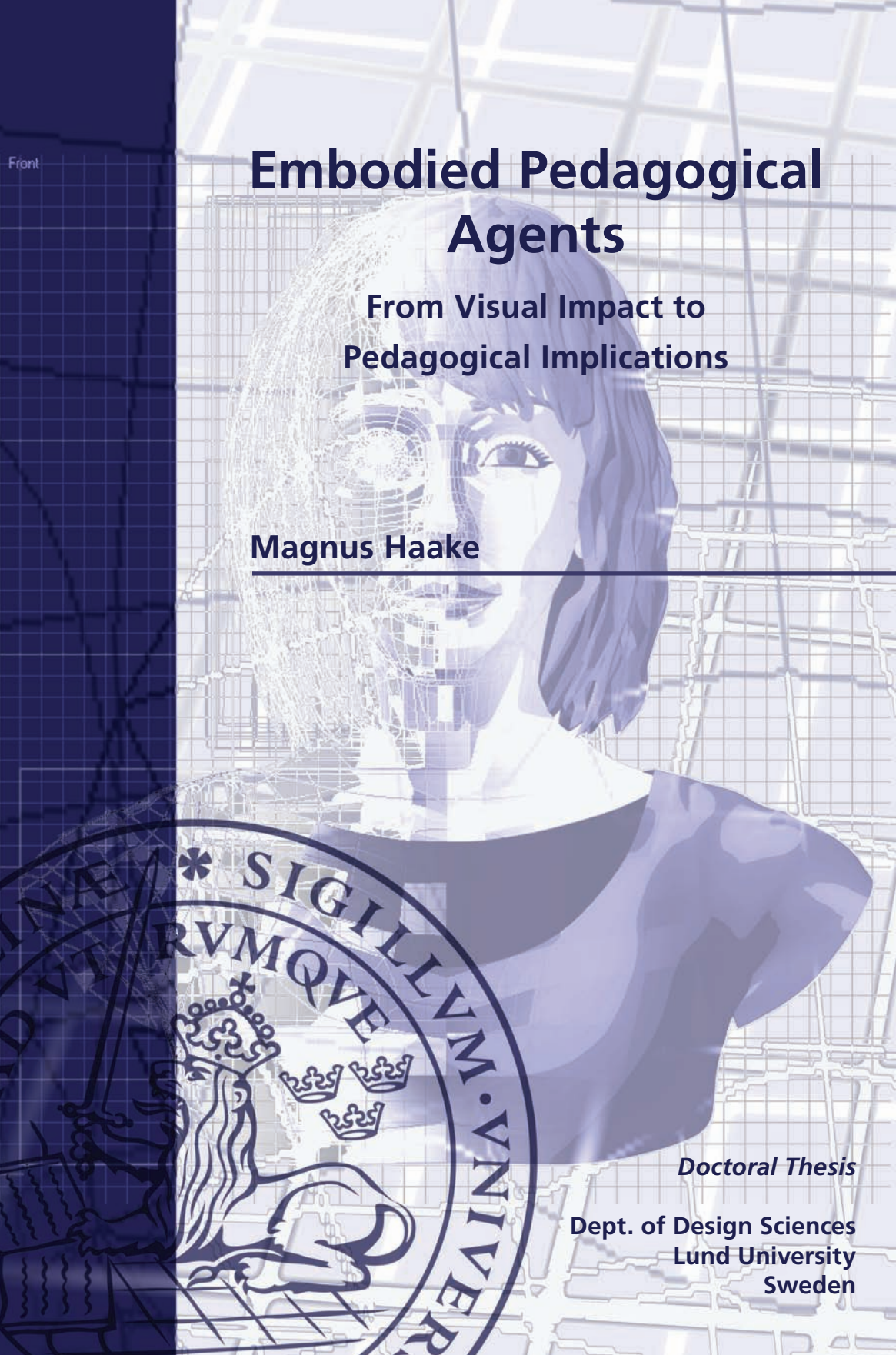
2009

Front

Embodied Pedagogical Agents

From Visual Impact to Pedagogical Implications

Magnus Haake



Doctoral Thesis

Dept. of Design Sciences
Lund University
Sweden

To Agneta – my everything …

… and to Sebastian and Igis – who, with their childish egoism, never have given me a real chance to confuse the reality of life with the virtuality of academia.

Embodied Pedagogical Agents

From Visual Impact to Pedagogical Implications

Magnus Haake

DOCTORAL THESIS

Department of Design Sciences
Faculty of Engineering
Lund University
Sweden

Magnus Haake:
Embodied Pedagogical Agents
From Visual Impact to
Pedagogical Implications

Div. of Ergonomics
Dept. of Design Sciences
Faculty of Engineering
Lund University
Sweden
http://www.design.lth.se/

ISSN 1650-9773
Publication 32
ISRN LUTMDN/TMAT --1032--SE
EAT 2009

ISBN 978-91-628-7804-7

Format: 154 x 233 mm
Typefaces: Adobe Garamond
Print: E-husets tryckeri, Lund, Sweden

Acknowledgements

I have had a position that is becoming rare in Sweden. As a faculty employed doctoral student with no specific project funding, I have found myself with a marvellous opportunity to pursue research out of pure curiosity. For this I am very much obliged to my supervisor Gerd Johansson that once offered me this position. Regarding the otherwise most peculiar Swedish academic system, it seems to me that most people carry out their research not thanks to, but rather 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 LUCS (Cognitive Science Lund University). In particular I want to thank our marvellous secretary Karin Öhrvik for helping us all with all our bizarre 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. I also want to thank Henrik Enquist for being even more of a sceptic than myself.

As to this very thesis, I owe a lot to my wife Agneta who has been my research colleague for many years. Since long 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, who have had their share of strange conversations around the dinner table. Especially, I remember Sebastian once asking: 'Why has everything something with design in it?'

Magnus Haake
Lund, 19/4 2009

Abstract

Embodied pedagogical agents – visually represented, computer generated characters in pedagogical roles, such as virtual instructors, mentors and learning companions – populate the digital society in increasing numbers. They are found in educational programmes from preschool to university, as well as in broader educational contexts in the form of medical counsellors, virtual health coaches, information guides, etc. Since long, they also appear in edutainment and infotainment settings.

The thesis addresses psychological, cognitive and social aspects of embodied virtual characters in pedagogical contexts with respect to 'static visual characteristics', i.e. underlying visual characteristics. On the basis of theoretical considerations as well as several empirical studies, it is argued that users' visual and aesthetic experience of embodied pedagogical characters (EPCs) is too important to be treated as a secondary issue – not the least with respect to central pedagogical goals to motivate and engage.

The main conclusions of the thesis are: (i) that visual aspects in EPCs can be related to pedagogical outcomes, (ii) that there may be significant relations between visual and social characteristics of EPCs and learner characteristics, and (iii) that EPCs may reproduce stereotypes from everyday real life human-human interaction, as well as from traditional visual media – but that they simultaneously harbour a considerable potential to challenge stereotypes.

As a tool for the research community, a framework of a visual graphical design space is proposed. The framework may scaffold the design of user evaluations, which often are blurred by uncontrolled visually related variables. With regard to weak user evaluations, the thesis furthermore argues that there is a problem with user evaluations based upon the concept of 'the User', i.e. an averaged standardized user. In order to reveal important correlations between agent characteristics and user characteristics, it is often necessary to identify relevant subgroups of users rather than to treat the whole user group as an 'averaged user'.

And finally – virtual agents may have a not yet thought of potential as research tools …

Sammanfattning

Virtuella pedagogiska karaktärer (*embodied pedagogical agents*) – visuellt representerade, datorgenererade karaktärer i pedagogiska roller, såsom virtuella instruktörer, handledare och lärkompanjoner – blir allt vanligare i det digitala samhället. De finns nu i lärprogram från förskola till universitet, liksom i bredare utbildningssammanhang i form av medicinska rådgivare, personliga tränare, informatörer, etc. De har också sedan länge förekommit inom *edutainment* och *infotainment*.

Avhandlingen diskuterar psykologiska, kognitiva och sociala aspekter hos virtuella karaktärer i pedagogiska sammanhang med avseende på så kallade ”statiska visuella egenskaper”, dvs. underliggande visuella egenskaper. Utifrån teoretiska överväganden samt empiriska studier hävdas att användares visuella och estetiska upplevelse av virtuella pedagogiska karaktärer (EPCs) är alltför viktig för att behandlas som en sekundär frågeställning – inte minst med hänsyn till viktiga pedagogiska egenskaper som motivation och engagemang.

De viktigaste slutsatserna i avhandlingen är: (i) att visuella aspekter i EPCs kan relateras till pedagogiska resultat, (ii) att det kan finnas viktiga samband mellan visuella och sociala egenskaper hos EPCs och egenskaper hos användare, och (iii) att EPCs reproducerar stereotypa föreställningar från det vardagliga samspelet mellan människor, likväl som stereotypa föreställningar från traditionell media – samtidigt som de besitter en förmåga att utmana stereotyper.

Som ett verktyg för forskare presenteras också ett ramverk för en visuell (grafisk) designrymd. Detta ramverk kan stödja utformning av användarutvärderingar som ofta blir osäkra på grund av att man inte kontrollerar olika visuella variabler. I samband med användarutvärderingar lyfter avhandlingen också fram problemet med föreställningen om *the User*, dvs. en genomsnittlig, standardiserad användare. För att upptäcka viktiga samband mellan egenskaper hos agenten och egenskaper hos användaren, är det ofta viktigt att identifiera relevanta undergrupper av användare i stället för att betrakta hela användargruppen som en medelvärdesbildad användare.

Till sist – visuella virtuella agenter kan även komma att få en viktig roll som forskningsverktyg …

.

Table of Contents

List of Included Papers

This thesis is based upon the following five papers, which will be referred to in the text by their Roman numerals (Paper I, II, III, IV & V).

The Papers

I. Design of Animated Pedagogical Agents – a Look at Their Look
Gulz, A. & Haake, M. (2006). *International Journal of Human-Computer Studies*, Volume 64(4), pp. 322-339.

II. Visual Stereotypes and Virtual Pedagogical Agents
Haake, M. & Gulz, A. (2008). *Educational Technology and Society*, 11(4), pp. 1-15.

III. A Look at the Roles of Look & Roles in Embodied Pedagogical Agents – a User Preference Perspective
Haake, M. & Gulz, A. (2009). *International Journal of Artificial Intelligence in Education*, 19(1), pp. 39-71.

IV. Visual femininity and masculinity in synthetic characters & patterns of affect
Gulz, A. Ahlnér, F. & Haake, M. (2007). In: A. Paiva, R. Prada & R. Picard (Eds.): *Lecture Notes in Computer Science (LNCS) 4738. Proceedings of the Second International Conference on Affective Computing and Intelligent Interaction*, Lisbon, Portugal, September 2007, pp. 654-665. Berlin/ Heidelberg, Germany: Springer.

V. Challenging Gender Stereotypes using Virtual Pedagogical Characters
Gulz, A. & Haake, M. (in press). In: Goodman, S. Booth & G. Kirkup (Eds.): *Gender Issues in Learning and Working with Information Technology*, IG Global.

Papers I and II
Both authors participated in all the research activities reflected and presented in the paper. Both authors are also equally responsible for the text, with the following exceptions: Haake had the major responsibility for the production and analysis of visual materials, and Gulz had the major responsibility for the pedagogical perspective.

Paper III
Both authors contributed equally to the design of the study as well as the collection and analysis of the empirical data. The software used in the study was produced by a collaborator (master student). The two authors are equally responsible for the text with the following exceptions: Haake is largely responsible for the theoretical framework of the visual design space presented and for the production and analysis of visual materials, and Gulz is largely responsible for the theoretical parts that are pedagogically oriented.

Paper IV
The three authors together with a fourth collaborator, (a master student), contributed equally to the studies presented in the paper. Haake designed and developed the main part of the digital applications used in the studies. The text was produced primarily by Gulz and Haake in collaboration with Ahlnér.

Paper V
Both authors contributed equally to the design of the study. Gulz had a major responsibility for conducting the study and Haake had a major responsibility for the production of the digital application used in the study. The authors contributed equally much to the analysis of the empirical materials (with assistance from a third collaborator). Both authors have been involved to an equal extent in the production of the paper with the exception that Haake has had a major responsibility for the visual material.

Other Relevant Work by the Respondent

Haake, M. (2004) *Smakens irrvägar: Reflektioner kring en kognitiv-biologisk ansats för upplevelser*. Course paper (Perspektiv på smak, 2003-04), Lund University, Sweden.

Haake, M. & Gulz, A. (2005). Aesthetic Stereotypes and Virtual Pedagogical Agents. In: *Abstracts of the International Conference on ICT, Media & Learning - Moving Technology, Changing Society (IML'05)*, Copenhagen, Denmark.

Haake, M. & A Gulz (2005). A look at the roles of look & roles in virtual pedagogical agents. *Shortpaper for IUI 2006*.

Gulz, A. & M Haake (2005). Social and visual style in virtual pedagogical agents. In: *Proceedings of UM 2005 Workshop on Adapting the Interaction Style to Affective Factors*, Edinburgh, Scotland.

Gulz, A. & Haake, M. (2006). Visual design of virtual pedagogical agents: Naturalism versus stylization in static appearance. In: *Proceedings of the 6th International Conference on Intelligent Virtual Agents (IVA'06), Lecture Notes in Computer Science*, vol. 4133, p. 455. Berlin/Heidelberg, Germany: Springer.

Gulz, A. & Haake, M. (2006). Visual design of virtual pedagogical agents: naturalism versus stylization in static appearance. In: *Proceedings of the 3rd International Design for Engagement Conference @ NordiCHI 2006 (iDec3)*, Oslo, Norway.

Gulz, A. & Haake, M. (2006). Virtual pedagogical agents – design guidelines regarding visual appearance and pedagogical roles. In: *Proceedings of the IV International Conference on Multimedia and Information and Communication Technologies in Education (m-ICTE2006)*, vol. III, pp. 1848-1852. Seville, Spain.

Haake, M. & A Gulz (2007). Virtual pedagogical agents: stylisation for engagement. *Interfaces Magazine*, vol. 70, pp. 12-13.

Johansson, P, Hall, L., Gulz, A., Haake, M. & Watanabe, K. (2007). Choice blindness and trust in the virtual world. In: *Technical report of Institute of Electronics, Information and Communication Engineers. (IEICE), Human Information Proceessing (HIP)*, vol.107, no. 60, pp. 83-86.

Gulz, A.,, M Haake & B. Tärning (2007). Challenging gender stereotypes using virtual pedagogical characters. In: *Seminar papers: GLIT – Gender, Learning and Information Technology*, Helsingborg, Sweden.

Gulz, A. & M Haake (2007). ICT, learning and the potential of androgynity. In: *Proceedings of the Educational Conference (Pedagogisk utvecklingskonferens)*, Lund University.

Gulz, A., M Haake & B. Tärning (2007). Visual gender and its motivational and cognitive effects – a usersStudy. *Lund University Cognitive Studies*, no. 137.

Gulz, A., & Haake, M. (in press). *Virtuella rollmodeller och androgynitetens potential*. In: J. Linderoth (Ed.), Anthology.

Labbé, D., Haake, M. & Gulz, A. (submitted). *Avatar Appearance in Computer Supported Learning*.

Lindström, P., Pareto, L., Haake, M. & Gulz, A. (submitted). *An Affective Side of Learning by Teaching*.

Gulz, A., & Haake, M. (2009). Benefits of digital pedagogical characters – a learning style perspective. *The Open Education Journal*, vol. 2, no. 2, pp. 34-41.

Part I

Embodied Pedagogical Agents: an Overview

This first part is a survey of embodied pedagogical agents, setting out on a journey from the hard computational roots of logic to socio-cultural implications.

Introduction: A New Social Actor

: -)

The Embodied Virtual Character – a New Social Actor

In the 1970s, a new kind of social actor saw the light – visually represented, computer generated *embodied virtual characters*. The immediate source of this was the advancement of video technology for visualizing animated virtual characters in commercial technical solutions. An immediate and important effect was that a larger group of people from then on had access to a digital arena and started to interact with (or mostly shoot down) these embodied virtual creatures such as the aliens in *Space Invaders* (Figure 1, left).

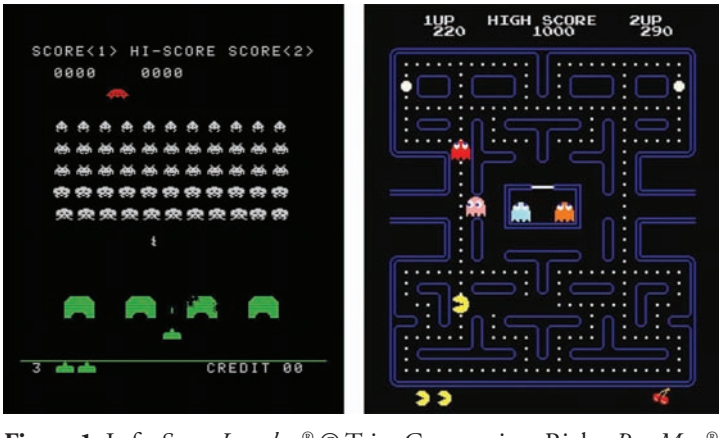


Figure 1. Left: *Space Invaders*® © Taito Corporation; Right: *Pac-Man*® © NAMCO BANDAI Games America Inc.

A step towards a more interactive and social embodied virtual character (or creature) came with the introduction of *Pac-Man* in 1980 (Figure 1, right). Pac-Man was a breakthrough in video game history, creating a whole new game genre and introducing an interactive (even if primitive) virtual world inhabited by 'hungry'

virtual creatures. Furthermore, Pac-Man initiated an early gender busting effect as it came to appeal to both males and females (Lammers, 1986).

Figure 2. Top left: IKEAs chat bot *Anna* © Inter IKEA Systems B.V.; Top middle: *Sam* © GNL, MIT Media Lab; Top right: Avatars in *World of Warcraft*™ © Blizzard Entertainment, Inc.; Bottom left: *Laura* © MIT Media Lab; Bottom middle: *AutoTutor* © University of Memphis, TN, Bottom right: USA; Characters from *FearNot* © eCIRCUS.

Now, thirty years later, people all over the world interact with embodied virtual characters in the form of *avatars* in multiplayer online games like *World of Warcraft*, interactive chat bots like *IKEAs Anna*, and embodied pedagogical agents like *Laura*, *AutoTutor*, and *FearNot* agents – to mention a few (Figure 2). During these years, the embodied virtual characters have developed from simple low resolution creatures like Pac-Man to increasingly autonomous and 'intelligent' high resolution characters that can be experienced in digital games. Furthermore, with every new generation of digital characters, the border between reality and virtuality is slowly becoming more and more diffuse.

Already in this short introduction, terminology abounds with concepts such as: agent, avatar, chat bot, embodiment, etc. – and since this field is relatively young and divergent, the use of these concepts is far from consistent and stringent. In what follows, I establish the basic terminology to be used in this thesis. The starting point is a short historical résumé of artificial intelligence.

Chapter 1:
The Computational Heritage

:-10

Artificial Intelligence – the Root of it All

In this thesis, I discuss psychological, cognitive and social aspects of embodied virtual characters in pedagogical contexts. However, without the development and progress in *Computer Science* and especially the field of *Artificial Intelligence (AI)*, there would be no virtual embodied pedagogical characters to talk (or write) about from psychological, cognitive, and social perspectives. To establish a foundation for this thesis and its discussions, I will thus start from the beginning with the longstanding idea of an artificial being.

A short pre-history of artificial intelligence

The idea of an artificial being can already be found in ancient Greek mythology, in the self-operating machines known as *automata* made by the divine blacksmith Hephaestus, as well as in the ivory statue of *Galatea* sculptured by Pygmalion and brought to life by Aphrodite (Crevier, 1993). These myths probably have even older precursors. From ancient Egypt testimonials originate of mechanical statues of the gods with deceit voice passages to the mouth, secretly controlled by priests (Brooks, 2002; Crevier, 1993). Later on through history there are many reports of artificial beings and artefacts, such as the *golem*, an animated artificial being in Jewish mythology.

In literature, Mary Shelley's *Frankenstein; or, the Modern Prometheus* (1818) became a milestone, shedding light on a fundamental topic in artificial intelligence: Can an artificial being have feelings? In his short novel, *I Robot* (1950), Isaac Asimov puts forth a set of ethics for robots which had great impact on later literature and film, as well as academia (Encyclopædia Britannica Online[1], 2009). Today, artificial intelligence is a well-known ingredient in popular media in numerous thrillers and science fiction movies like *Star Trek*, *Terminator*, *Alien*,

[1] *Encyclopædia Britannica Online*, henceforth abbreviated 'EBO'

Blade Runner, and *Matrix* (to mention a few). Notable is the recurrent conflict between (artificial) logic-rational reasoning and (human) emotional reasoning that is used as a both humorous and philosophical ingredient. One nice example is the internal logic-emotional conflict in the half human, half 'Vulcan' (alien) *Mr Spock* in the space odyssey series *Star Trek*.

Turning to actual implementations with relevance for the field of artificial intelligence, the mechanical clock (15th - 16th centuries) was a big step in the construction of mechanical devices and was soon followed by mechanical animals (16th century) and later on toys (18th century). In the meantime, Blaise Pascal constructed one of the first digital mechanical calculators in 1642-44. In the 19th century, Charles Babbage and Ada Lovelace worked on the principles for a programmable mechanical calculating machine (EBO, 2009).

The beginning of the 20th century saw the breakthrough of modern formal logic with seminal works like *Principia Mathematica* by Bertrand Russell & Alfred Whitehead (1910-13) and the prelude to modern computer science through names like Alan Turning (EBO, 2009).

During the mid 20th century, scientists from different academic fields (mathematics, psychology, engineering, economics, and political science, among others) began to discuss the possibility of constructing an artificial intellect. Shortly thereafter, forces united and the academic field of *Artificial Intelligence* is generally acknowledged to have been established 1956.[2]

Artificial intelligence from 1956 to the present

Artificial intelligence has had a turbulent history since 1956 – pendulating between periods of spiralling enthusiasm and paralyzing set-backs. The following résumé relies to a large extent on the book *AI: The Tumultuous History of the Search for Artificial Intelligence* by Daniel Crevier (1993).

The initial hype

The first period following the establishment of artificial intelligence as a field of research was enthusiastic and expansive. Supported by generous governmental funding, AI laboratories were founded at MIT, Carnegie Mellon and Stanford. Within a short time, new and seemingly astonishing computer programs were

[2] The academic field of Artificial Intelligence is generally acknowledged to have been established at the Dartmouth Conference of 1956. The proposal for the Conference included the following assertion: '*Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it.*' (Crevier, 1993, p. 48).

solving problems in algebra, proving logical theorems and speaking English. This initial success prompted optimistic prognoses like:

> *[M]achines will be capable, within twenty years, of doing any work a man can do.*
> *Herbert Simon (1956), quoted in Crevier (1993, p. 109)*

> *Within a generation [...] few compartments of intellect will remain outside the machine's realm – the problem of creating 'artificial intelligence' will substantially be solved.*
> *Marvin Minsky (1967), quoted in Crevier (1993, p. 109)*

However, the vision of actually putting together an artificial intellect was (and is) all but humble, and the initial hype eventually faded away. The early programs that seemed so astonishing could in many cases only operate at very trivial levels with well-defined framings in relation to the problems they set out to solve. There were several reasons for these disappointments, among others:

- The problem of intractability and the combinatorial explosion, that is the fact that many problems exhibit a logarithmical explosion of alternatives to evaluate, outnumbering the capacity of even a utopian super computer.
- Another problem was to be called Moravec's paradox. Contrary to earlier assumptions, the human ability for conscious, logic-rational reasoning requires comparatively little computation, whereas the unconscious sensomotoric skills and instincts require enormous computational resources.[3]

These two problems led to another main problem of AI: How does everyday reasoning function? The early approach was to regard human reasoning as high level, formal symbolic reasoning, but the failure of the first AI hype to model functional algorithms to solve complex problems gave way to theories of human reasoning in terms of unconscious, intuitive, emotional and embodied 'know how' rather than 'symbol processing'.

The first AI winter and the second hype

By the mid 70s, the generous undirectional governmental (and military) funding in both the UK and the USA were drastically cut due to disappointment with the meagre outcome as to functional systems. This was the first 'AI winter'.

One of the branches of AI that survived during coming years was *expert systems* (programs that simulated the knowledge and analytical skills of one or more human experts). In the early 80s, expert systems went into commercial success and

[3] In Moravec's own words: '*[I]t is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility.*' (Moravec, 1988, p. 15).

AI research experienced a revival. However, as an ironic joke about the difference between the rational (AI) and the emotional (humans) – AI research was caught in a second spiralling hype of unrealistic expectations.

The second AI winter and AI after the mid-90s

By the end of the 80s the limitations of expert systems started to show. Even though an expert system might, for example, be more accurate in diagnosing than most physicians, it could suddenly miss a diagnosis by miles if a situation was not foreseen and implemented. Part of the explanation for this relates to the problem of common sense reasoning, which remains one of the persistent AI problems. An expert system can only have knowledge of a relative limited and specified world, which makes it very vulnerable to situations that touch or expand the border of the defined knowledge domain (i.e. the *Frame Problem*). These knowledge constraints also makes it very difficult for an expert system to make holistic common sense judgements whether – as in the example above – a diagnosis is plausible or not. [4]

Fair or not, once again hyped expectations ended in drastic cuts in funding, resulting in a second and longer AI winter – and still today AI, as an academic field of its own, suffers from this disrepute.

At the same time, behind the scenes AI has attained broad success with powerful algorithms that solve difficult computational problems in other areas of technology such as data mining, industrial robotics, regulatory systems ... and not the least computer games.[5] Probably the major consequence of the second AI winter was that AI was fragmented and joined other research fields where it continued under new names like knowledge-based systems, computational intelligence, and intelligent agents. New multidisciplinary research fields also emerged including intelligent virtual agents, affective computing, and educational technologies, that all have a bearing on this thesis. The next chapter deals with the definition of intelligent virtual agents and the somewhat confusing terminology in the field(s).

---

[4] For example: *MYCIN* (Stanford University) was an early medical expert system for treating blood infections. It performed roughly equal to human specialists in blood infections and better than general physicians. However, if it were told that a patient (who had received a gunshot wound) was bleeding to death, the program would attempt to diagnose a bacterial cause for the patient's symptoms (EBO, 2009).

[5] To quote a colleague (Björn Breidegard, Certec, Lund University): '*Artificial intelligence? You know, everyone is fiddling with it – but no one admits it*.'

Chapter 2:
From Intelligent Agents to Embodied Pedagogical Characters

: - §

The 'Intelligent Agent'

First of all, keep in mind that we are basically dealing with computer algorithms (i.e. program code that generates experiences of simulated virtual beings). Secondly, we need to disentangle the fundamental concept of an 'agent'.

Agents

The notion of an 'agent' may take on a multitude of interpretations depending on the context – but let's start with a dictionary definition.

> **1:** one that acts or exerts power
> **2 a:** something that produces or is capable of producing an effect : an active or efficient cause b: a chemically, physically, or biologically active principle
> **3:** a means or instrument by which a guiding intelligence achieves a result
> **4:** one who is authorized to act for or in the place of another: as **a:** a representative, emissary, or official of a government <crown *agent*> <federal *agent*> **b:** one engaged in undercover activities (as espionage) : spy <secret *agent*> **c:** a business representative (as of an athlete or entertainer) <a theatrical *agent*>
> **5:** a computer application designed to automate certain tasks (as gathering information online)
> *Merriam-Webster Online Dictionary (2009)*

Of these standard definitions, the first two with their emphasis on action and effect are of particular interest: '1: one that acts or exerts power', and '2a: something that produces or is capable of producing an effect'.

In the domain of Computer Science, the notion of an agent (or *software agent*) on the other hand generally adheres to the last definition above: '5: a computer application designed to automate certain tasks'.

Combining these interpretations of an agent, this thesis defines 'an agent' as 'a software application capable of operating on its own and of performing action'.

Intelligent agents

A specific group of software agents are denoted 'intelligent agents' or 'autonomous agents'. This corresponds to software applications that 'behave' more or less autonomously and with some kind of decision capacity or intelligence. For example, computer viruses, worms, and other forms of computer malware may operate on their own with different strategies in order to infect computers (thereby causing trouble like their biological counterparts). Accordingly, an intelligent agent can be seen as a digital actor, operating in a digital environment – and though this notion of an intelligent agent is situated in a virtual, digital context, its actions and behaviours may affect and continue in the real world (like the effect of a computer virus causing a hard disk to crash.)

In this thesis, the concept of an intelligent agent is even more delimited with respect to the focus on computer generated, humanlike characters. The intelligent agent in this case represents an immaterial computer application with simulated humanlike, cognitive capabilities[6] such as perception, communication, planning, reasoning, learning, decision making, and emotions. To be able to implement such applications, one needs to engage solutions and strategies from the field of artificial intelligence.

Following this reasoning, an intelligent agent is defined as 'a software application with an ability to act on its own and to reproduce human behavioural and cognitive capacities by means of artificial intelligence'. However, (since the breakthrough of artificial intelligence, mimicking human abilities always seems to be predicted to come twenty years in the future) we will in practice be satisfied with all kinds of 'tricks of the trade' in order to provide an ever so faint experience of interaction with another living existence. The shorter term 'agent' consequently (in this thesis) also denotes a computer application that may behave as an intelligent agent, but without any actual artificial intelligence employed, in that the perceived intelligent behaviour of the agent is based on simple sets of predetermined conditions or similar strategies.

Summing up, the concept of an 'intelligent agent' is associated with or encompasses ideas of digital technology, virtuality, autonomity, intelligence, interaction,

[6] In this thesis, 'cognition' takes on the broad interpretation from cognitive science, encompassing a wide range of human aspects related to thinking and behaviour, not the least emotions. This is in contrast to cognitive psychology where cognition often is used to refer to the 'intellect' in contrast to motivation and emotion.

and living, but it does not necessarily encompass aspects such as embodiment, character, and visualization.

Embodiment

In Cognitive Science, theories of embodied cognition challenge the idea of the human brain as an isolated and independent intelligent entity. In short, embodied cognition proposes that the human mind is largely determined by the human body. All aspects of human cognition, such as reasoning, communicating, decision making, and perception have evolved under the constraints of the evolution of the human body and its capacities (e.g. Brooks 1999; Dourish, 2001). In this approach, the idea of an isolated brain is meaningless. Without access to perceptual channels and a body, the brain would in a short time deflate and eventually pass away.

In line with the theories of embodied cognition, a pure self-operating intelligent agent would be of limited interest for the scope of this thesis. A virtual character requires some means to communicate. An early and simple method was command line keyboard input together with screen text output, which was hardly successful for a broader public in terms of efficiency and smoothness. As humans, we are by evolution specialized to communicate with other humans by means of spoken language, facial expressions, body language, tactile/physical contact, and not the least, olfaction.[7] Furthermore, humans appear as a very egocentric species, eager to anthropomorphize both living creatures and lifeless objects. Consequently, an anthropocentrically oriented interaction with an intelligent agent could benefit from the potential richness and flexibility of human communicative strategies. Following this line of reasoning to its end, the optimal embodied intelligent agent is a robot not distinguishable from a human.

At least in the near future, much of this is still pure science fiction. What we can do is to render animated graphical representations of human beings, creatures (or whatever) on a display ... or build spectacular robots for the entertainment industry. By this we access some of the potential power of facial expressions and body language in addition to the more immaterial qualities of spoken language, such as a face in which to embed the spoken message with a smile or subtle wrinkles around the eyes – the ever so important context.

Doing so, we may obtain a visual intelligent agent with the communicative potential power of embodiment. However, human communication is not spoken language and gestures alone – we also make heavy use of mental conceptualiza-

[7] Some theories of human cognitive evolution emphasize the development of social interaction and language as main driving forces (e.g. Whiten & Byrne, 1997).

tions like preconceptions, stereotypes, and social constructions in order to situate, contextualize, and facilitate communication (even if this is a double edged sword). An embodied intelligent agent will automatically be loaded from the start with ideas, preconceptions, and stereotypes invoked by its visual body as well as its voice (see also Part II, Chapter 8).

To sum up, with the embodiment of the intelligent agent, we have the basic foundation of a 'new social actor' – *the embodied virtual character* – presented in the first lines of this thesis. Next, I present an outline of the diverging branches of embodied virtual characters and associated terminology.

Embodied Virtual Characters and Conceptual Confusion

Avatars, agents, bots (and mannequins)

Depending on the approach, two main categories of embodied virtual characters can be identified. One is the 'avatar', which is controlled by a human and used as her or his virtual representation in a digital arena (e.g. the characters used in online games like *World of Warcraft*).

The second main category is the character that can 'act' on its own and 'react' on user input/contextual input. This is the 'intelligent agent' (or 'autonomous agent') discussed above. In everyday use, the shorter form 'agent' is common when the experience of a computer generated (intelligent) agent is more in focus than the implementation of any actual 'intelligent' algorithms. The 'bot' is also included in this second category, and may be defined as an 'intelligent agent' situated in an Internet context. For example, there are embodied 'chat bots' that can answer questions and assist in other ways like the virtual assistant *Anna* on IKEAs web-site (see Figure 2). There are also the 'chat bots' that are designed to simulate intelligent conversation with humans

Outside the scope of this thesis are the 'virtual mannequins'. These are digital versions of mannequins used, for example, in computer supported ergonomic design of driver environments, computer simulated crash tests, virtual reality oriented medical education ... and they are (of course) now also used in fashion design.

Conversational agents

Conversational agents constitute a branch of intelligent agents designed to simulate conversation in natural language by means of artificial intelligence (with echoes of the old *Turing Test* of artificial intelligence, originally set up as a natural language test mediated by keyboard and screen interaction.)

Embodied conversational agents constitute one of the dominating fields of embodied intelligent agent research, which also appropriated an abbreviation of its own: *ECAs*.

From the perspective of the anthropocentric embodied intelligent agent domain, natural language involves the embodied and contextualized aspects of facial expressions and gestures. In line with this, the overall goal of the research on embodied conversational agents is a smooth and natural interaction. However, while natural language generation has made quite impressive progress when it comes to synthetic speech – natural language understanding remains bothersome and the general input method is still the keyboard.

Affective & relational agents

Affective agents

Another vivid research domain overlapping the domain of intelligent agents is affective computing (Picard, 1995), which focuses on technologies that can recognize, evaluate, respond to, and generate emotionally related information. For example, there are systems that collect and analyze physiological data like heart rate frequencies, eye movements, and voice characteristics. These data are then evaluated based on an emotional model of a human (bored, confused, engaged, etc.), upon which an appropriate system response is calculated.

Such systems belong to the field of artificial intelligence under the name of 'affective systems'. Consequently the corresponding agents are categorized as intelligent agents under the name of 'affective agents' (i.e. intelligent agents that can detect and react upon a user's emotional state and behaviour). Adding a body gives us embodied affective agents that are able to express humanlike emotions by means of a happy smile or eyes filled with sorrow.

Relational agents

Relational agents make up a subgroup of affective agents (Bickmore & Cassell, 2001). These are intelligent agents designed to develop and maintain long time socio-emotional relations with users. They can save memories (data) of past interaction sequences with the user which then can be processed in combinations with computational models of long time humanlike relational processes.

Pedagogical agents

This category – which also is the focus of this thesis – addresses intelligent agents that have a pedagogical agenda. Historically, they originate from a long tradition of intelligent tutoring systems (ITS) emerging during the 1960s (c.f. Chapter 3).

Since then the power and complexity of tutoring computational models have evolved, paralleled by technological advances in computational power. Given this development, along with the preceding survey on intelligent agents and embodiment, it is no surprise that pedagogical systems during the 1990s started to explore embodiment and the potential of social and relational contexts. This development has also been influenced by the growing awareness of socio-cultural aspects in pedagogy, something that will be further discussed in Chapter 4.

In this thesis, the focus is on embodied intelligent agents in pedagogical settings. They will be denoted *embodied pedagogical agents (or characters)*, with the term 'pedagogical' being more constrained and specific then the broader term 'intelligent'.

To finish this terminological review which (with some deviations) has led us to embodied pedagogical agents, one can in a more straightforward manner define them as 'visually represented, computer generated characters in pedagogical roles, such as virtual instructors, mentors and learning companions, encountered in different digital environments' (c.f. *Paper III: Haake & Gulz, 2009, p. 39*).

Embodied virtual characters in general

As indicated, virtual characters (in the sense of computer generated, visible, more or less humanlike characters) have become an increasingly important ingredient of virtual environments. In order to better understand these novel arenas, the terminology reviewed above can help to:

i) distinguish between avatars (that are controlled by a user) and agents (that are controlled by a system),
ii) identify and distinguish between conversational, affective, relational, and pedagogical agents, and
iii) recognize 'embodiment' as an expansion to intelligent agent systems that adds the potential of embodied human social communicative strategies.

Additionally, one should mention that the notation 'agent' more often is used when one wants to emphasize the intelligence and autonomy of an (embodied) virtual character. Correspondingly, the notation 'character' is more often seen when focusing on the embodiment. The notion 'character' is furthermore often used in a more comprehensive sense, encompassing all the diverging variations of visually represented, computer generated, humanlike artefacts (intelligent or not).

Turning to the broader literature, one may encounter notions such as: animated agent/character, synthetic agent/character, interface agent/character, etc. In accordance with the preceding classifications, these notations are – more or less deliberately – used to emphasize certain attributes. Thus, an 'animated character'

may be used to emphasize the qualitative aspects of animated characters, with a focus on algorithms that generate gestures and movements in real time. This is not to be confused with an 'animated character' in, for instance, a Pixar/Disney movie. These latter characters are pre-rendered animations belonging to a fix story, with no interaction, flexibility or intelligence whatsoever. However, such characters, pre-rendered into sets of alternative, optional, combinable sequences, may comprise the underlying embodiment of intelligent agents. They can also be used as test dummies in order to examine, for example, the social and communicative effects in the interaction between humans and embodied characters in the role of artificial social actors.

Intelligent Agents vs. Embodied Pedagogical Characters

As has been described, there is a multitude of different computer based intelligent humanlike artefacts, ranging over a vast spectrum. To wrap up this chapter, I will recapitulate what distinguishes them from each other – and what unifies them.

- *Complexity:* How complex and flexible are the algorithms and models that control and/or generate the behaviour?
- *Real time generation:* Does the system use pre-rendered materials, and in what ways?
- *Humanlikeness:* How humanlike is the appearance, movements, gestures, and behaviour?
- *Input:* What kind of input is used (keyboard text input, spoken language and voice characteristics, visual signals such as facial expressions and gestures, physiological data, etc.)? How flexible and complex are the algorithms and models that recognize and evaluate this input?
- *Output:* What kind of output is used (text on screen, voice and sounds, facial expressions, body language, etc.)? How complex and rich is this output?
- *Interaction:* How rich and advanced is the interaction (communication) between human and agent? That is to say, what potentials and constraints are there for any mutual influence regarding the dialogue and behaviour of the user and the agent respectively?

A Final Reminder ...

Writing and reading about embodied intelligent agents may prophesy a fantastic and impressive as well as a terrifying and scary future, but as far as artificial intelligence is concerned, we are far from being able to simulate or understand human cognition in its broad sense.

... or Two?

Embodied pedagogical agents are not going to replace teachers, but hopefully help them to personalize the learning during the everyday activities.

Chapter 3: Artificial Intelligence and Learning

: – ¤

Artificial Intelligence and Learning – an Historical Résumé

The use of technology to support or assist pedagogical activities is probably as old as pedagogy itself – at least if the use of a stick to punish lazy youngsters can be seen as a use of technology. A more meaningful example is the use of primitive tools to preserve and recapitulate objects and ideas as graphical pictograms and ideograms. Later on in history, the gradual evolution of modern writing systems and the technique of using parchment and paper are indeed cases of technology supported learning – followed by the invention of the printing press. In modern history, it is interesting to note how new technologies immediately generate new ideas of technology supported learning. A splendid example is the granting of radio broadcasting licenses to colleges, universities, and school boards by the US federal government from the 1920s to the 1940s. The venture ended in a total failure only to be carried on to the new media of television (Nasseh, 1997).

The present-day, rapid development in information technology has consequently generated lively activity involving research and development around technology supported learning. Comprehensive notions like *e-Learning* (*electronic Learning*) and *Educational Technology* signify broad fields that embrace all kinds of digitally computer related ideas and applications connected to pedagogic issues. When it comes to this thesis, the main focus is on the branch of artificial intelligence known as *Intelligent Tutoring Systems* (*ITS*).

Intelligent tutoring systems (ITS)[8]

Among ideas and applications following the first boom of artificial intelligence in the 1960s was that of intelligent machine assisted learning. As a branch of

[8] This part on Intelligent Tutoring Systems (ITS) builds on Haake (2006) and Crevier (1993).

artificial intelligence, it has followed the general progress and ups and downs in this field (c.f. Chapter 1). An initial approach was to provide students with a personal and individualized *intelligent tutoring system* (*ITS*), constructed around 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.

The early systems that emerged during the 1960s and early 1970s were relatively simple, and the main innovative power was perhaps the development of user interaction techniques such as video screens, keyboards, and speech synthesizing. The intelligence was based upon domain specific data, organized in databases with relatively simple rules for the interaction, prompting the student through static tutoring plans. The learning situations were thus much on the terms of the systems, and eventually the general problems in artificial intelligence to progress beyond the initial achievements also affected the research and development of intelligent tutoring systems.

Following the revival of artificial intelligence with the expansion of expert systems during the 1980s, the intelligent tutoring systems took a leap forward with the development of computational models for the user/student and pedagogic strategies. Then again, as the problems of further expanding the models of human cognitive capacities into more advanced and complex models were severely underestimated, research and development within intelligent tutoring systems levelled out.

After this, the AI branch of intelligent tutoring systems transformed under influences from other fields such as pedagogy and human-computer interaction. Today, with the expanding potential of graphical user interfaces and video graphics, the old idea of intelligent tutoring systems is a part of the novel and vivid research fields of *e-Learning* and *Educational Technology* mentioned above.

Present-day Artificial Intelligence in Learning

A cognitive approach to artificial intelligence and education

As suggested above, the early rational track of artificial intelligence, inspired by theories of modern formal logic, could not provide any substantial computational models of human cognition or human learning. This disappointment eventually brought the rational paradigm and the idea of making artificial 'human' intelligence into disrepute.

Since then, there has been a growth in theories of embodied cognition and an increasing awareness of affective and emotional aspects of human cognition. Today, a more multifaceted picture has emerged of a gradually evolved human cognition, situated in a complex and dynamic reciprocal system of brain and body, as well as a surrounding world and culture. This view of cognition as inherently embodied and situated is far from the earlier formal and rational traditions in science (see Chapter 5).

Artificial intelligence in education

Concerning artificial intelligence, these new perspectives of embodied and situated cognition have led to new approaches manifested in computational fields such as affective computing. Here we also find the technological roots of embodied virtual characters in their guise as social actors in digital learning systems.

Chapter 4:
Technology Supported Social Learning

: - 9

Social Learning

In present-day pedagogy, a view on knowledge as socially constructed is integrated in central pedagogical theories. This approach emphasizes the processes of teaching as well as of learning as inherently social, involving interaction, communication, negotiating, and sharing (Gulz, 2004; Gulz & Haake, 2009)[9].

The obvious contrasting example is the idea of a prototypical learner as a student in a study chamber, employed in an individual intellectual pursuit of reading and writing. However, for the everyday student, learning is probably better described as a dynamic alternating process of socially oriented learning activities (lectures, group assignments, co-studying, etc.) together with periods of more individual endeavours of reading and reflection.

An important aspect of social learning is the recognition of the physical place (e.g. classrooms, group rooms, and cafés) as an arena that both initiates and supports social activities.

Lessons from Distance Learning

While technology can be designed to support social group activities like multiple user workplaces (Dillenbourg, 1999), the development of distance learning technologies over the Internet has received much attention. The most simple form is more or less a transcription of the older mail correspondence courses based on written instructions and assignments passing back and forth. A well-known concern with this basic form of distance learning is the relatively high percentage

[9] This is neither a new nor anarchistic idea. The antique Academy of Plato was in much built around the philosophical dialogue, which truly is a social approach to learning (EBO, 2009).

of students that sooner or later drop out, never finishing the course (Tyler-Smith, 2006). With the exception of the obvious unrealistic expectation of being able to set off time for home studies while working and raising a family, the absence of a social context in the actual learning situation may well be an important factor. It can be argued that social relations in the form of teachers and fellow students provide a social framing that scaffolds continued studying even when it becomes quite demanding. One can here speculate about rewarding social affordances like 'studying together with fellow students' and 'getting help with difficult parts' as well as avoiding social distress like 'not disappointing the teacher' and 'not being excluded from the community of fellow students'.

Educational Technology and Social Learning

Educational technologies are only tools, but they are flexible ones. This potential to be flexible and the awareness of social dimensions and emotional factors in learning have led to a strong attention on technological supports that situate learning in social contexts. Scanning the fields of Educational Technology and e-Learning brings forth themes such as: social learning in multimedia, enabling social presence in virtual environments, personalized web companions, computer-supported social interaction, character enhanced systems (Gulz, 2004).

To address social learning contexts in educational technologies, two main strategies can be distinguished (*ibid.*). In *extrinsically social learning systems*, the role of the system is to provide room and support for social activities, which are created and supplied by the learners (e.g. distance learning systems with discussion forums and online chats). The contrast is *intrinsically social learning systems*, where social interaction and context is provided within the system by means of social arenas populated with social actors, for example *The Exercise Advisor Laura* (Bickmore, 2003) and *FearNot* (VITEC/eCircus).

Extrinsically social systems are in their basic form relatively easy to develop as they can be built on standard solutions for mail services, discussion forums and online chats. Intrinsically social systems, on the other hand, must deal with social interaction in other ways, which often means advances algorithms and complex architectures by means of artificial intelligence.

Another important difference between these two kinds of systems that must not be forgotten regards the single user. The extrinsically social system requires at least two users for social interaction to take place, whereas the intrinsically social learning system in itself can engage a single, individual user in social interactivity with a single, individual user.

The potential of Intrinsically Social Learning Systems

Both extrinsically and intrinsically social environments have their potentials as well as problems, and one can of course consider systems employing both strategies to a greater or lesser extent. The following discussion focuses on the potential of intrinsically social learning environments that support social interaction by means of embodied pedagogical agents.

Since intrinsically social learning systems have the specific quality to engage a single, individual user in social interaction, it is possible to adapt such a system to the changing needs of an individual student. The system can thus be tailored with regard to teaching strategies (troublemaker, challenger, co-operator), roles (coach, instructor, companion), and so on. With regard to the daily classroom situation, the social prerequisites as to classmates and teachers are relatively constrained. However diligent a teacher may be, it is simply impossible to meet every individual student based on that student's personal needs as to learning style and need for personal attention. From this perspective, the intrinsically social learning system can be a complement to a teacher with the potential to engage in dynamically tailored learning without any restrictions as to time.

Character enhanced intrinsically social learning systems

Proposed benefits of character enhancement

The possible benefits of character enhanced intrinsically social learning systems are all but systematically explored and the results are often ambiguous. With this reservation, there are several proposed advantages and Gulz (2004) lists and discusses six kinds of benefits that can be seen as central for character enhancement. They are summarized below with some updates and will be followed by a discussion as to empirical evidence:

i) *Increased motivation:* Character enhancement may prompt students to stay on and involve themselves in a learning environment by means of motivation. Central factors that may increase motivation are experiences of entertainment, likeability, and engagement. Other factors often related to motivation are the lifelikeness of the character, the 'Persona Effect'[10], and the ability of the character to show and elicit emotions (e.g. Lester et al., 1997; Moreno 2004; Moundridou & Virvou, 2002).

[10] The 'Persona Effect': That the presence of a lifelike character can have strong (positive) effects on students' perception of their learning experience (Lester et al., 1997)).

ii) *Increased sense of ease and comfort:* The addition of social characters may have a positive relaxing effect on the student, making her or him feel more comfortable and more at ease with the learning tasks and the learning environment (e.g. Moundridou & Virvou, 2002; van Mulken et al., 1998).

iii) *Stimulation of essential learning behaviours:* Learning can be described as the employment of different basic strategies and the presence of a social actor may have a positive stimulating effect as to exploration, cooperation and reflection (e.g. Blair et al., 2006; Johnson et al., 2003). Attention, as elusive as it may be, is an important and central concept in learning. From the perspective of social learning, character enhancement may have the potential to capture and increase, as well as guide the attention of the learner (e.g. André et al., 1998; Rickel & Johnson, 2000).

iv) *Enhanced flow of information and communication:* Building on the idea that humans are evolutionarily fine tuned with regard to social interaction with other humans, a social actor with human-like communicative capabilities may pave the way for a smooth, rich and efficient flow of information. A central aspect of this approach is the importance of facial expressions and body language to reinforce, clarify, and consolidate the spoken dialogue as well as to provide feedback (e.g. Cassell & Thorisson, 1999; Massaro et al., 2000; Oviatt & Adams, 2000).

v) *Gains in terms of memory, problem solving and understanding:* Positive effects in terms of improved memory, problem solving, knowledge transfer, and understanding may follow from character enhancement (e.g. Blair et al., 2006; Johnson et al., 2003; Moreno et al., 2001).

vi) *Fulfilling the need for deeper personal relationships in learning:* Finally, in addition to prompting, encouraging, and scaffolding different behaviours and experiences as to learning, a social actor may also promote experiences of deeper personal relations. In the form of relational agents (see pages 13) it is possible to establish and even maintain long-term qualitative personal relationships between a student and a social learning system (e.g. Bickmore & Picard, 2003; Moreno et al., 2001; Veletsianos & Miller, 2008).

Empirical support for the benefits of character enhancement

When it comes to empirical support for the listed benefits of character enhancement, there is a general problem in the absence of long-term studies.[11] The absolute majority of studies are short-term, often evaluated by means of quantitative

[11] Among important exceptions one can find Bickmore (2003) and Veletsianos & Miller (2008).

measures. The quality of these studies is often problematic as to what is actually measured in combination with problematic interference effects between different benefits and short-term effects of novelty (e.g.'Hawthorne effects')[12] (Dehn & van Mulken, 2000; Gulz, 2004).

Another problem in the empirical outcomes is the omission to examine the distribution of positive or negative effects in relation to distinct groups. Instead there is a tendency to merge the test group into an averaged user (see Chapter 5). More explicitly, by correlating the examined group as to communicative style, pedagogical style, gender, etc., one can find one group showing positive effects and appreciation as to social interactive learning and one group with negative responses and irritation as to the presence of social actors (e.g. Bickmore, 2003; *Paper III: Haake & Gulz, 2009*). These findings will be further discussed in Chapter 8 where the papers included in this thesis this thesis are presented.

Finally, the empirical outcomes (even though ambiguous) support an optimistic approach to character enhancement in intrinsically social learning systems. The results do not suggest that one should think of character enhancement as a general solution for social learning – but as a support for certain students where the potential positive effects of social characters relate to the individual learner and the specific learning situation. This also implies the important potential of computer based pedagogical systems: The same learning material can be presented in different ways, with or without social context, and – in the case of social character – with characters differing and tailored to different needs and situations.

Summary: Technology Supported Social Learning

There is much in favour for the endeavour to develop technology supported social learning systems. When it comes to the prospect of actually developing systems capable of inducing the possible benefits, one should remember that we are far from being able to simulate or understand human cognitive and communicative capabilities.

[12] The 'Hawthorne effect': The Hawthorn study sought to identify those aspects of a job that were most likely to boost worker productivity. At the study's onset, it was thought that economic factors would have the greatest influence on productivity. The results were surprising: productivity increased, but for reasons unrelated to economics. Ultimately, researchers concluded that job performance improved because more attention [i.e. the study (*author's comment*)] was being paid to the workers (EBO, 2009).

At the same time, social learning and character enhancement in pedagogical systems is not merely academic research. There have been several more or less advanced commercial systems around for a good while. Thus, considering the severe knowledge gap as to long-term use of character enhanced social learning systems, it is not necessarily due to the underlying technology – but rather to an unfortunate lack of systematic studies.

Chapter 5: The Troublesome Concept of Rationality

The concept of rationality is as old as science itself. Basically, any field of modern academic science is bound to articulate its knowledge and research by means of rational (logical) coherent reasoning.

At the same time, the phenomena of study in modern science (except for some disciplines in natural sciences) are seldom accessible and describable as strict logic-rational systems. At best they may subdue to statistical probabilities, but often enough the articulation of knowledge resides within multiple and even ambiguous argumentative and descriptive perspectives.

To me, this implies a scientific paradox: Science (as a whole) is bound to articulate its knowledge by means of rational reasoning, though the phenomena of study seldom will fall apart into rational describable models.

About Rationality

The domain of embodied virtual characters is intimately connected to computer science and artificial intelligence and, accordingly, the domain as a whole is computationally oriented. This is only natural, as there would be no embodied agents without the continuous development of even more complex and sophisticated algorithms in addition to the constant increase in computational power.

At the same time, the domain of embodied virtual characters is a cross-scientific field, combining artificial intelligence with fields like cognitive science, social psychology, linguistics, neurobiology, and graphic design. This inevitably leads to a variety of approaches and – perhaps more troublesome – diverging views on the scientific issue of rationality.

A short history of rationalism

In Western science, the Greek heritage has imposed the pervasive idea that all phenomena ultimately are addressable by rational (logical) reasoning. For instance, Aristotle's work on logic (syllogism) and his argument that humans are 'rational animals' have had a tremendous impact. This rationalistic aspiration has manifested itself throughout history in endeavours to access by reason the world, the human mind, and the existence of God (e.g. philosophical rationalists[13] like Descartes, Leibniz, and Kant). Omnipotent as well as impressive, these projects were in the end only mental constructs – but they have continued to nourish the philosophical-scientific community of the 20th century. Influential philosophers of logic and mathematics (e.g. Frege, Russell, Wittgenstein, and Gödel) made groundbreaking contributions in the rationalistic traditions and set the foundation of modern formal logic and programming (c.f. Chapter 1).

At the doorstep to the second half of the 20th century, science as a practice seemingly relied on a firm and stable ground of rational reasoning and experimentation.[14] Soon, however, postmodernistic critique and deconstruction inundated the scientific community, digging trenches between scientific disciplines.

The non-rational, rationalizing human

Eventually, as the problems and set-backs of the rational analytical approaches to human cognition in fields like artificial intelligence (c.f. Chapter 1) and cognitive psychology (see below) have piled up, the old scientific assumption of the 'rational human' can hardly be upheld in science today.[15] One way to deal with this dilemma is concisely expressed in a quotation from the science fiction writer Robert Heinlein, '*Man is not a rational animal; he is a rationalizing animal.*' (Heinlein, 1953, 14). This approach can be described as follows:

1) Humans have no dedicated logical-rational cognitive processing modules in the brain. That is to say, humans do not process (logical) information in ways similar to electronic logical circuits.
2) Humans can behave in a logical-rational manner, given certain conditions.

[13] Rationalism: the philosophical view that regards reason as the chief source and test of knowledge – holding that reality itself has an inherently logical structure (EBO, 2009).

[14] However, already in the first half of the 20th century, the inherent limitations of formal logic and mathematics had been demonstrated by, for example, Gödel (*incompleteness theorems*) and Ramsey, Church and Turing (*undecidable problems*) (EBO, 2009).

[15] An interesting argument for the non-rational human is given in Johansson et al. (2005) where they cleverly exposed the human feebleness for introspectively derived reasons when failing to detect mismatches between intention and outcome, an effect they call choice blindness.

The first argument (1) relies, among others, on recent theories of human evolution and embodied and situated cognition (Clark, 1999; Lindblom & Ziemke, 2003). The second argument (2) is easily exemplified with everyday mathematical problem solving.

The benefit of this alternative view of human cognition is that it may open up new perspectives and approaches, especially in research disciplines constrained by explicit or implicit assumptions of human rationality.

Rationalism and artificial intelligence

Regarding artificial intelligence, the field was born in the positivistic academic atmosphere of the mid 20th century. From the beginning it was tightly connected to the new branch of cognitive psychology, which was approaching human cognition by means of implicit rational model building – such as the persistent hypothesis that human decision-making could be accessed by formal models (Sahlin, Wallin & Persson, 2009).

Thus, given the academic society of the time, it is no wonder that the belief in the potential of formal logic rationality and human rationality was so strong at the onset of artificial intelligence.

> *It is not my aim to surprise you – but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, [...] the range of problems they can handle will be coextensive with the range to which the human mind has been applied.*
> *Herbert Simon (1957), quoted in Crevier (1993).*

As argued in earlier passages of this thesis, many aspects of human cognition are not accessible or solvable by formal logic. At the same time, the inherent nature of computation is a logical manipulation of formal symbols. For embodied virtual characters as a research field, this implies a somewhat paradoxical situation of non-rational rationality.

The Rationalistic Approach and Its Consequences

Taking a closer look at the field of embodied virtual characters, let us propose a hypothetical rationalistic agenda – (i) and (ii) below – and then explore the consequences of this agenda with respect to applied artificial intelligence in education (c.f. *Paper I: Gulz & Haake, 2006a*).

i) The quest for the 'perfect' artificial human.
ii) The concept of 'the user'.

The quest for the 'perfect' artificial human

Following a long tradition of philosophy and science, one of the ambitions of artificial intelligence is to uncover human cognition with the overriding aim of constructing a 'perfect' artificial replication of a human being (i.e. an embodied agent and eventually a robot or cyborg). This is science *per se*, striving for knowledge for the sake of knowledge itself.

In present artificial intelligence there may be a dilemma as to the possible implicit concept of the 'rational human'. In a field this focused on rational computational models of human cognition – there is an augmented hazard with regard to the scientific (and human) tendency to mistake the map for the territory.

The concept of 'the user'

In order to enable human-agent interaction, there is a need to implement computational models of the intended human user. Such user models are often *generalized* models of human cognition. Real humans, on the other hand, exhibit an *individual diversity* with respect to cognitive and emotional behaviours and an astonishing flexibility as to communication. To engage in social interaction with other humans means to guess, negotiate, reconsider, fill in, make mistakes, confirm, correct, etc. (Hansson et al., 1979).

Once again, there is an apparent risk to mistake the map for the territory. The everyday case with single (general) user models may transfer to the evaluation situation and a treatment of test groups as generalized single humans. As a consequence, evaluations risk missing significant correlations as to, for example, cognitive sub-groups of pedagogical and communicative preferences (e.g. *Paper III: Haake & Gulz, 2009*).

Applied artificial intelligence in education.

The two discussed dilemmas could be regarded as internal issues, but they are not. The field of educational technology is directed towards functional and efficient applications and systems. What if the agendas of (i) and (ii) above will not do the trick? What opportunities are there today or in the near future to implement functional and efficient applications of, for instance, embodied pedagogical agent systems?

Another concern is the view of contextual phenomena when employing an embodied pedagogical agent to interact with a human learner. A cartoonish rendering of an embodied pedagogical agent with an emphasized, caricatured personality and pedagogical profile would be far from a realistic simulation of a real human being. On the other hand it may prove very efficient for interactive social

learning (c.f. Chapter 8 and *Paper I: Gulz & Haake 2006a*, *Paper III: Haake & Gulz, 2009*).

Both these concerns imply a shift from the agenda to simulate 'real' human communication, towards a focus on algorithms that may scaffold and even augment human-agent communication. Actually, a contextual perspective of this kind may have even more dramatic consequences than a change as to the computational problems of interest. It may even (in some cases) erase the need for new computational approaches all together.

The computational trap ...

The development of a computational algorithm to control the smile in a virtual face can easily diverge into increasingly more complex computational real-time models. An alternative approach is to turn to social psychology that can provide evidence that modifications of a face itself (as to gender, age, attractiveness, etc.) may totally change or overrule the effect and quality of the smile algorithm. The task at hand will then shift from (i) a development of a computational algorithm to control the transition of vertex points in a mesh to (ii) a matter of graphic design to model 'by hand' a face that is appropriate as to holistic and contextual aspects (e.g. support clear human facial expressions, handle socio-cultural impact on visual design, experiment with different graphic and artistic interpretations). Eventually, there may not even be a need for a new and better computational algorithm.

... and cultural aspects

Different perspectives regarding computational (rational) approaches and cultural aspects of human-agent interaction is another topic to discuss. The effect of appearance and voice on the cognitive processes may not be readily accepted, but is nevertheless empirically well established. To pick an example, there is much research that supports the halo effect (i.e. that good-looking people are attributed other positive traits such as being sociable, intelligent, and interesting, whereas unattractive people are considered less socially competent, dishonest, and psychologically unstable) (Brigham, 1980; Langlois et al., 2000). Similarly, socio-cultural effects also adhere to voices. In their book, *Wired for Speech*, Clifford Nass and Scott Brave (2005) convincingly demonstrate how cultural constructed conceptions relate to voice characteristics such as gender, dialects, and sociolects.

Rationality and Embodied Pedagogical Agents

Within modern academic science, as a common societal endeavour, the articulation of knowledge and research by means of rational coherent reasoning can be regarded as a fundamental prerequisite. Also a visually oriented domain as design (in its academic form) eventually aligns with the rational discourse to be able to discuss and argue around design objects, demonstrators, and design concepts.

Rationality is also an insidious concept. The necessity to align with rational reasoning in the explicit articulation of knowledge may easily be confused with an implicit idea of a rational universe, where all phenomena are inherently understandable and describable by means of rational reasoning.

The lesson (to easily forgotten) is: '*Never mistake the map for the territory*'. Science is model building – the ultimate 'Truth' is something else.

The practical consequences of this perspective on 'rationality' with regard to the field of embodied pedagogical agents, can be summarized as follows: The research on embodied pedagogical agents, as a cross-scientific field, has a built-in tension between: (i) implicit rational disciplines like artificial intelligence, (ii) relativizing and constructivistic socio-cultural fields of social interaction and communication, and (iii) non-academic practices like graphic design. If asking for trouble, it is there to be found – but from another perspective there are magnificent challenges to deal with.

Finally, it is not a question of right and wrong. All perspectives have their strengths as well as shortcomings and complement each other. It's a question of proportions ... and watching out for the rational ghost.

Part II

Embodied Pedagogical Agents: From Visual Impact to Pedagogical Implications

Part II presents a summary of the five papers appended to this thesis, with a main focus on the visual impact of embodied pedagogical agents with regard to pedagogical implications. Concerning details and in depth reasoning, I refer to the appended papers which represent the actual scientific work behind this thesis.

I also want to add a personal remark. The research presented in this second part is an inherently joint endeavour by me and my wife and colleague, Agneta Gulz, covering some seven years of a variety of projects and studies. We have also had much valuable help of several master students.

Chapter 6:
A Graphical Design Space of Embodied Pedagogical Agents

{ : -)

As argued in Part I (Chapter 5), more design-oriented aspects of embodiment such as appearance often fall outside the computational practice. Yet these are aspects that form our first impressions and influence the human-agent interaction.

In order to handle such aspects, I begin this second part of the thesis with an outline of a graphical design space of static visual characteristics of embodied pedagogical agents. (The version outlined here is a summary of the framework presented in Paper III.)

Visual Aspects of Embodied Pedagogical Agents

Dynamic and static visual characteristics

First of all, visual appearance can be divided into the two sub-categories of *dynamic visual characteristics* and *static visual characteristics*. The dynamic visual characteristics relate to animated expressions mediated by facial displays, body and hand gestures, postures, movements, etc. The static visual characteristics refer to visual aspects of the underlying character, for example, face and body shapes, hair, clothes, attributes, colours, and graphic style.

The problematic static visual characteristics

From a pedagogical perspective, there is reason to believe that the visual characteristics of an agent may considerably affect learner expectations, attitudes, understanding and motivation (c.f. Chapter 4). Similar influences are well known in related domains such as theatre, film animation, advertising, as well as in social psychology (e.g. Gard, 2000; Lassetter, 1987; Schneider, Hastorf & Ellsworth, 1979).

In the field of embodied agent research, the importance of dynamic (computational) visual aspects is generally agreed upon and much effort has been

directed to research and development with respect to facial expressions, gestures, and movements.

In contrast, the static visual characteristics have received relatively sparse attention, which may be due to their non-computational nature (c.f. pages 29-32). A consequence of this neglect is that static visual characteristics often are approached in a simplified manner that overlooks the complexity of visual aspects and eventually results in misconceptions and problematic over-generalizations.

With regard to this problematic neglect, the aim of this chapter is to present an outline of a graphical design space of static visual characteristics of embodied pedagogical agents.

(*A more elaborated discussion of these topics can be found in Chapter 8.*)

Design Dimensions of Static Visual Characteristics

The static visual characteristics constitute an immense and complex multidimensional design space that may be described as a set of more or less elusive, changing, context-dependent qualities. Such a perspective can hardly be submitted to any kind of rational, analytical deconstruction and the focus thus needs to be on high-level design aspects. In order to relate to this multidimensional design space, a comprehensive framework will be laid out with three basic high-level design considerations (with regard to static visual characteristics of embodied pedagogical agents): *basic model*, *physical properties* and *graphical style*.[1]

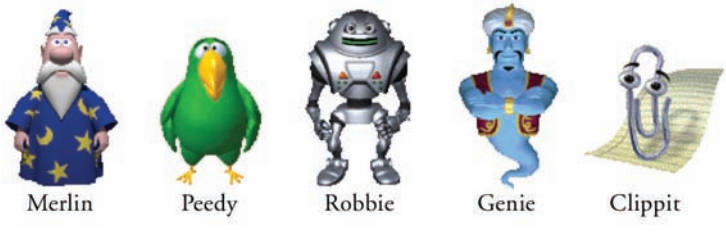


Figure 3. The characters from the *MS Agent Package* and the *MS Office Assistant* exemplifying the proposed basic model: *Merlin* (a human); *Peedy* (an animal/creature); *Robbie* (a combination of a human, inanimate parts, and a fantasy concept); *Genie* (a combination of a human and a fantasy concept); *Clippit* (an inanimate object).

[1] These design considerations are partly elaborated from the discussions of a more integrated design approach for virtual characters reported in Gratch et al. (2004) and further developed in Ruttkay et al. (2004).

Basic model

The first high-level design topic to consider is the *basic model* or constitution of a character (an embodied pedagogical agent), which can be described in relation to four basic conceptual entities: *a human*, *an animal or creature*, *an inanimate (non-living) object*, *a fantasy or fiction*, or *a combination of these entities* (Figure 3).

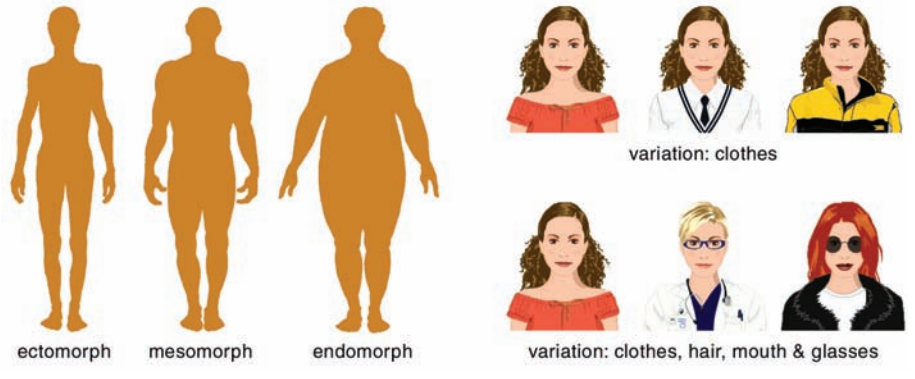


Figure 4. Examples of physical properties. Left: the three body types of *ectomorph*, *mesomorph*, and *endomorph* identified by Sheldon et al., (1940). Upper right: three different outfits transforming the experience of the one and same character as to personality, social position, education, etc. Lower right: variations of the same character by means of outfit, hair style, hair colour, use of lipstick, and use of glasses, dramatically changing the experience of the character. (The six characters to the right were put together by means of the *SitePal Demo Tool*, www.sitepal.com.)

Physical properties

The second high-level design topic addresses *physical properties* such as: body type, face shape, skin colour, hair cut, hair colour, clothes and accessories (Figure 4).

An important aspect of this second design topic is that the physical properties of an embodied pedagogical agent (as well as a human) always carry social, cultural, psychological and affective baggage, that is, there is no such thing as a *visually neutral* character (e.g. *Paper IV: Gulz, Ahlnér & Haake, 2007*; *Paper II: Haake & Gulz, 2008*; Isbister, 2006).

Graphical style

The two previous high-level design topics can to some extent be analytically approached and described. This is more difficult with the third high-level design topic of *graphical style* (Figure 5).

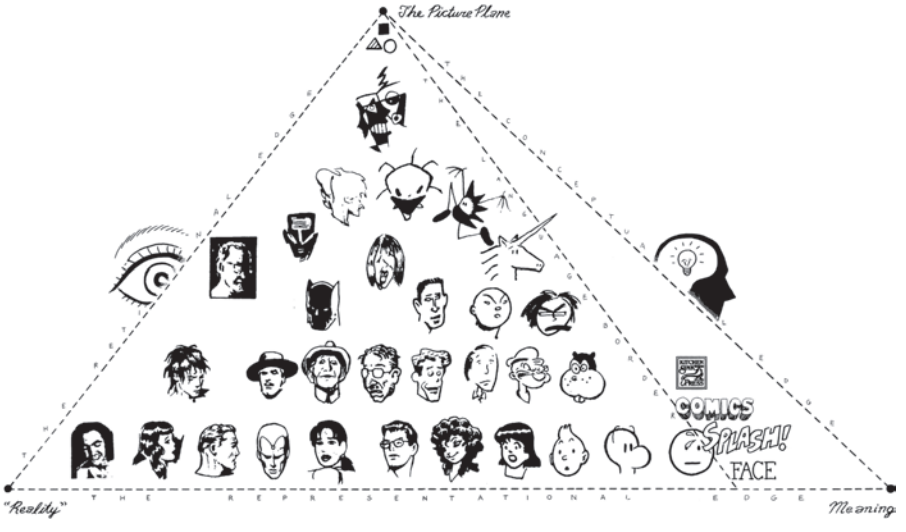


Figure 5. An edited reproduction of the design space of visual iconography (*The Big Triangle*) described by Scott McCloud in his book *Understanding Comics* (McCloud, 1993). While this is an ingenious way to set up the design space of graphical style, there is no predictable power to correlate a certain graphical style to a well-defined response.

Here one enters the full complexity of the graphical design space where small changes in, for example, the qualities of the line (Figure 6), shadings, and proportions may completely change the visual and cognitive experiences in diverse and not the least unpredictable directions (McCloud, 1993).

In spite of this, some aspects of graphical style may be both possible and worthwhile to explore. With regard to embodied pedagogical agents, there are at least two aspects of interest: *degree of detailedness* and *degree of naturalism*.

Graphical style: degree of detailedness

A regular photo of a face is highly detailed. By straightforward means of reduction, such a photo can be transformed into a contour line representation or a two level posterized representation of the very same face (Figure 7, left). In everyday design practice, however, a reduction of details usually goes hand in hand with more complex changes and modifications of the graphical expression or style (Figure 7, right).

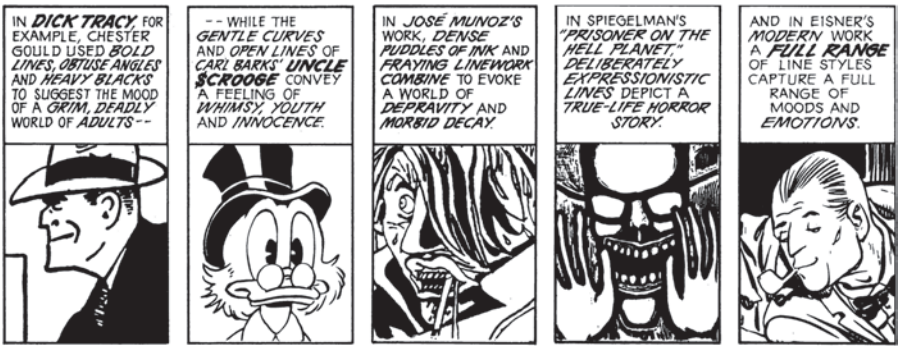


Figure 6. The inked line has almost unlimited possibilities with a potential of tremendous visual power (McCloud, 1993, p. 126).

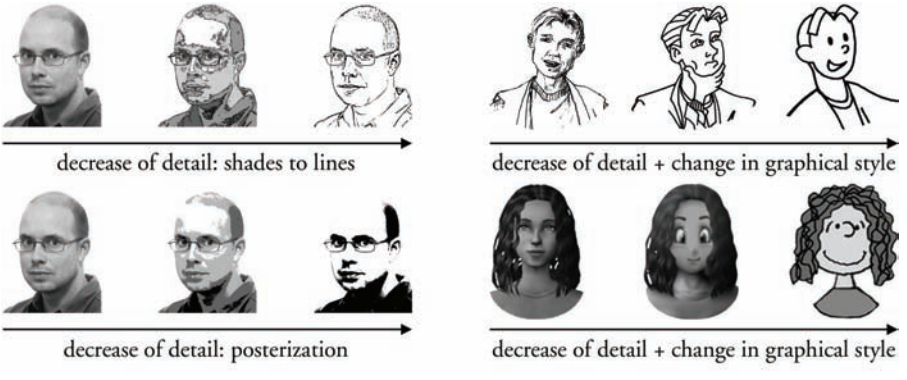


Figure 7. Graphical style: degree of detailedness. Left: two examples of straightforward reductions of details; Right: two more complex examples where reduction of details are combined with divergent (non-computational) changes as to graphical style.

The degree of detailedness affects cognitive processing. For example, the reduction of details can promote increased distinctness of facial expressions which may support a more rapid and accurate processing and interpretation (e.g. Cook, 1979; Isbister, 2006). This simplification may also (as discussed in more detail in Chapter 8), facilitate subjective self-identification.

Graphical style: naturalism vs. stylization

The same conceptual character, with respect to its basic model and physical properties, may be visually represented in numerous variations (Figure 8). By manipulating the qualities of line, shape and colour – each single variation (graphical style) may convey its own complex, dynamic, cultural and context dependent impact on the unique interpretation processes of the individual perceiver.

A constructive approach with regard to pedagogical effects can be to collapse the design space into a two dimensional map of *naturalism* and *stylization*. Naturalism here constitutes a well-defined endpoint of an otherwise immense and diverging design space of different stylized graphical expressions (Figure 8).

Figure 8. A design space of naturalism-stylization. The two figures to the left (the naturalistic endpoint) differ in degree of detailedness (c.f. Figure 7, left) but both are referred to as rather naturalistic. This design space is derived out of the pictorial plane (*the Big Triangle*) by Scott McCloud (1993) but is simplified in order to emphasize the dichotomy of naturalism vs. stylization (c.f. Figure 5).

The four stylized representations in Figure 8 above vary in expressive style inspired by: *Peanuts* (simplified, whimsy and humorous); *Dragonball* (cute, emotional and friendly); *Hernandez* (underground, rebellish); *Picasso* (abstract and intellectual).

Note: this simplified design space (Figure 8) does not convey any actual information about realative cognitive effects between different kinds of stylization – something that seems out of reach (at least from an academic perspective).

Chapter 7:
What About Realism?

: - ?

Relying on the framework of a graphical design space of embodied pedagogical agents just outlined – it is now time to approach the concept of visual realism.

(This abbreviated presentation is based on a discussion to be found in Paper III.)

A Definition of a Realistic (Human) Embodied Agent

Considering the embodied agent literature, the term 'visual realism' or 'visually realistic' is used in remarkably diverging and confusing ways. In an attempt to grasp a common concept of a *realistic (human-like) embodied agent*, I propose the following definition:

i) modelled upon the basic model of a 'prototypical' human
ii) adequate and relevant as to physical properties of a 'prototypical' human
iii) fully detailed (no reduction of detailedness)
iv) naturalistic (not graphically stylized).

The underlying problem is that many studies, when presenting their material and conclusions, neglect to specify what is actually meant by the term 'realism' (or 'realistic'). Sometimes the comparisons regard a character modelled upon a human versus a character modelled upon a fantasy concept, though both are perfectly naturalistic in relation to their basic models and could be equally realistic given a fictive context of a science fiction movie (Figure 9). On other occasions, the difference is in the degree of detailedness, yet combined with differences related to naturalism-stylization that are, however, neither recognized nor problematized. Despite such lack of deeper and more detailed analyses, the results are often generalized in terms of 'visual realism' versus 'visual non-realism'.

Figure 9. A study by van Vugt et al. (2007) compared three settings: non-character, realistic character (left), and un-realistic character (right). The hypothesis was that there would be effects of 'design-realism', but the effects of the two characters (above) did not differ. (In the framework of this thesis, the characters differ with regard to: basic model, clothes, and graphical style (soft rendering vs. facets).

Figure 10. The interactive information assistant Olga (developed by KTH, SU, SICS & Nordvis AB).

An illustrative example is reported by Gustavsson and Czarniawska (2004). At a conference, there was a discussion on the development of the interactive assistant *Olga* (Figure 10). The linguists in the development team opted for a more realistic Olga, since they insisted that in order for Olga's speech to be understood, Olga must be made as humanlike as possible. Olga's lip movements, in particular, had to correspond to those of a living human. The designers in the development team, on the other hand, were of the opinion that the comic styled Olga was actually more humanlike and attractive than the more realistic 3D-Olga, who they thought looked like Frankenstein's monster (an opinion shared by most of the conference audience).

The point of this example is that *none* of the Olgas (Figure 10) are realistic but instead constitute two different visually stylized representations that both could be referred to as:

i) modelled upon the same combination of a human and a fantasy concept of an alien from outer space,
ii) comparable as to physical properties (with the exception of the neck and the tie),
iii) rather comparable as to a low degree of detailedness,
iv) differences as to graphical style: 2D rendered, a bit squarish/angular shapes, some visual freedom as to postures *versus* 3D rendered, a bit softer/rounder shapes, constrained (stiff) as to postures.

This example indicates that the notion of realistic is (often) more of a non-reflected idea than an actual and explicit visual quality ('*If it's 3D, it's realistic!*').

The underlying problems

The underlying problems regarding the (mis)conception of realism are at least two:

i) The concept of realism is used without definition or specification (as discussed above).
ii) The confusion falls back upon a deeper problem: a widespread ignorance concerning 'deeper' aesthetic/graphical qualities and their relation to human experience (as seen in the Olga-example above).

In most studies, the inherent graphical qualities of the presented characters are not problematized. That is to say, the studies employ rather arbitrary representations, for example, one arbitrary 3D rendered character is compared to another arbitrary cartoonish character. Thereafter the results are generalized to the whole groups of 3D characters and cartoonish characters respectively.

The point is that all too often the representations exhibit uncontrolled and unproblematized variations as to the graphical framework presented above – not the least when it comes to artistic qualities. For example, as with the study of van Vugt et al. (2007) commented on in Figure 9, there are many possible factors involved in the comparison between a relatively standarized and normal 3D rendering and an rather odd representation. The difference in such a case probably lies more in the artistic qualities. The comparison regards a kind of standard soft 3D rendering of a business character with a suite versus a faceted 3D rendering of an artistically quite bizarrely composed un-realistic(?) fantasy character with odd clothes. What then is really measured?

(The same reasoning can be applied to the study on androgyny by Nowak & Rauh (2008) presented in the next chapter (page 52).

Chapter 8:
From Visual Impact to Pedagogical Implications

: - !

With the presentation of a graphical design space and a discussion on the use of visual realism, it is now time to present an overview of the central research topics and the corresponding findings in the papers included in this thesis.

The Importance of Visual Aspects

The research behind this thesis started more than seven years ago, when it soon became obvious that visual or graphical issues – as to the basic appearance of the embodied agent itself (i.e. the *static visual characteristics*) – were often left out in computational oriented approaches such as: user modelling, animation algorithms for gestures, facial expressions and gaze, language processing, and affective computing. In particular, comparative studies that systematically varied these visual aspects were hard to find (*Paper I*). At the same time it seemed unlikely that visual aspects should not have an impact on the experience of embodied pedagogical agents given the collected experience and knowledge in practices like computer games, films, performing arts, and cartoons. Furthermore, there was also an abundance of empirical evidence in social psychology (*Paper I*).

As discussed in Part I, the field of embodied (pedagogical) agents is inherently connected to artificial intelligence and computational approaches. In such contexts, visual appearance and similar considerations might sometimes be seen as amusing tinkering with aspects outside the scientific scope. A personal reflection is that this mirrors the trap of the rational paradox (page 31). Visual design of embodied (pedagogical) agents is not accessible or describable by means of

rational models or schemas (or computational 'rational' algorithms) and thus visual design falls outside the scientific scope of the field.[2]

Today, empirical results that address aspects of visual design and the impact on the experience of the interactive embodied agent are no longer 'that odd' in the community of embodied (pedagogical) agents. There is also additional support in experiments demonstrating that aesthetic factors can have a positive impact on cognitive tasks (e.g. Norman, 2002; Lavie & Tractinsky, 2004).

The impact of physical appearance

Findings from social psychology

For a good while now, there has been ample evidence in social psychology[3] that the appearance and observable physical cues of other people profoundly affect our judgements. It has also been shown that similar responses apply to embodied virtual (pedagogical) agents (*Paper I-V*).

Figure 11. Disney has long excelled in the reproduction (and production?) of stereotypical instances such as *the princess* (*Paper I*).

An interesting argument for the impact of visual appearance is its role in 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 body shape, height, gender, race, physical attractiveness, hair, make-up, clothing, and so on. Noteworthy is also a quotation from Berscheid & Walster (1974): '*[O]ur 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.*' (*ibid.*, p. 159). The crucial issue is that regardless of how accu-

[2] To quote the (early) Wittgenstein: '*Wovon man nicht sprechen kann, darüber muss man schweigen.*', (Ludwig Wittgenstein, Logish-philosophische Abhandlung (Tractatus Logico-Philosophicus), 1921, Proposition 7). Note, however, that this statement was later rejected by the (latter) Wittgenstein. (In English: '*Whereof one cannot speak, thereof one must be silent.*', Translation by C. K. Ogden.)

[3] E.g. Schneider, Hastorf, & Ellsworth (1979; Schneider (2003).

rate such insights are, people do build them. In folk psychology this is recognized in idiomatic warnings such as: 'Don't judge a book by its cover.' (*Paper I*).

Furthermore, impressions of other people's personalities based on physical appearance may not only persist, but even increase over time (Mathes, 1975). Once again, in folk-psychology, you find notions like: '*First impressions are lasting.*' (*Paper I*).

Given the assumption that people fall back on social strategies when interacting with embodied agents in ways similar to real life interaction (Reeves & Nass, 1996), it is an intriguing idea that a principle like, 'First impressions are lasting' may bear upon embodied agents as well.

Evidence of visual impact of embodied pedagogical agents

An illustrative example regarding the effects of visual impact in the use of embodied pedagogical agents is reported by de Rosis et al. (2004). A virtual character was designed for a natural-language interface for a legal information system. Initially the character was designed as a very attractive young female assistant, since the developers assumed that the typical users of the system were 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 (*Paper I & II*).

The Power of Simplification (iconization)

A question of communication

The discussion on rationality in Part I (Chapter 5) contained a passage on computational and cultural considerations with regard to the embodiment of virtual agents (pages 31-32). An aspect of that discussion addresses the issue of resource demanding naturalistic 3D rendering versus simplified cartoon rendering in the presentation of embodied agents. Given the steady technical progress, issues of real time 3D rendering may only be a matter of time. On the other hand, as proposed in the discussion, there is more to visual presentation than computational power. The main objective for an embodied pedagogical agent is to engage in social interaction with the learner and thus strengthen the social and communicative features of a pedagogical system. This means that the basic concerns have to address the qualities of the pedagogical interaction (communication) itself, that is, how do different visual representations relate to positive pedagogical effects

such as motivation, stimulation, ease, flow, problem solving, understanding, etc. (c.f. Part I, Chapter 4).

Simplification (iconization)

The concept of simplification in the embodied (pedagogical) agent domain addresses the possible benefits of using a simplified (iconized) visual representation for the agent (Figure 12), (c.f. Chapter 6 and Figure 7).

Figure 12. Degrees of simplification (iconization) from highly stylized (left) to near naturalistic (right), (see also Figure 7).

There are several arguments pro and con the idea of simplification, some claiming that pictorial realism (i.e. naturalism) is a necessary condition for meaningful human co-operation with an animated agent (e.g. Welch et al., 1996; Nass, Isbister & Lee, 2000).

Outside the world of embodied agents, comic artist and writer Scott McCloud (1993) argues – in contrast – that subjective audience involvement can be increased by iconization. According to McCloud (*ibid.*), people will more easily get involved with and also be likely to project themselves into a visually simplified character than a highly detailed and naturalistic character. The highly detailed and naturalistic character is more of a visual and socio-emotional fact (an object associating to 'the other'), which does not leave much for a user to elaborate on and fill in. A stylized character, on the other hand, invites subjective identification and elaboration by the user, who may fill in from his or her own personal and subjective experiences. In McCloud's wordings, the stylized character is '*[A]n empty shell that we inhabit*.' (*ibid.*, p. 36).

The underlying mechanism brought forth by McCloud (*ibid.*) is that the concept and image of oneself (subject) is highly iconic in contrast to that of other people (objects) in one's environment. Therefore, identification and social affinity with an agent come more naturally and effortlessly in response to a simplified (iconic) agent. This, in turn, may increase the impact that the agent has on users (*Paper I* and Figure 13).

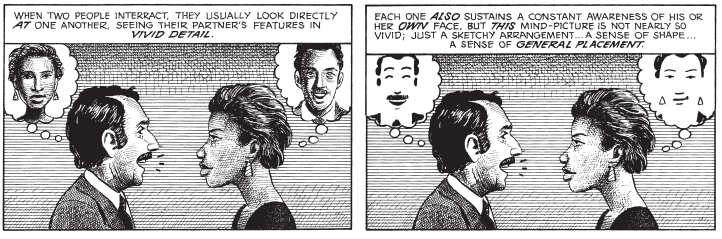


Figure 13. People who interact experience each other's faces in a detailed, realistic manner, whereas they only have a schematic, iconic representation of their own faces (McCloud, 1993, p. 35-36). © Scott McCloud & HarperCollins Publishers, Inc.

Iconization vs. naturalism

A direct pedagogical implication of the reasoning above can be exemplified with a virtual teacher in contrast to a virtual learning companion. If the design consideration of iconization-naturalism is applied, the teacher character (relating to an instance of the other, an object) might benefit from more naturalism in the representation. A learning companion character (with the potential of a subjective identification with 'oneself') may, on the other hand, benefit from a more simplified (iconized) representation (*Paper I-III*).

So far the theoretical reasoning. When it comes to empirical findings, the picture gets more complicated with diverging results in different studies. Part of this confusion often lies in very weak and non- reflected interpretations of the confusing concept of visual realism as discussed earlier (Chapter 7). In addition, we also find generalizations neglecting further contextual aspects as, for example, graphical style in relation to pedagogical role (as in the example above).

In one of our own studies (*Paper III*) we included 'pedagogical role' and 'communicative style' as contextual variables. The results of the study supported the idea that iconization may correspond to pedagogical role and communicative style.

Referring to anecdotic references from both studies and the domain of computer games, I would here like to speculate on two further factors: (i) *age dependence* in that younger students (first to sixth grade) may have a preference for more stylized (cartoonish) characters, and (ii) *artistic and graphical qualities* as to total experience and communications of emotional states. In the latter, the graphical style of *Manga* [4] (with explicit techniques for emotional expressiveness) has seen worldwide success.

[4] *Manga* has come to represent a widespread graphical style, though the actual meaning of the word is 'Japanese comics' (or literally: 'aimless pictures'), (EBO, 2009).

Iconization, similarity attraction, and role modelling

Other aspects of simplification (iconization) relate to identification issues: *similarity attraction* and *role modelling* (*Paper V*).

Similarity attraction

Similarity attraction has proved to be a working factor in embodied pedagogical agent settings. Lee & Nass (1998) report on participants perceiving an agent to be more socially attractive, trustworthy and intelligent when there was an ethnical match. Another example is the case with the legal information system presented on page 47.

Figure 14. The virtual coach (young and cool female) used in Baylor, Rosengard-Kima & Plant (2006).

More recent effects of similarity attraction are reported by Baylor, Rosengard-Kima & Plant (2006). In an engineering tutorial, young female students were most strongly affected as to motivation and self-efficacy regarding engineering subjects by virtual coaches that were similar to themselves (or similar to how they would like to be) – in this case female, young and cool (Figure 14).

This points at a dilemma in embodied agent design. One solution is to design multiple agents to always have one that may work on the basis of similarity attraction. This may, however, underpin negative aspects of group mentality in terms of 'we and them' or alignment to detrimental stereotypical behaviours. An alternative approach may be to design simplified (iconized) characters with a lesser degree of visual clues as to ethnicity, body constitution, social status, etc. (This seems to be Disney's recipe in recent years, where ethnic markers as skin tone are held back (Engholm, Michelsen, 1999).)

Role modelling

Role modelling is another fruitful concept in the research of embodied pedagogical agents. Opportunities for role modelling are known to strengthen development in pedagogical terms (Bandura, 1977).

One parameter that is known to influence the strength of a role model is attractiveness. If a role model is perceived as attractive the behaviour of the model is more often imitated (Rommes et al., 2007). Studies by Baylor and her collaborators (Baylor & Plant, 2005; Baylor, Rosenberg-Kima & Plant, 2006) have

highlighted the importance of images and alternative cultural role models for engineering students. Also other researchers have suggested that more physically attractive and glamorous female role models might change the negative prototypes of computer scientists (Coltrane and Adams, 1997).

In comparison with traditional pedagogical media, virtual settings provide specific potentials for role modelling and identification. By offering a broad range of alternative pedagogical characters with regard to visual style, personality, and social identity, it can be possible for a larger number of students to identify with a virtual pedagogical character (find a role model) and to engage in the learning task.

However, the use of role modelling is complex and comprises hidden dilemmas. For example, the connection between role modelling and visual stereotypes comes forth in the article by Baylor, Rosengard-Kima & Plant (2006) in which it was demonstrated that the use of virtual pedagogical coaches portrayed as *young and attractive females* (see Figure 14) can increase the willingness of female students to apply for technical courses and help to increase their self-efficacy. At the same time it was found that these positive results stemmed from a conception of a 'female, feminine, young and attractive' engineer as less competent than a 'real, typical male engineer'. The prejudice that females, and in particular females with a more pronounced feminine appearance, are less competent in technical domains seems to spill over to the virtual area, generating increased self-efficacy of the kind '*If she is able to do it, I can do it!*'

This implies a potential conflict between a short-term pedagogical goal of recruitment and boosted self-efficacy in female students, and a long-term pedagogical goal of changing rather than reproducing gender prejudices and stereotypes.

Simplification and androgyny

Thus, similarity attraction and role modelling can have both positive scaffolding as well as negative detrimental effects. Is there a way to 'have your cake and eat it'? In *Paper V: Gulz & Haake (in press)* we investigated the possibility of using androgyny to counter detrimental effects of gender prejudices.

Androgyny is a complex issue. Not being able to decide whether someone is a man or a woman is known to induce insecurity and unease in many people (Brave and Nass, 2005). Attractiveness (mentioned above), however, can be a factor here due to its role in well-functioning role models and cultural images (Rommes et al., 2007). In addition to this, attractiveness may even overrule the 'unease' of androgyny (Hess, 2007).

By designing the pedagogical characters in a naturalistic, but simplified form and thereafter render them in a 'flat' posterized visual graphical style (c.f. Figure 7, left) – we could manipulate and express aspects of androgyny and attractiveness in combination with a satisfying visual (aesthetic) result (Figure 15, left).

This simplified graphical style of the characters might furthermore facilitate identification and thus counter the 'unease' of androgyny … and as for the results, the study gave reason to believe in this possibility (see pages 61-64).

Figure 15. Left: Androgynous agents used in *Paper V: Gulz & Haake (in press)*; Right: Androgynous avatars used in Nowak & Rauh (2008). Note that the two left agents share the same basic 3D mesh. The main differences are the clothes and hair, with only minor modifications as to make-up (lips, eye liners, and eye brows) and light angle.

In a widely noticed study by Nowak & Rauh (2008) the authors argued that androgyny decreases avatar credibility. However, the study has shortcomings as to the actual avatars that were evaluated and therefore the generalization that are made of the results is unsupported (Figure 15, right). Referring to the Gulz & Haake study above, there are several visual factors beside androgyny that may count for a decrease in credibility such as experiences of attractiveness, graphical (aesthetic) qualities, oddness, etc. (Figure 15, right).

Summary: simplification (iconization)

There is reason to believe that well-designed cartoonish simplification (iconization) can have positive qualities in the representation of embodied pedagogical characters. It may ease subjective identification by using generalized, symbolic features. Another aspect is that there are less visual obstacles for subjective identification.[5] This may also be used to more specifically defuse visual clues of ethnicity and gender in order to decrease detrimental prejudices.

A more technically oriented argument for visually simplified characters is that it may decrease the mismatch between highly naturalistic characters as to visual

[5] Thanks to Henrik Enquist (Certec, Lund University) for this alternative view.

appearance (static visual characteristics) and stiff, awkward as well as resource demanding animations. In less naturalistic settings, the degrees of freedom for animation are substantial higher as long as the visual experience functions as a whole. All in all, I find it a bit distressing and a misuse of resources that such a large part of the research field of embodied pedagogical agents seem so stuck to the idea of all-over 'realism'.

Visual Stereotypes and Embodied Pedagogical Agents

As already argued for, the visual perception of other people plays a central role in human-human social interaction with profound effects on attitudes as well as behaviour. Accordingly, different visual aspects can be assumed to be critical for a smooth and effective human-agent communication and the potential to engage and motivate a learner.

In order to pursue the examination of these issues, it can be fruitful to change the perspective to that of *visual stereotypes* in relation to embodied pedagogical agents. This approach can be used to especially address the questions of to what degree and in what circumstances visual issues have a pedagogical impact regarding embodied pedagogical agents.

Visual characteristics

In order to approach visual stereotypes in relation to visual properties in embodied pedagogical agents, one can rely on the basic categorizing of visual aspects into dynamic and static characteristics as outlined in the proposition for a graphical design space for embodied pedagogical agents (Chapter 6).

i) *Dynamic visual characteristics* such as facial expressions and gaze, gestures, and movements.

ii) *Static visual characteristics* such as body and face properties, skin, hair, clothes and attributes (i.e. the underlying graphical entities).

In the following discussion, the focus is on the static visual properties as a source for the construction of visual stereotypes. The reason for this is that the dynamic qualities (as mentioned previously), are relatively well researched. This is also the case for the non-academic practice of animation, as exemplified in the unofficial canon *The Illusion of Life: Disney Animation* by Frank Thomas and Ollie Johnston (1984), two then leading Disney animators. The explicit articulation of the book is on the art of animation and illusion of life (*dynamic aspects*) while issues of the underlying visual form (*static aspects*) only are addressed implicitly – even if their importance certainly are acknowledged (*Paper I: Gulz & Haake, 2006a*).

Cognitive implications of visual stereotypes

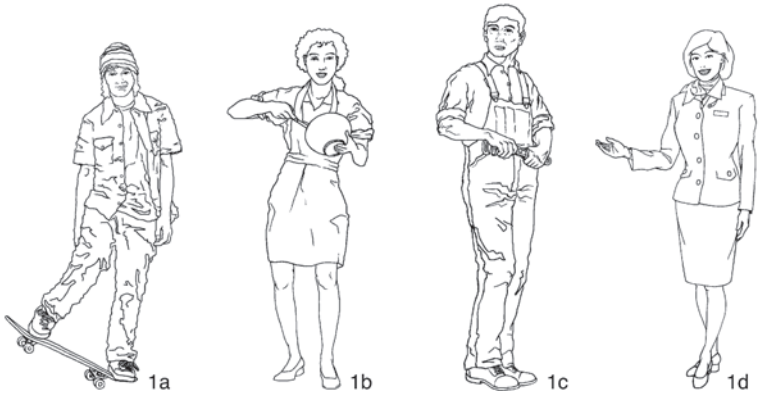


Figure 16. Examples of visual stereotypes

Figure 16 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 our sense of the term, will then 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. In this way, the visual input activates expectations of 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 and cultural baggage.

In the everyday use of the term, stereotypes convey a negative connotation of human behaviour, but it is not that simple. From a cognitive perspective, the use of stereotypes is probably of outermost importance in order to interact with other people. They function as cognitive *short cuts* to make action and life manageable for human beings. Instead of becoming overwhelmed by analyses, thoughts and questions regarding people that we encounter, we make use of their visual appearance to quickly situate them in order to focus on interaction as such (Brewer, 1988; Laurel, 1993). In this way, visual stereotypes frame peoples' expectations, and also support the building of common references for the conversation. 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 (Smith & Medin, 1981). This was also one of the 'not thought of' (nasty) problems of artificial intelligence as commented by Moravec (see footnote 3 on page 7).

Virtual Visual Stereotypes

In their seminal book, *The Media Equation*, Reeves & Nass (1996) convincingly demonstrate how social behaviour and strategies from the real world often are reproduced in interaction with computers. It has also been shown that these findings apply to the reproduction of real world visual stereotypes in the interaction with virtual characters.

A possible conclusion is then that positive as well as negative aspects of real-life visual stereotypes are reproduced in interactive visual media as cognitive tools for handling a complex social environment together with their problematic normative consequences.

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 opportunity 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.

This portrayal of the ideal can be taken one step further with interactive computer media. A key difference lies in what is 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. It is not all that inconceivable that the distance between users or learners and those 'ideal super people' will narrow or even diminish. The consequences of such a supposition are certainly not evident – but in an era desperately pursuing perfection in appearance, this might have detrimental effects on people's self-image and self-esteem (Figure 17).

Figure 17. The on-line 3D chat *imvu*® (www.imvu.com; © IMVU) explicitly exploits the idea of escaping the reality and one's real self to enter a world of 'perfect' people (with 'perfect' lives).

Chapter 9
Empirical Findings

: -$

The empirical findings of this thesis can be found in Paper III-V of the appended papers. The studies behind these three papers are also reported in other papers presenting additional empirical findings. Those papers are also mentioned in this presentation.

Furthermore, all studies are the collaborative work of myself, Agneta Gulz (my wife and colleague), and others. Given this, I will in this chapter keep to 'passive tense' and/or 'first person plural'.

The Studies

Paper III: *A Look at the Roles of Look & Roles in Embodied Pedagogical Agents – A User Preference Perspective*

The paper reports on a user study addressing three aspects of embodied pedagogical agents: visual static appearance, pedagogical role, and communicative style. Additional findings from this study are presented in Gulz (2005), Gulz & Haake (2005; 2006b; 2009). (A special thanks to Martin Jonasson for collaboration in the study.)

Study

In the study, 90 school children (aged 12-15) were (one-by-one) introduced to a dummy multimedia program were they were presented with either an instructor or a learning companion condition. This part of the presentation was communicated by the multimedia program by means of a pre-recorded voice.

The participants were then to make a preference choice between eight visually different embodied pedagogical agents presented on the screen. The eight agents were designed upon four young grown-ups, two males and two females. Their physical appearance was chosen to be relatively neutral.[6] Each person was

[6] The notion "neutral" should here be interpreted as: not visually deviating and/or conspicuous for the participants in the study.

then rendered into either a more naturalistic (detailed & 3D) or a more stylized (simplified & cartoonish) version (see Figure 18).

After the preference choice, the participants were asked whether they preferred a strictly task-oriented or a task- and relation-oriented pedagogical agent.

Figure 18. The dummy multimedia program used in the study: Top left: data collection of participants and setting for instructor/companion version; Top right: start page, Bottom left: introduction page; Bottom right: Preference choice page.

Results

The goal of the study was to explore possible relations between the three aspects mentioned above with respect to user preferences. The results were:

i) when the agent was introduced as a learning companion, female students displayed a significant tendency to choose a more stylized visual character,
ii) when the agent was introduced as a learning companion, female students displayed a significant tendency to choose a task- & relation-oriented agent, and
iii) in the case when students had chosen a more stylized character, there was a significant tendency to choose a task- and relation-oriented agent.

Comments

One limitation of the study was that the variables pedagogical role and communicative style of the agents were not implemented as actual algorithm driven agent behaviour. This limitation lowers the ecological value, but may on the other hand capture some additional knowledge on a more generalized, conceptual and articulated level.

Another concern of the study is the age group (12-15 years). Other age groups might have given different results. A personal speculation is that the relative preference for cartoonish stylized agents would increase with younger age groups.

Regarding the results, the most prominent and central finding is that a deconstruction of the user group with respect to parameters such as (in this case) communicative style and gender can give significant results as to specific preferences like pedagogical role and visual style of the agent – a topic that has been discussed earlier in the thesis.

Given the more specific focus of this thesis, the results gave support for the assumption that the graphical style of the embodied pedagogical agent can have significant impact on the experience of the user, and that iconized characters should be considered in cases where subjective identification and social interaction is more emphasized.

Paper IV: *Visual Femininity and Masculinity in Synthetic Characters & Patterns of Affect*

The paper reports on a study investigating stereotypical predictions with respect to visual femininity and masculinity in embodied pedagogical agents. (A special thanks to Dan Holmér for collaboration in the study.)

Study

Forty adolescents (aged 13-18) and 40 adults (aged 25-65) participated. In each age group there was an equal number of female and male participants.

In the study, the participants were presented to two computer characters (one female and one male medical doctor) speaking about diurnal rhythm – a topic selected to be relatively gender neutral. The two characters were randomized from a set of two female doctor characters (one more feminine-looking and one more neutral) and two male doctor characters (one more masculine-looking and one more neutral), see Figure 19.

Relevant visual cues were systematically varied, whereas voice, spoken content, linguistic style and role of the characters were held constant. After the session the participants were interviewed about their impressions.

Figure 19. Screen shots of the four characters used in the study. Upper left: (more feminine-looking), upper right: (more neutral), lower left: (more neutral), lower right: (more masculine-looking).

Results

The results both followed and contradicted stereotype predictions:

i) The two female characters differed in accordance with stereotype predictions (e.g. the more feminine-looking female was perceived as significantly more personal, pleasant, and less intelligent) with the exception of competence-related traits.
ii) The two male characters differed very little with respect to stereotype predictions.
iii) The pattern for male versus female characters was slightly in opposite to stereotype predictions.

Comments

The study was inspired by a study reported in Voelker (1994) showing that users of a digital system perceived a more masculine-sounding female voice as more persuasive and intelligent than a corresponding but more feminine-sounding

female voice. Our study instead explored whether a parallel pattern of affectively coloured evaluations can be elicited when femininity and masculinity are manipulated via visual cues instead of via voices.

The hypothesis regarding visual stereotype prejudices were correct when comparing the two female characters, whereas there were only small variations between the two male characters.

The interesting, but puzzling result was that no difference in stereotype prejudices were found when comparing the female versus the male characters – this in contradiction to the initial hypothesis.

A possible explanation for the weak or even contradictory results was that gender stereotypes here were overridden by a medical doctor stereotype. In the Swedish society this is a high status profession, with its practitioners ascribed expertise, knowledge and intelligence.

A replication of the study but now with voices only (no visual cues) gave a possible additional explanation. The male voice had a regionally coloured dialect, whereas the female voice was that of standard Swedish. Theory says that dialects are perceived as warmer, kinder, more emotional and more naïve, whereas the official standard version of a language is associated with high status, authority, professionalism, and credibility. Thus, the original experimental setup featured a high-status female voice associated with stereotypic male traits versus a low-status male voice associated stereotypic female traits. In light of this, the 'out-levelled' results in the original study may be due to the fact that dialect and gender stereotypes in this case counteracted one another.

A lesson learned from this study is that different traits in embodied (pedagogical) characters can interact in ways hard to predict.

Paper V: Challenging Gender Stereotypes using Virtual Pedagogical Characters

This last paper reports on a study exploring motivational and cognitive effects of more neutral or androgynous-looking versus more feminine-looking and masculine-looking virtual characters. Additional findings from this study are presented in Gulz, Haake & Tärning (2007a; 2007b; 2007c). (Thanks to Adam Altmejd, Aron Vallinder, Betty Tärning and Thérèse Deutgen for collaboration in the study.)

Study

158 students, aged 17-19, encountered one out of four virtual presenters informing about university programmes at Lund University (Figure 20).

Figure 20. Screenshot from the multimedia presentation with the more androgynous or neutral young woman (FA) presenting the programme in computer engineering at Lund University.

The virtual presenters were visually manipulated to represent gender stereotypicality versus androgyny as follows: a more feminine-looking female character (FF), a more neutral or androgynous-looking female character (FA), a more neutral or androgynous-looking male character (MA), and a more masculine-looking male character (MM), see Figure 21.

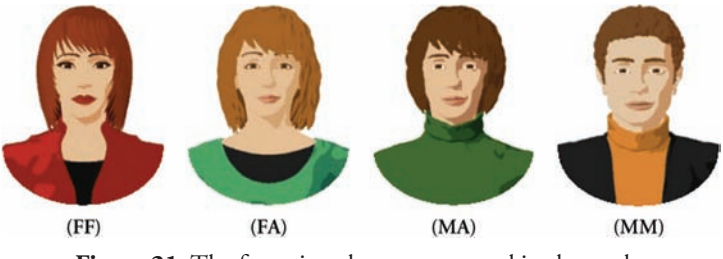


Figure 21. The four virtual presenters used in the study.

The two more neutral or androgynous-looking characters (FA & MA) were developed out of an identical bust (3D mesh). Compared to MA, FA had: (i) longer hair, (ii) more regular and slightly plucked eye brows, (iii) different neutral clothing, (iv) painted eye lashes, and (v) a slightly lighter colour scheme.

All other variables were kept constant except the voice (a recording of a female voice), which was digitalized into a female and a male voice.

An important part of the study examined whether the different virtual presenters had any effect on the participating students' attitude to a university programme in computer engineering. The attitudes were measured before and after the multimedia presentation. The relative difference in attitude was denoted the presenters' *implicit* influence.

After this, the students' *explicit* attitudes towards the different virtual characters were measured by means of preference ranking and verbal articulation.

Results

All in all, there are several reports from the study (see above), and the focus group interviews are still waiting to be reported. Below follows a summary of the findings presented in *Paper V*.

i) The two more neutral or androgynous-looking characters were generally preferred as presenters when ranked explicitly, in particular so among female participants.
ii) In explicit gender related arguments, FA was (without exception) promoted as presenter in terms of being a woman, whereas FF was generally dismissed for being a woman ('this kind of woman').
iii) Explicit comments of attractiveness and plainness were evenly distributed over the four characters (with the exception of FF who received one third of the attractiveness/non-attractiveness comments). Notably, the students diverged considerable in their opinions regarding the four characters.
iv) The implicit influence of the presenters was measured (as described above) as the relative change in attitude towards the presented education. For female students, the positive implicit influence was evenly distributed over the characters (which can be contrasted to their strong explicit preference for FA & MA). For male students it was MM & FF that had a strong implicit influence, whereas it was FA & MA that were preferred in their explicit ranking and reasoning. That is, there was a considerable mismatch between the implicit and explicit preferences.

Comments

Two central results of this study were: (i) that 'more androgynous' characters indeed were possible to use as virtual presenters, and (ii) that there was a considerable mismatch between the students *explicit* (conscious) preferences and argumentation versus the *implicit* (unconscious) impact of the characters, and especially so for the male students.

The virtual characters that presented a university programme in computer engineering can probably be used to also present other educational domains. They may be of particular interest when dealing with gendered occupational choices.

At the same time one should be aware of mismatches between explicit actions and implicit reactions. Androgyny, role modelling, and cultural images are complex phenomena and must be handled with cautious reflection.

A Comment on Methodology

Methodology is a delicate subject to which whole dissertations can be devoted. With our cross-scientific approach, we pick up methods from quite different academic areas and traditions and have little understanding for the sometimes infected dispute on qualitative versus quantitative methods. Our research projects usually start with an idea or fascination related to some phenomenon. Next step is to formulate some research question – and out from this, hypotheses and methods emerge depending on the phenomenon of interest and the context at hand. With regard to the research questions we are usually dealing with, we employ a mix of qualitative and quantitative methods and aim at measuring both explicit (conscious) actions and implicit (unconscious) reactions (c.f. *Paper V*).

A drawback with this approach is that we do not have specialist expertise in the methods (methodologies) we use. This also means that we, especially with regard to quantitative data collection and data analysis, usually rely on relatively simple and robust experimental designs that require larger data sets (i.e. more test participants). It takes more work, but the quality of the results may even be better.

Finally, as a comment on the eternal disputes on methodology and different scientific approaches and traditions, we propose that the basic and common practice of science (except a creative brain) can be formulated as follows:

i) be as careful and precise as possible during the research process, and
ii) document and report as much as possible of the research process in a clear, transparent, and sincere manner.

In this way, other researchers as well as the surrounding society can evaluate the research and even form different opinions based on the documentation of the research process, the data, and the results. Any critique will then (hopefully) concern the specific study, not a general method.

After all – research is about gaining knowledge of phenomena around us, and in this endeavour methods can be of outmost importance, but in the end they are ... only methods.

Chapter 10
Discussion

8~}

In this final discussion I comment on the central topics of this thesis and eventually end with a reflection.

A General Discussion

The discussions in this thesis have addressed psychological, cognitive and social aspects of embodied virtual characters in pedagogical contexts with respect to static visual characteristics. The virtual characters themselves are, however, computational algorithms and the research area is heavily computational. This makes the study of human–agent interaction a both fascinating and delicate cross-scientific venture. If one aims at more than retrospective descriptive analyses, it is necessary to encompass scientific traditions, methods and skill from many different academic domains – including technology and computation.

Embodiment

The concept of embodiment is a prerequisite for this thesis. With the breakthrough of embodiment, a whole new spectrum of visually grounded communication strategies was at hand. As we have seen, humans intuitively make use of inter-human social communicative strategies. This anthropomorphic reflex is unconscious and probably biologically hardwired. An argument for this is the view of a by evolution tuned human social intelligence, relying on a whole set of advanced and complex strategies to communicate and navigate in a social world.

This close connection between anthropomorphized human–agent interaction and everyday human–human interaction brings with it the whole baggage of social strategies in the form of stereotypes, prejudices, identification, role modelling, etc. – with all its potentials and risks.

Virtual embodied characters as a research tool

The visual nature of embodiment sets the focus on the visual characteristics of the communication between human and agent as well as between humans. Virtual

embodied characters have a large potential in providing knowledge about human cognition and behaviour. In contrast to real human beings, they can be precisely controlled in interactive experimental settings. For example, it is now possible to examine stereotypes of visual clues, body language, and voice by letting (human) participants encounter the one and same interactive embodied character – with relevant variables (clothes, hair cut, body language, voice) systematically manipulated. This has not been possible before, since not even trained actors exactly can repeat or modify their behaviour, body language, or voice between test rounds.

As for a future research agenda – here I have more than I bargained for.

The Visual Impact and Pedagogical Implications

Compared to other aspects of embodied pedagogical agents, static visual characteristics are sparsely researched. As for today, we have little knowledge of the impact of visual parameters in pedagogical settings. A glance at other visual media such as theatre, film, and computer games with their strong emphasises on visual aspects makes it on the other hand a both interesting and urgent question.

In an attempt to advocate the static visual qualities, I have in this thesis proposed a framework for a visual (graphical) design space in order to describe what visual properties are actually measured and manipulated in different studies. A future goal is to use this framework to review the research on embodied pedagogical characters. As for now, I have in the thesis commented on the conceptual chaos in the use of 'realism'.

Simplification (iconization)

Visual simplification (or iconization) has been a central topic in this thesis. With regard to the computational dominance of this field, its rationalistic tradition, and the aim of artificial intelligence to construct an artificial intellect mimicking human cognition, it seems as if 'human naturalism' is an inherent implicit goal of the field per se.

An alternative standpoint is to focus on the development of engaging, effective and smooth interactive embodied pedagogical agents that works in real pedagogical contexts. After all, the main objective for embodied pedagogical agents is to engage in social interaction with learners in order to strengthen the social and communicative features of pedagogical systems.

In the light of this, visually simplified, more or less cartoonish representations of embodied pedagogical characters may have considerable potentials with their

possibility to afford subjective identification and thus engagement and motivation. Simplified characters may have similar effects with regard to role modelling and group identification in terms of gender, age, social status, ethnicity, etc.

Virtual Stereotypes

Regarding the 'social baggage' of embodiment mentioned above, this thesis has looked closer on stereotypes. Simultaneously as virtual characters have the potential to counter stereotypes, the general trend seems to rather be a further exploitation. A troublesome aspect here is that virtual characters do not have the same constraints as to behaviour and visual appearance as humans do. They can – maybe not now, but in the future – be modelled into a sort of magnificent, wonderful super beings with perfect lives. This is nothing new – society has since long created heroes and more lately movie stars and rock stars, but whereas these 'traditional' superstars were imaginary products, these new virtual characters has the power of dynamic interactivity, blurring the border between reality and fantasy.

Closing

To finally end this thesis, I put forth two reminders and one strange reflection.

Reminder 1
Embodied pedagogical agents are no a digital monsters threatening to replace teachers – they are only a tool with the potential to help teachers in daily classroom activities. Nothing can replace a good human teacher, but we will never have enough teachers for our youths – and here they can be a valuable complement.

Reminder 2
I surely can get carried away by the possibilities of virtual characters, but in much they are more of an artifact of the future than of today's reality. Creating an artificial being with the social and communicative capabilities of a human is, I believe, far beyond the horizon. In the near future we will probably have to rely more on the human capacity and willingness to engage in virtual creatures, than on the capacity of these creatures.

A strange reflection

I have sometimes wondered over the shortcomings of artificial intelligence when it comes to modelling human cognition in its broad sense with emotions, intuition, and social and communicative abilities. On the other hand – are humans that fantastic? If we take a closer look at any human, she or he is so full of strange behaviours, daily mistakes and slips, ambiguous emotions, cognitive breakdowns, incomprehensible utterances, contradictory thoughts, etc.

Considering this: 'Why do we demand so much of these virtual characters? Why do we want them to be so perfect?'

This is to compare apples to oranges. Maybe it is the capacity to communicate and engage in social activity, in spite of all strange and inconsistent events that constantly surround us, that constitutes the human. And the quest for a computational perfect being is in much the ghost of rationality.

References

André, E., Rist, T. & Müller, J. (1998). Integrating reactive and scripted behaviours in a life-like presentation agent. In: K. Sycara & M. Woolridge (Eds.), *Proc. of the Second Int. Conf. on Autonomous Agents (AGENTS'98)*, pp. 261–268. New York, NY: ACM Press.

Bandura, A. (1977). *Social Learning Theory*. Prentice Hall.

Baylor, A. & Plant, A. (2005). Pedagogical agents as social models for engineering: The influence of appearance on female choice. In: *Proc. of Artificial Intelligence in Education 2005*, Amsterdam, the Netherlands.

Baylor, A., Rosenberg-Kima, R. & Plant, A. (2006). Inteface agents as social models: The impact of appearance on females' attitude toward engineering. In: *Proc. of CHI 2006*, Montreal, Canada.

Berscheid, E. & Walster, E. (1974). Physical attractiveness. In: L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, 7, pp. 158-216. New York, NY: Academic Press.

Bickmore, T. & Cassell, C. (2001). Relational agents: a model and implementation of building user trust. *Proc. of the ACM SIGCHI Conf. on Human Factors in Computing Systems (CHI-01)*, pp. 396-403. Seattle, WA.

Bickmore, T. (2003). *Relational Agents: Effecting Change through Human-Computer Relationships*. PhD Thesis, Massachusetts Institute of Technology (MIT), Cambridge, MA.

Bickmore, T. & Picard, R. (2003). Subtle expressivity by relational agents. In: *Proc. of the CHI 2003 Workshop on Subtle Expressivity for Characters and Robots*. Fort Lauderdale, FL.

Blair,K., Schwartz, D., Biswas, G & Leelawong, K. (2006). Pedagogical agents for learning by teaching: teachable agents. *Educational Technology & Society, Special Issue on Pedagogical Agents*, pp. 62-66.

Branham, S. (2001). Creating physical personalities for agents with faces: modelling trait impressions of the face. In: *Proc. of the UM 2001 Workshop on Attitudes, Personality and Emotions in User-Adapted Interactions*. Sonthofen, Germany. Retrieved July 18, 2005, from: http://www.di.uniba.it/intint/sonthofen/brahnam-f.pdf

Brewer, M. (1988). A dual process model of impression formation. In: R. Wyer & T. Srull (Eds.), Advances in Social Cognition, 1, pp. 1-36. Hillsdale, NJ: Erlbaum.

Brigham, J. (1980). Limiting conditions of the 'physical attractiveness stereotype': attributions about divorce. *J. of Research in Personality*, 14, pp. 365-375.

Brooks, R. (1999). *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press.

Brooks, R. (2002). Flesh and Machines: How Robots will Change Us. New York, NY: Pantheon Books.

Cassell, J. & Thórisson, K. (1999). The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, 13, pp. 519-538.

Clark, A. (1999). Embodied, situated, and distributed cognition. In: W. Bechtel & G. Graham (Eds.), *A Companion to Cognitive Science*, pp. 506-525. Malden, MA: Blackwell.

Coltrane, S. & Adams, M. (1997). Work-family imagery and gender stereotypes: television and the reproduction of difference. *J. of Vocational Behavior*, 50, 323-347.

Cook, M. (1979). *Perceiving Others: The Psychology of Interpersonal Perception*. London, UK: Methuen.

Crevier, D. (1993). *AI: the Tumultuous History of the Search for Artificial Intelligence*. New York, NY: BasicBooks.

Engholm, I. & Michelsen, A. (1999). Designmaskinen: design af den moderne verden. Copenhagen, Denmark: Gyldendahl.

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

de Rosis, F., Pelachaud, C. & Poggi, I. (2004). Transcultural believability in embodied agents: a matter of consistent adaptation. In: S. Payr & R. Trappl (Eds.), *Agent Culture: Designing Human-Agent Interaction in a Multicultural World*, pp. 75-105. Mahwah, NJ: Laurence Erlbaum Associates.

Dillenbourg, P. (1999). *Collaborative Learning: Cognitive and Computational Approaches*. Oxford, England: Pergamon.

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, MA: MIT Press.

Encyclopædia Britannica Online / EBO (2009). *Articles*. All articles retrieved February to April 2009, from: http://search.eb.com.ludwig.lub.lu.se/eb/[*article-id*]

Lammers, S. (Ed.), (1986). *Programmers at Work*. Buffalo, NY: Microsoft Press.

Gard, T. (2000). Building character. *Gama Network: www.Gamasutra.com*. Retrieved July 18, 2005, from: http://www.gamasutra.com/features/20000720/gard_pfv.htm

Gratch, J., Marsella, S., Egges, A., Eliëns, A., Isbister, K., Paiva, A., Rist, T. & ten Hagen, P. (2004). Design criteria, techniques and case studies for creating and evaluating interactive experiences for virtual humans. In: *Report (Paper 04121): Dagstuhl seminar 04121 on Evaluating Embodied Conversational Agents*. Schloss Dagstuhl, Germany.

Gulz, A. (2004). Benefits of virtual characters in computer based learning environments: claims and evidence. *Int. J. of Artificial Intelligence in Education*, 14, pp. 313-334.

Gulz, A. (2005). Social enrichment by virtual characters – differential benefits. *J. of Computer Assisted Learning*, 21(6), pp. 405-418.

Gulz, A. & Haake, M. (2005). Social and visual style in virtual pedagogical agents. In: *Proc. of the UM 2005 Workshop on Adapting the Interaction Style to Affective Factors*, Edinburgh, Scotland.

Gulz, A. & Haake, M. (2006a). Design of animated pedagogical agents – a look at their look and visual form. *Int. J. of Human-Computer Studies*, 64(4), pp. 322-339.

Gulz, A. & Haake, M. (2006b). Virtual pedagogical agents - design guidelines regarding visual appearance and pedagogical roles. In: A. Méndez-Vilas, A. Solano Martin, J. Mesa Gonzalez, J. Mesa Gonzalez (Eds.), *Current Developments in Technology-Assisted Education, Proc. of mICTE 2006*, vol. III, pp. 1848-1852, Seville, Spain. Badajoz, Spain: FORMATEX.

Gulz, A. Haake, M. & Tärning, B. (2007a). Challenging gender stereotypes using virtual pedagogical characters. *Project Report for the Helsingborg 2007 Symposium on Gender, Learning and IT*, Helsingborg, Sweden.

Gulz, A., Haake, M. & Tärning, B. (2007b). Digitala lärmiljöer och androgynitetens potential. In: *Proc. of Utvecklingskonferens LU: Att tänka om – ett kvalificerat akademiskt lärarskap*, Lund, Sweden.

Gulz, A., Haake, M. & Tärning, B. (2007c). Visual gender and its motivational and cognitive effects- a user study. *Lund University Cognitive Studies*, 137.

Gulz, A., Ahlner, F. & Haake, M. (2007). Visual femininity and masculinity in synthetic characters & patterns of affect. In: A. Paiva, R. Prada & R. Picard (Eds.), *LNCS 4738: Proc. of the Second Int. Conf. on Affective Computing and Intelligent Interaction*, pp. 654-665, Lisbon, Portugal. Berlin/Heidelberg, Germany: Springer.

Gulz, A., & Haake, M. (2009). Benefits of digital pedagogical characters – a learning style perspective. *The Open Education J.*, 2, pp. 34-41.

Gustavsson, E. & Czarniawska, B. (2004). Web Woman: the on-line construction of corporate and gender images. *Organization*, 11(5), 651-670.

Haake, M. (2006). *Virtual Pedagogical Agents: Beyond the Constraints of the Computational Approach*. Licentiate Thesis, Lund University, Sweden.

Haake, M. & Gulz, A. (2008). Visual stereotypes and virtual pedagogical agents. *Educational Technology & Society*, 11(4), pp. 1-15.

Haake M. & Gulz, A. (2009). A Look at the roles of look & roles in embodied pedagogical agents – a user preference perspective. *Int. J. of Artificial Intelligence in Education*, 19, pp. 39-71.

Hansson, G., Jansson, T., Linell, P. & Sjölander S. (1979). *Kommunikation mellan människor*. Linköping: Linköping University.

Heinlein, R. (1953). Assignment in Eternity. New York, NY:Baen.

Hess, U. (2007). The assumption that 'attractiveness may overrule feelings of uneasiness' comes from a discussion with prof. Ursula Hess (Professor of Psychology, University of Quebec, Canada). Discussion at ACII 2007, Lisbon, Portugal.

Isbister, K. (2006). *Better Game Characters by Design – A Psychological Approach*. San Francisco, CA: Morgan Kaufmann.

Johnson, W., Shaw, E., Marshall, A. & LaBore, C. (2003). Evaluation of user interaction: the case of agent Adele. In: *Proc. 2003 Int. Conf. on Intelligent User Interfases (IUI 2003)*, pp. 93-100. New York, NY: ACM Press.

Langlois, J., Kalakanis, L., Rubenstein, A., Larson, A., Hallam, M. & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*. 126, pp. 390-423.

Lassetter, J. (1987). Principles of traditional animation applied to 3D computer animation. *ACM SIGGRAPH Computer Graphics*, 21(4), pp. 35-44. New York, NY: ACM Press.

Laurel, B. (1993). *Computers as Theatre*. Addison-Wesley, Reading, MA.

Lavie, T. & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *Int. J. of Human-Computer Studies*, 60(3), pp. 269- 298.

Lee, E.-J., Nass, C. (1998). Does the ethnicity of a computer agent matter? An experimental comparison of human-computer interaction and computer-mediated communication. In: *Proc. of the WECC 1998 Workshop on Embodied Conversational Characters*, pp. 123-128, Tahoe City, CA.

Lester, J., Converse, S., Kahler, S., Barlow, T., Stone, B. & Bhogal, R. (1997). The persona effect: affective impact of animated pedagogical agents. In: *Proc. of the ACM SIGCHI Conf. of Human Factors in Computing Systems (CHI-97)*, Atlanta, GA.

Lindblom, J. & Ziemke, T. (2003). Social situatedness of natural and artificial intelligence: Vygotsky and Beyond. *Adaptive Behaviour*, 11(2), pp. 79-96.

Massaro, D., Cohen, M., Beskow J. & Cole, R. (2000). Developing and evaluating conversational agents. In: J. Cassell, J. Sullivan, S. Prevost & E. Churchill (Eds.), Embodied Conversational Agents, pp. 287-318. Cambridge, MA: MIT Press.

Mathes, E. (1975). The effects of physical attractiveness and anxiety on heterosexual attraction. *J. of Marriage and the Family*, 37, pp. 769-773.

McCloud, S. (1993). *Understanding Comics: The Invisible Art*. New York, NY: HarperPerennial.

Moravec, H. (1988). *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, MA: Harvard University Press.

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

Moreno, R. (2004). Immersive agent-based multimedia environments: Identifying social features for enhanced learning. In H. Niegemann, F. Leutner & R. Brünken (Eds.), *Instructional Design for Multimedia Learning*, pp. 9-18. Münster, Germany: Waxmann.

Merriam-Webster Online Dictionary (2009). *Agent*. Retrieved February 25, 2009, from: http://www.merriam-webster.com/dictionary/agent

Moundridou, M. & Virvou, M. (2002). Evaluating the persona effect of an interface agent in a tutoring system. *J. of Computer Assisted Learning*, 18, pp. 253-261.

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

Nass, C. & Brave, S. (2005). Wired for Speech. Cambridge, MA: MIT Press.

Nasseh, B. (1997). *A Brief History of Distance Education*. Ball State University. Retrieved March 24, 2009, from: http://www.bsu.edu/classes/nasseh/study/history.html

Norman, D. (2002). Emotion & design: attractive things work better. *Interactions*, 9, pp. 36–42.

Nowak, K., & Rauh, C. (2008). Choose your 'buddy icon' carefully: the influence of avatar androgyny, anthropomorphism and credibility in online interactions. *Computers in Human Behavior*, 24(4), pp. 1473-1493.

Oivatt, S. & Adams, B. (2000). Designing and evaluating conversational interfaces with animated characters. In: J. Cassell, J. Sullivan, S. Prevost & E. Churchill (Eds.), Embodied Conversational Agents, pp. 319-345. Cambridge, MA: MIT Press.

Picard, R. (1995). Affective computing. *M.I.T Media Laboratory Perceptual Computing Section Technical Report*, no 321.

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

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

Rommes, E., Overbeek, G., Engels, R.. & Scholte, R. (2007). 'I'm not interested in computers': gender-based occupational choices of adolescents. *Information, Communication & Society*, 10(3), pp. 299-319.

Ruttkay, Z., Dormann, C. & Noot, H. (2004). Embodied conversational agents on a common ground. In: Z. Ruttkay & C. Pelachaud (Eds.), *From Brows to Trust: Evaluating Embodied Conversational Agents*, pp. 27-66. Norwell, MA: Kluver Academic Publishers.

Sahlin, N-E., ·Wallin, A. & Persson, J. (2009). Decision science: from Ramsey to dual process theories. *Synthese*.

Schneider, D. (2003). *The Psychology of Stereotyping*. New York, NY: Guilford Press.

Smith, E. & Medin, D. (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.

Schneider, D., Hastorf, A. & Ellsworth, P. (1979). *Person Perception*. Reading, MA: Addison-Wesley.

Sheldon, W., Stevens, S. & Tucker, W. (1940). *The Varieties of Human Physique: An Introduction to Constitutional Psychology*. New York, NY: Harper & Brothers.

Thomas, F. & Johnston, O. (1984). *The Illusion of Life: Disney Animation*. New York, NY : Abbeville Press.

Tyler-Smith, K. (2006). Early attrition among first time eLearners: a review of factors that contribute to drop-out, withdrawal and non-completion rates of adult learners undertaking eLearning programmes. *J. of Online Learning and Teaching*, 2(2), pp. 73-85.

van Mulken, S., André, E. & Müller, J. (1998). The persona effect: how substantial is it? In: H. Johnson, L. Nigay, L. & C. Roast (Eds.), *People and Computers XIII: Proc. of HCI'98*, pp. 53-66. Berlin: Springer.

van Vugt, H. Konijn, E., Hoorn, J., Keur, I . & Eliëns, A. Realism is not all! User engagement with task-related interface characters. *Interacting with Computers*, 19, pp. 267-280.

Veletsianos, G. & Miller, C. (2008). Conversing with pedagogical agents: a phenomenological exploration of interacting with digital entities. *British J. of Educational Technology*, .39(6), pp. 969-986.

Voelker, D. (1994). *The Effects of Image Size and Voice Volume on the Evaluation of Represented Faces*. PhD Thesis, Stanford University, Palo Alto, CA.

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

Whiten, A. & Byrne, R. (Eds.) (1997). Machiavellian Intelligence II: Extensions and Evaluations. New York, NY: Cambridge University Press.

Paper I

Design of Animated Pedagogical Agents – a Look at Their Look

Gulz, A. & Haake, M. (2006)
Int. J. of Human-Computer Studies,
vol. 64, no. 4, pp. 322-339

doi:10.1016/j.ijhcs.2005.08.006
www.elsevier.com

Design of animated pedagogical agents – a look at their look

Agneta Gulz[a], Magnus Haake[b]

[a] *Division of Cognitive Science, Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden*

[b] *Department of Design Sciences, Lund University, P.O.Box 118, 221 00 Lund, Sweden*

Original paper published in:
Int. J. Human-Computer Studies, vol. 64, no. 4, (2006), pp. 322-339.

Abstract: A w ell-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 chose to enter the arena.

The view of learning as a social phenomenon is also evidenced in approaches in which an *already populated social arena* is created by technology – *intrinsically social* or *content oriented approaches*. Examples are scenario-based systems inhabited by social characters (Schank and Neaman, 2001) and systems using *animated pedagogical* agents, which is the focus of this article (e.g. André et al., 1998; Cassell et al., 1994; Paiva and Machado, 1998; Cassell and Tho´ risson, 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 peda-

[1] There 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).

gogical agent to the interface provides elements of embodiment, visibility and personality. In addition to the ability to communicate in an intelligent manner, a pedagogical agent should, according to Lester et al. (2001), have socio-emotive abilities and be lifelike. It should be visually present, by means of gestures, facial expressions and so on, and have a rich and interesting personality Lester et al., 2001. Consequently, the addition of animated agents to intelligent tutoring systems opens up the possibility for learners to have a personal relationship and an emotional connection with the agent, which in turn may promote interest in the learning task (Moreno et al., 2001). The addition of animated agents to intelligent tutoring systems can, in other words, be seen as an attempt to fulfil the need for a social context for learning in these systems (Kearsley, 1993).

Several educational systems and prototypes that incorporate various kinds of animated pedagogical agents exist today.[2] In some systems the animated agent primarily acts as *teacher*, *instructor* or *coach* – e.g. *AutoTutor* (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* (Oivatt and Adams, 2000); *Herman the Bug* (Lester et al., 1997); *Talking Head* (Moundridou and Virvou, 2002); *Laura* (Bickmore, 2003). In other systems the animated agent primarily acts as *a learning companion* – e.g. *Trouble Maker* (Aïmeur et al., 1997) and *Steve*[3] (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) (Figure 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

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

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

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

[5] As an illustration of the vagueness, *Gandalf* (Cassell and Tho´ risson, 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'.

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.

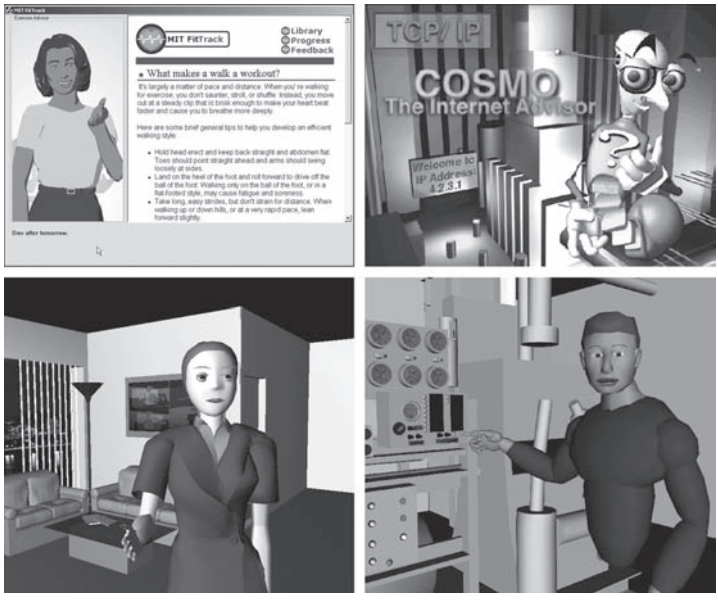


Figure 1. The agents Laura (upper left), Cosmo (upper right), Rea (lower left), and Steve (lower right). Laura: © Affective Computing Group, MIT Media Lab; Cosmo: © IntelliMedia, North Carolina State University; Rea: © GNL Group, MIT Media Lab; Steve: © CARTE, University of Southern California.

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 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.

[6] Compare Dehn and van Mulken (2000) who equate engagingness and entertainment value.

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.[7] It is within the second category that we point out the visual rendering or look of agents as a neglected design element.

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).

[7] 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.

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[8] are the following:

– *Movement characteristics*, in particular gestures, hand movements (Lester et al., 1998, 2000; Cassell et al., 2000; Massaro et al., 2000).

[8] We are not discussing design features related to the intelligent tutoring system that underlies the agent.

- *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.

3. A Look at Look in the Literature on Animated Pedagogical Agents

3.1. Personality

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

André et al. (2000) also explored look in relation to perception of the personality dimensions: disagreeable, agreeable, extrovert and introvert.[9] 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 person-

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

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

3.2. Facial expressions

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

Figure 2. The two faces used in the study of Walker et al. (1994). Left: pleasant/neutral expression; right: stern expression (compare the eyebrows). © Walker et al. (1994) & Association for Computing Machinery (ACM).

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

[10] The process may not be verbally articulated but well documented in the form of character descriptions and 'model sketches' (Thomas and Johnston, 1984).

same underlying image (Figure 2). (As mentioned previously, the stern face created a bad impression but was good for productivity.) Thus, the visual form and style of faces is not included in Walker et al.'s, 1994, empirical explorations. However, the authors do acknowledge the issue in posing some important questions: 'Should human facial realism be a goal? If so, whose face would appear?' Furthermore, they mention that it is established that attractive faces improve people's response to advertisements.

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

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

3.3. Appearance

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

Churchill et al. (1998, 2000) claim that the design of the character's appearance is central to charting out what design issues ought to be explored for situated conversational characters.[11] The authors explicitly mention the design of *the Look*[12] and propose five 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 staticity (i.e. the extent to which the

[11] Situated or embodied conversational agents are 'specifically conversational in their behaviour and specifically humanlike in the way they use their bodies in conversation' (Cassell et al., 2000).

[12] In (2000) the authors replaced the term 'look' with the term 'embodiment'.

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 involve computation. For instance, some of the research on design of gestures for agents is based on ethnographic and sociological research (O'Neill-Brown, 1997; Cassell, 2000; de Rosis et al., 2004). Nevertheless, most of the work on the design of gestures in agents strives for computational algorithms, and the same goes for work on facial expressions and voices. But there is hardly any corresponding work to be found on visual rendering.[13]

[13] One 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

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 animated agents, that is, on different faces, bodies and clothing; on degrees of iconicity in the visualization; and on visual styles. With respect to the pedagogically central potential of animated agents to be *engaging* – to increase involvement and contribute to the impact of learning activities – this research gap may be unfortunate. When people interact with real people, there is ample evidence that interpretations of appearance and observable physical cues profoundly affect both beliefs and behaviour (Milord, 1978; Schneider et al., 1979; McArthur, 1982; Kalick, 1988).

specific physical personalities. She expresses, however, the reservation that altering the physical personality of an agent in this way could lead to user confusion.

[14] These evaluations had been assessed in pre-studies.

[15] Swedish proverb. (*Döm inte hunden efter håren!*)

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

A possible explanation for the role of visual appearance is that it is significant for the representation of personality. Branham (2001) borrows the drama theory term *physical personality* of a character to refer to the aspects of appearance, which immediately and with no acquaintance, produce an impression of personality, and which initiate a set of attitudes and expectations. Among those aspects are many visual aspects such as shape, height, sex, race, physical attractiveness, hair, clothing, makeup, facial shape, facial hair, and so on. In contrast to how the wealth of such visual cues is immediately picked up in the encounter with another human being, the slow linear stream of spoken information is incredibly small[16] (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, Lassetter (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.' (Lassetter, 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 [...] still [be]

[16] Even though, of course, some voice characteristics are also quickly picked up.

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 Figure 3.

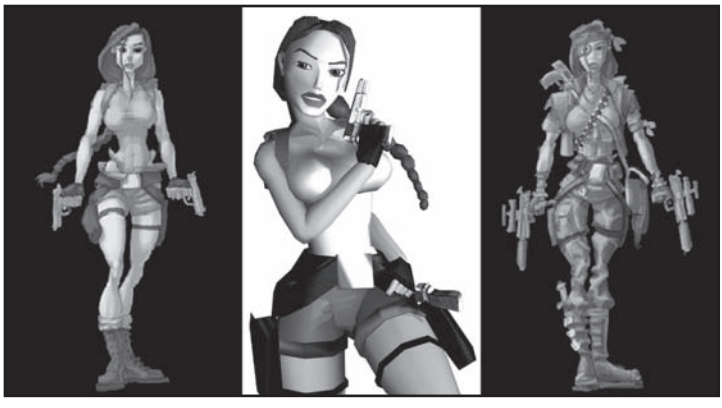


Figure 3. Lara Croft (middle) – main character of the game Tomb Raider. Left and right are two concept versions of Lara Croft illustrating the concept of a simplified, yet powerful design (left) with strong and immediate visual characteristics versus an overworked design (right) creating visual bewilderment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

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,[17] 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 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.'

[17] The reference is also frequently found in the literature on animated pedagogical agents.

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 (Figure 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]

Figure 4. Example of how 'degree of iconicity', i.e. the dimension 'iconic-realistic', can be visualized, from left to right: (highly) iconic, semi-iconic, and realistic (with a photorealistic representation corresponding to highly realistic). The illustrated figures (pedagogical agents) are taken from a user preference study (Gulz and Haake (2005)). © Magnus Haake.

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 (Figure 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 practi-

[18] At 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).

cal necessity – Disney has used it for decades with impressive results with when it comes to engagingness of characters.[19]

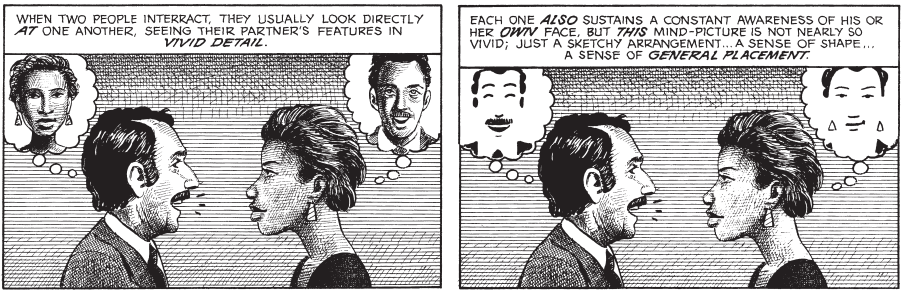


Figure 5. People who interact experience each other faces in a detailed, realistic manner, whereas they only have a schematic, iconic representation of their own faces (McCloud, 1993, p. 35–36). © Scott McCloud & HarperCollins Publishers, Inc.

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, 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.[20] 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

[19] And 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.

[20] For instance, the most common skin tone in recent years is semi-brown (Engholm and Michelsen, 1999).

studies is different user groups. There may be differences between cultures and sub-cultures in responses to realistic versus iconic animated agents.[21] 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 (Figure 6) are considered more naive, honest, warm and kind-hearted than others. They are also seen as weaker

[21] As a parallel regarding anthropomorphic versus non-anthropomorphic agents, O'Neill-Brown (1997) suggests that preferences may have a cultural basis. She refers to anecdotal evidence from the US market indicating that US users do not tend to favour anthropomorphic agents, while Japanese users prefer them.

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 Figure 6.

Figure 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 interact 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.

A third aspect found to have an impact is the facial expression of the neutral face. We are, of course, affected by dynamical facial expressions, such as smiles and frowns – but also by the expression of the face in a resting position. For instance, the lips naturally turn upwards on some faces in resting position. Such people are viewed more positively. They are considered friendly, kind, easygoing, and non-aggressive (Branham, 2001). In a similar vein 'faces that have features indicative of anger or hostility, e.g. low-lying eyebrows, thin lips, and withdrawn corners of the mouth, are perceived to be more threatening, aggressive, and dominant' (Branham, 2001, p. 3). Furthermore, similar facial expressions and an identical verbal utterance posed may be perceived as suggesting quite different motives depending upon the face of the person who is speaking. Branham (2001), reports a scenario where people are shown a photograph of a secretary and given a written description of the dialogue between the secretary and her employer, a young man. The secretary has just been thanked for taking a dictation and asks if there is anything else her employer would like her to do. When asked what the secretary

means by her question, people vary in their responses depending upon which photograph representing the secretary they are shown. 'Some faces consistently suggest seductive motives, others ambitious intentions, and some a polite way of taking leave.' (Branham, 2001, p. 3).

6.3. Body and costume design

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

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

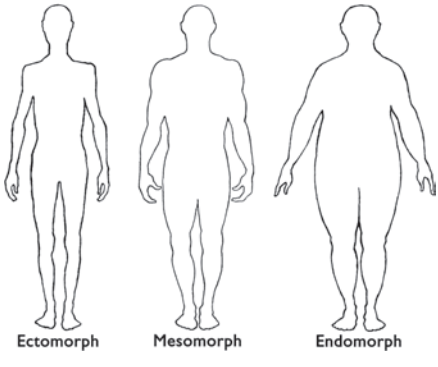


Figure 7. The three body stereotypes (somatotypes) defined by Sheldon, et al. (1940). From left to right: ectomorph: narrow shoulders and hips, thin narrow chest and abdomen, thin legs and arms, very little muscle or body fat; mesomorph: broad shoulders and narrow hips, muscular body, strong forearms and thighs; endomorph: wide hips and narrow shoulders, fat body, fat upper arms and thighs, slim wrists and ankles. © Magnus Haake.

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 (Figure 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 adventuresome, as well as less intelligent, more intolerant, and temperamental (Ryckman et al., 1991). In his book The Psychology of Stereotyping, Schneider (2003) reports on powerful

sub-types, where couch potatoes are rated negatively, while Santa Claus types are rated more positively, both belonging to the fat body type. Surprisingly enough, not many height stereotypes are reported, although this is an obvious feature of our physical presence. It should be noted that all related stereotypes concerning the body vary depending on gender, culture, fluctuations in the market, etc. For example, Schneider (2003), mentions a report about a shifting of the ideals of feminine beauty towards a more Rubenesque type in the US as the affluent 1920s turned into the depression of the 1930s (Fallon, 1990).

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

In the world of animation, Thomas and Johnston (1984), rarely explicitly articulate issues of how to design the figure and costume. But the issues are implicitly handled between the lines and in many of the examples illustrating the step-bystep sketching and development of different figures. An example is the many pictures illustrating the visual look of *the three Good Fairies* in *Sleeping Beauty*,where thereisalsoa 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 Figure 8.

Figure 8. Sketch of the three Good Fairies in Sleeping Beauty with their hats (Thomas and Johnston, 1984, p. 403). © Disney Enterprises, Inc.

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 Figure 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.

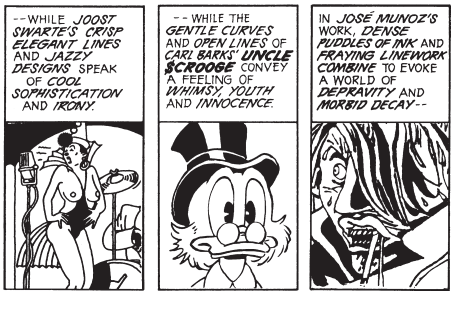


Figure 9. The simple inked line has tremendous visual power and almost unlimited possibilities (McCloud, 1993, p. 126). Left: © Joost Swarte; middle: © Disney Enterprises, Inc.; right: © José Munoz and Carlos Sampayo. Text by Scott McCloud (McCloud, 1993, p. 126). Reproduced with permission from HarperCollins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

Moving further to computer generated three-dimensional renderings there is actually no line at all, only the boundary between adjacent fields or volumes. Here visual style can be discussed in terms of form, colour, texture, lightning, 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 fre-

quent 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 engagingness (Laurel, 1993). In the context of animated pedagogical agents, the motivation to use and return to use a program is critical. This implies that visual experiences are too important to be treated as something to turn to when one has first solved other, more important issues.

We are certainly not denying that there *are other* fundamental and pressing issues that need to be solved.[22] But we maintain that visual rendering issues are pressing and fundamental, and need to be seriously approached as well. Arguments of the kind that 'We are now working on the underlying engine, we first need a car that is capable of running – after that we can think about the design of the chassis' are flawed, as the parallel does not hold. Function and visual appearance are more intimately connected with one another in the case of educational programs than in the case of cars. A car has the undisputed basic function of transporting a user from one place to another. Even if a car does not visually attract or appeal to a user – or to users at all – it still, as long as it runs, performs its most fundamental task. But an educational program based on animated agents that, because of poor visualization, does not at all appeal to a user – is not motivating, has little impact – might not be used at all, even if the underlying intelligent tutoring system is well functioning. Furthermore, within the area of computer games one is repeatedly advised not to start out with any programming before the visual aspects and the story are thoroughly worked out.[23] The conceptualization of visual aspects as make-up which may be added on the surface at the end is inadequate.

Without seriously addressing aesthetic and visual aspects in research on animated pedagogical agents, the gap between academic human-computer interaction in general and experience oriented issues risks to be reproduced in this area of animated pedagogical agents. For decades, academic human-computer interaction has focused rather narrowly on *usability* measured in terms of efficiency and effectiveness, (e.g. Norman, 1990; Shneiderman, 1992; Nielsen, 1994; Preece et al., 1994), and the research and development methods have typically not been broad enough to deal with social and psychological factors, and even less with aesthetic and visual aspects. This is understandable, as efficiency and effectiveness are more easily approached by methods of quantitative measurement. Recently,

[22] Not the least, issues on the intelligent tutoring system level, concerning the content of the support and the competence level of agents.

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

however, some of the central researchers in the area have begun to address aesthetics as an aspect of user experience (Nielsen, 2002; Norman, 2002; Preece et al., 2002). The lesson learned in the human-computer interaction domain is that with a narrow concept of usability, it is hard to gain credibility outside of pure academia. As Cloninger (2000) puts it, this limited approach has, among others, resulted in a sometimes affected schism between academics advocating usability and graphic designers concerned with aesthetics within the multimedia and web domains. The root of the problem might be that the design and marketing of information technology artefacts deals with a wide range of product qualities, including aspects of technical features, usability and pleasure[24] – and without methods to systematically handle the whole package, the marketing of products tends to focus on selling aspects of technical features and aesthetics, neglecting or twisting the knowledge gained in academic research, whether it is functionality, usability or something else.

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

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

Today much of the work is pursued in an engineering spirit, with work on implementing particular details of agents, with little regard for the goals of animated agents as a whole, that is, the goals of the design process. By asking *what kinds of animated pedagogical agents we actually want to achieve*, aspects of visual experiences will also be placed on the agenda. *Look* is too important to either

[24] In his proposal for *the new human factors*, Jordan (2000) describes a hierarchy of consumer needs, with functionality at the base, usability in the middle, and pleasure at the top. Pleasure is here defined as '[...] the emotional, hedonic and practical benefits associated with products' (Jordan, 1999).

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

Ahola, A. (Forthcoming). *Doctoral dissertation*. Dept. of Psychology, Stockholm University, Sweden.

Aïmeur, E., Dufort, H., Leibu, D., Frasson, C., 1997. Some justifications for the learning by disturbing strategy. In: *Proceedings of the Eight World Conference on Artificial Intelligence in Education*. IOS Press, Japan, pp. 119–126.

Alley, T., 1988. Social and applied aspects of face perception. In: Alley, T. (Ed.), *Social and Applied Aspects of Perceiving Faces, Resources for Ecological Psychology*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 1–10.

André , E., Rist, T., Müller, J., 1998. Integrating reactive and scripted behaviours in a life-like presentation agent. In: Sycara, K., Woolridge, M. (Eds.), *Proceedings of the Second International Conference on Autonomous Agents (AGENTS'98)*. ACM Press, New York, NY, pp. 261–268.

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

André , E., Rist, T., van Mulken, S., Klesen, M., Baldes, S., 2000. The automated design of believable dialogues for animated presentation teams. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 220–255.

Badler, N., Bindiganavale, R., Allbeck, J., Schuler, W., Zhao, L., Palmer, M., 2000. Parameterized action representation for virtual human agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 256–284.

Ball, G., Breese, J., 2000. Emotion and personality in a conversational agent. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 189–219.

Bardack, N., McAndrew, F., 1985. The influence of physical attractiveness and manner of dress on success in a simulated personnel decision. *Journal of Social Psychology*, 125, 777–778.

Bécheiraz, P., Thalmann, D., 1998. A behavioural animation system for autonomous actors personified by emotions. In: *Proceedings of the 1st Workshop on Embodied Conversational Characters (WECC'98)*. AAAI and ACM, Tahoe City, CA, pp. 57–65.

Berscheid, E., Walster, E., 1974. Physical attractiveness. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, 7. Academic Press, New York, NY, pp. 158–216.

Beskow, J., McGlashan, S., 1997. Olga – A conversational agent with gestures. In: *Proceedings of the IJCAI'97 Workshop on Animated Interface Agents: Making them Intelligent*. Nagoya, Japan.

Bickmore, T., 2003. *Relational agents: effecting change through human-computer relationships*. MIT Ph.D. Thesis, MIT Media Lab, Cambridge, MA. http://web.media.mit.edu/bickmore/bickmore-thesis. pdf (2005-07-18).

Branham, S., 2001. Creating physical personalities for agents with faces: modeling trait impressions of the face. In: *Proceedings of the UM2001 Workshop on Attitudes, Personality and Emotions in User-Adapted Interactions*. Sonthofen, Germany http://www.di.uniba.it/intint/sonthofen/brahnam-f.pdf (2005-07-18).

Brigham, J., 1980. Limiting conditions of the 'physical attractiveness stereotype': attributions about divorce. *Journal of Research in Personality*, 14, 365–375.

Butler, C., Ryckman, R., Thornton, B., Bouchard, R., 1993. Assessment of the full content of physique stereotypes with a free-response format. *Journal of Social Psychology*, 133, 147–162.

Cassell, J., 2000. Nudge nudge wink wink: elements of face-to-face conversation for embodied conversational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 1–27.

Cassell, J., Tho´ risson, K., 1999. The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, 13, 519–538.

Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Beckett, T., Douville, B., Prevost, S., Stone, M., 1994. Animated conversation: rule-based generation of facial expression, gesture, and spoken intonation for multiple conversational agents. In: *Proceedings of the 21st Annual International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'94)*. ACM Press, New York, NY, pp. 413–420.

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhja´ lmsson, H., Yan, H., 1998. An architecture for embodied conversational characters. In: *Proceedings of the 1st Workshop on Embodied Conversational Characters (WECC'98)*. AAAI and ACM, Tahoe City, CA, pp. 21–29.

Cassell, J., Bickmore, T., Campbell, L., Vilhja´ lmsson, H., Yan, H., 2000. Human conversation as a system framework: designing embodied conversational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 29–63.

Churchill, E., Prevost, S., Bickmore, T., Hodgson, P., Sullivan, J., Cook, L., 1998. Design issues for situated conversational characters. In: *Proceedings of the 1st Workshop on Embodied Conversational Characters (WECC098)*. AAAI and ACM, Tahoe City, CA, pp. 149–158.

Churchill, E., Cook, L., Hodgson, P., Prevost, S., Sullivan, J., 2000. 'May I help you?': designing embodied conversational agent allies. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 64–94.

Cloninger, C., 2000. Usability experts are from Mars, graphic designers are from Venus. *A List Apart*, http://www.alistapart.com/articles/ marsvenus/ (2005-07-18).

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

de Rosis, F., Pelachaud, C., Poggi, I., 2004. Transcultural believability in embodied agents: a matter of consistent adaptation. In: Payr, S., Trappl, R. (Eds.), *Agent Culture: Designing Human-Agent Interaction in a Multicultural World*. Laurence Erlbaum Associates, Mahwah, NJ, pp. 75–105.

Dillenbourg, P., 1999. *Collaborative Learning: Cognitive and Computational Approaches*. Pergamon, Oxford.

Disney-Pixar, 2003. *Finding Nemo*. Feature Film 2003 from Disney-Pixar, http://www.pixar.com/featurefilms/nemo/index.html (2005-07-18).

Dryer, D., 1999. Getting personal with computers: how to design personalities for agents. *Applied Artificial Intelligence*, 13, 273–295.

Elman, D., 1977. Physical characteristics and the perception of masculine traits. *Journal of Social Psychology*, 103, 157–158.

Engholm, I., Michelsen, A., 1999. *Designmaskinen – design af den moderne verden*. Gyldendahl, Copenhagen, Denmark.

Fallon, A., 1990. Culture in the mirror: sociocultural determinants of body image. In: Cash, T., Pruzinsky, T. (Eds.), *Body Images: Development, Deviance, and Change*. Guilford Press, New York, NY, pp. 80–109.

Gamasutra, 2004. Features. *Gama Network: http://www.Gamasutra. com*, http://www.gamasutra.com/php-bin/ article_display.php (2005-07-18).

Gard, T., 2000. Building character. *Gama Network: www.Gamasutra. com*, http://www.gamasutra.com/features/20000720/gard_pfv.htm (2005-07-18).

Gulz, A., Haake, M., (2005). Social and Visual Style in Virtual Pedagogical Agents. In: *Proceedings of the Workshop on Adaptation to Affective Factors, in conjunction with the 10th International Conference on User Modelling (UM'05)*. Edinburgh, Scotland, 24–29 July, 2005 (in press).

Hellström, Å., Tekle, J., 1994. Person perception through facial photographs: effects of glasses, hair, and beard on judgments of occupation and personal qualities. *European Journal of Social Psychology*, 24, 693–705.

Hietala, P., Niemirepo, T., 1998. The competence of learning companion agents. *International Journal of Artificial Intelligence in Education*, 9, 178–192.

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

Isbister, K., Nass, C., 1998. Personality in conversational characters: building better digital interaction partners using knowledge about human personality preferences and perceptions. In: *Proceedings of the 1st Workshop on Embodied Conversational Characters (WECC'98)*. AAAI and ACM, Tahoe City, CA, pp. 105–113.

Iwawaki, S., Lerner, R., 1976. Cross-cultural analyses of body-behavior relations: III. Developmental intra-and inter-cultural factor congruence in the body build stereotypes of Japanese and American males and females. *Psychologia*, 19, 67–76.

Johnson, K., Roach-Higgins, M., 1987. Dress and physical attractiveness of women in job interviews. *Clothing and Textiles Research Journal*, 5, 1–8.

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

Jones, W., Hansson, R., Phillips, A., 1978. Physical attractiveness and judgments of psychotherapy. *Journal of Social Psychology*, 105, 79–84.

Jordan, P., 1999. Pleasure with products: human factors for body, mind and soul. In: Green, W., Jordan, P. (Eds.), *Human Factors in Product Design: Current Practice and Future Trends*. Taylor & Francis, London, pp. 206–217.

Jordan, P., 2000. *Designing Pleasurable Products: An Introduction to the New Human Factors*. Taylor & Francis, London.

Kaiser, S., 1985. *The Social Psychology of Clothing and Personal Adornment*. Macmillan, New York, NY.

Kalick, S., 1988. Physical attractiveness as a status cue. *Journal of Experimental Social Psychology*, 24, 469–489.

Kearsley, G., 1993. Intelligent agents and instructional systems: implications of a new paradigm. *Journal of Artificial Intelligence in Education*, 4, 295–304.

Keller, J., 1983. Motivational design of instruction. In: Reigeluth, C. (Ed.), *Instructional Design, Theories and Models: An Overview of their Current Status*. Erlbaum, Hillsdale, NJ, pp. 383–434.

Kerlow, I., 2003. *The Art of 3-D Computer Animation and Effects (Third ed.)*. John, New York, NY.

Koda, T., 1996. *Agents with faces: a study on the effect of personification of software agents*. Master's thesis, MIT Media Lab, Cambridge, MA.

Koda, T., Maes, P., 1996. Agents with faces: the effects of personification. In: *Proceedings of the Fifth IEEE International Workshop on Robot and Human Communication (RO-MAN'96)*. Tsukuba, Japan, pp. 189–194.

Koschman, T., 1996. *CSCL: Theory and Practice of an Emerging Paradigm*. Erlbaum, Mahwah, NJ.

Langlois, J., Kalakanis, L., Rubenstein, A., Larson, A., Hallam, M., Smoot, M., 2000. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390–423.

Lassetter, J., 1987. In: *Principles of traditional animation applied to 3D computer animation, ACM SIGGRAPH Computer Graphics*, 21. ACM Press, New York, NY, pp. 35–44.

Laurel, B., 1993. *Computers as Theatre*. Addison-Wesley, Reading, MA.

Laurillard, D., 1993. *Rethinking University Teaching: A Framework for the Effective Use of Educational Technology*. Routledge, London.

Lee, E.-J., Nass, C., 1998. Does the ethnicity of a computer agent matter? An experimental comparison of human-computer interaction and computer-mediated communication. In: *Proceedings of the 1st Workshop on Embodied Conversational Characters (WECC '98)*. AAAI and ACM, Tahoe City, CA, pp. 123–128.

Lester, J., Stone, B., Converse, S., Kahler, S., Barlow, S., 1997. Animated pedagogical agents and problem-solving effectiveness: a large-scale empirical evaluation. In: *Proceedings of the Eighth World Conference on Artificial Intelligence in Education*. Kobe, Japan, pp. 23–30.

Lester, J., Towns, S., Callaway, C., Fitzgerald, P., 1998. Deictic and emotive communication in animated pedagogical agents. In: *Proceedings of the First Workshop on Embodied Conversational Characters (WECC'98)*. AAAI and ACM, Tahoe City, CA, pp. 67–76.

Lester, J., Zettlemoyer, L., Grégoire, J., Bares, W., 1999. Explanatory lifelike avatars: performing user-designed tasks in 3d learning environments. In: *Proceedings of the Third International Conference on Autonomous Agents*. ACM Press, New York, NY, pp. 24–31.

Lester, J., Towns, S., Callaway, C., Voerman, J., Fitzgerald, P., 2000. Deictic and emotive communication in animated pedagogical agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 123–154.

Lester, J., Callaway, C., Grégoire, J., Stelling, G., Towns, S., Zettlemoyer, L., 2001. Animated pedagogical agents in knowledge-based learning environments. In: Forbus, K., Feltovich, P. (Eds.), *Smart Machines in Education*. AAAI/MITPress, Menlo Park, CA, pp. 269–298.

Liggett, J., 1974. *The Human Face*. Stein and Day, New York, NY.

Lindegaard, G., Dudek, C., 2003. What is this evasive beast we call user satisfaction? *Interacting with Computers*, 15, 429–452.

Magli, P., 1989. The face and the soul. In: Feher, M. (Ed.), *Fragments for a History of the Human Body: Part Two*. Zone, New York, NY, pp. 466–489.

Malone, T., 1981. Toward a theory of intrinsically motivating instruction. *Cognitive Science*, 4, 333–369.

Massaro, D., Cohen, M., Daniel, S., Cole, R., 1998. Developing and evaluating conversational agents. In: *Proceedings of the 1st Workshop on Embodied Conversational Characters (WECC098)*. AAAI and ACM, Tahoe City, CA, pp. 137–147.

Massaro, D., Cohen, M., Beskow, J., Cole, R., 2000. Developing and evaluating conversational agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 287–318.

Mathes, E., 1975. The effects of physical attractiveness and anxiety on heterosexual attraction. *Journal of Marriage and the Family*, 37, 769–773.

McArthur, L., 1982. Judging a book by its cover: a cognitive analysis of the relationship between physical appearance and stereotyping. In: Hastorf, A., Isen, A. (Eds.), *Cognitive Social Psychology*. Elsevier, New York, NY, pp. 149–209.

McCauley, L., Gholson, B., Xiangen, H., Graesser, A., 1998. Delivering smooth tutorial dialogue using a talking head. In: *Proceedings of the First Workshop on Embodied Conversational Characters (WECC'98)*. AAAI and ACM, Tahoe City, CA, pp. 31–37.

McCloud, S., 1993. Understanding Comics. HarperPerennial, New York, NY.

Milord, J., 1978. Aesthetic aspects of faces: a (somewhat) phenomenological analysis using multidimensional scaling methods. *Journal of Personality and Social Psychology*, 36, 205–216.

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

Moundridou, M., Virvou, M., 2002. Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning*, 18, 253–261.

Mulford, M., Orbell, J., Shatto, C., Stockard, J., 1998. Physical attractiveness, opportunity, and success in everyday exchange. *American Journal of Sociology*, 103, 1565–1592.

Nass, C., Gong, L., 1999. Maximized modality or constrained consistency. IN: *Proceedings of the AVSP'99 Conference*. Santa Cruz, CA, pp. 1–5.

Nass, C., Steuer, J., Tauber, E., 1994. Computers are social actors. In: *Proceedings of the CHI 1994 Conference on Human Factors in Computing Systems (CHI'94)*. ACM Press, New York, NY, pp. 72–78.

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

Nielsen, J., 1994. *Usability Engineering*. Morgan Kaufman, San Francisco, CA.

Nielsen, J., 2002. User empowerment and the fun factor. *Jacob Nielsen's Alertbox*. http://www.useit.com/alertbox/20020707.html (2005-07-18).

Noma, T., Badler, N., 1997. A virtual human presenter. In: *Proceedings of the IJCAI'97 Workshop on Animated Interface Agents: Making them Intelligent*. Nagoya, Japan, pp. 45–51.

Norman, D., 1990. *Design of Everyday Things*. Currency and Doubleday, New York, NY.

Norman, D., 2002. *Emotion & design: attractive things work better. Interactions*, 9, 36–42.

O'Neill-Brown, P., 1997. Setting the stage for a culturally adaptive agent. In: Dautenhahn, K. (Ed.), *Socially Intelligent Agents, Papers from the 1997 AAAI Fall Symposium, Technical Report FS-97-02*. AAAI Press, Menlo Park, CA, pp. 93–97.

Oivatt, S., Adams, B., 2000. Designing and evaluating conversational interfaces with animated characters. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 319–345.

Paiva, A., Machado, I., 1998. Vincent, an autonomous pedagogical agent for on-the-job training. In: Goettl, B., Halff, H., Redfield, C., Shute, V. (Eds.), *Proceedings of the Fourth International Conference on Intelligent Tutoring Systems*. Springer, Berlin, pp. 584–593.

Parise, S., Sproull, L., Kiesler, S., Waters, K., 1996. My partner is a real dog: cooperation with social agents. In: *Proceedings of the ACM 1996 Conference on Computer Supported Cooperative Work (CSCW'96)*. ACM Press, New York, NY, pp. 399–408.

Poggi, I., Pelachaud, C., 2000. Performative facial expressions in animated faces. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), *Embodied Conversational Agents*. MIT Press, Cambridge, MA, pp. 155–188.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T., 1994. *Human-Computer Interaction*. Addison-Wesley, Reading, MA.

Preece, J., Rogers, Y., Sharp, H., 2002. *Interaction Design*. John, New York, NY.

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

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

Rickenberg, R., Reeves, B., 2000. The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In: *Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems (CHI'00)*. ACM Press, New York, NY, pp. 49–56.

Ryckman, R., Robbins, M., Thornton, B., Kaczor, L., Gayton, S., Anderson, C., 1991. Public self-consciousness and physique stereotyping. *Personality & Social Psychology Bulletin*, 17, 400–405.

Schank, R., Neaman, A., 2001. Motivation and failure in educational simulation design. In: Forbus, K., Feltovich, P. (Eds.), *Smart Machines in Education*. AAAI/MITPress, Menlo Park, CA, pp. 37–69.

Schneider, D., 2003. *The Psychology of Stereotyping*. Guilford Press, New York, NY.

Schneider, D., Hastorf, A., Ellsworth, P., 1979. *Person Perception*. Addison-Wesley, Reading, MA.

Shaw, E., Ganeshan, R., Johnson, W., Millar, D., 1999. Building a case for agent-assisted learning as a catalyst for curriculum reform in medical education. In: Lajoie, S., Vivet, M. (Eds.), *Proceedings of the Ninth World Conference on Artificial Intelligence in Education AIED'99, Frontiers in Artificial Intelligence and Applications*. IOS Press, Amsterdam, pp. 509–516.

Sheldon, W., Stevens, S., Tucker, W., 1940. *The Varieties of Human Physique: An Introduction to Constitutional Psychology*. Harper, New York, NY.

Shneiderman, B., 1992. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley, Reading, MA.

Sproull, L., Subramani, M., Kiesler, S., Walker, J., Waters, K., 1997. When the interface is a face. In: Friedman, B. (Ed.), *Human Values and the Design of Computer Technology*. Cambridge University Press, New York, NY, pp. 163–190.

Takeuchi, A., Naito, T., 1995. Situated facial displays: towards social interaction. In: Katz, I., Mack, R., Marks, L., Rosson, M., Nielsen, J. (Eds.), *Human Factors in Computing Systems: CHI'95 Conference Proceedings*. ACM Press, New York, NY, pp. 450–455.

Taylor, I., McInnes, F., Love, S., Foster, J., Jack, M., 1998. Providing animated characters with designated personality profiles. In: *Proceedings of the 1st Workshop on Embodied Conversational Characters (WECC'98)*. AAAI and ACM, Tahoe City, CA, pp. 89–94.

Terry, R., Krantz, J., 1993. Dimensions of trait attributions associated with eyeglasses, men's facial hair, and women's hair length. *Journal of Applied Social Psychology*, 23, 1757–1769.

Thomas, F., Johnston, O., 1984. *The Illusion of Life: Disney Animation*. Abbeville Press, New York, NY.

Tractinsky, N., Katz, A., Ikar, D., 2000. What is beautiful is usable. *Interacting with Computers*, 13, 127–145.

van Mulken, S., André ,E., Müller, J., 1998. The persona effect: how substantial is it? In: Johnson, H., Nigay, L., Roast, C. (Eds.), *People and Computers XIII: Proceedings of HCI'98*. Springer, Berlin, pp. 53–66.

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

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

Zebrowitz, L., 1997. *Reading Faces: Window to the Soul?* Westview Press, Boulder, CO.

Paper II

Visual Stereotypes and Virtual Pedagogical Agents

Haake, M. & Gulz, A. (2008)
Educational Technology & Society
vol. 11, no. 4, pp. 1-15

www.ifets.info

Visual Stereotypes and Virtual Pedagogical Agents

Magnus Haake[a], Agneta Gulz[b]

[a] *Department of Design Sciences, Lund University, P.O.Box 118, 221 00 Lund, Sweden*

[b] *Division of Cognitive Science, Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden*

Original paper published in:
Educational Technology & Society, vol. 11, no. 4, (2008), pp. 1-15.

Abstract: The paper deals with the use of visual stereotypes in virtual pedagogical agents and its potential impact in digital learning environments. An analysis of the concept of visual stereotypes is followed by a discussion of affordances and drawbacks as to their use in the context of traditional media. Next, the paper explores whether virtual pedagogical characters introduce anything novel with regard 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, television and video. It is proposed that novel affordances, as well as novel drawbacks, indeed are being introduced with the use of visual stereotypes in virtual characters. The conclusion of the paper is that knowledge on these matters can be useful both for developers of educational systems and for educators in enabling them to strengthen some pedagogical settings and activities.

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

1. 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 in the roles of virtual medical counsellors, physical exercise coaches and guides on city homepages, and they also appear in edutainment and infotainment settings. Furthermore, virtual characters are essential ingredients in digital environments used by young people, such as (of course) games and edutainment applications as well as chat systems and mobile phone applications (e.g. *imvu* at http://www.imvu.com).

According to Moreno & Flowerday (2006), we see more and more frequently that effective multimedia lessons are combined with the presence of virtual pedagogical agents. It is a qualified guess that various forms of digital educational environments and systems will increasingly incorporate these characters – virtual instructors, learning companions, coaches, mentors, etc. – concurrently with tool-kits for designing characters becoming increasingly available and widespread.

Furthermore, in recent years there has been a growing focus on the social dimension of human interaction with virtual pedagogical agents, with a variety of social competences being simulated in various agents (Baylor & Kim, 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 & Nass (1996) based on extensive 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 & 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. These 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 little attended to in research on virtual agents (Gulz & Haake, 2006a; Gulz & Haake, 2006b).

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 & Walster (1974) note that, "… our appearance telegraphs more information about us than we would care to reveal on a battery of personality inventories, intelligence tests, and character scales. From flame-coloured hair through flat feet, few aspects of appearance fail to provide kernels of folk insight into another's nature." (ibid., p. 159).

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, roughly standing for a culturally shared socio-cognitive schema that summarizes our beliefs of other people and acting as a default setting in social perception processes. Since our focus is 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 and the potential impact of such use in educational technology with respect to teaching and learning. First, 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 an attempt to answer this question, we also present and discuss some of the (few) empirical studies carried out in the area. Second, 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 drawbacks. Before entering these two issues, we provide a background by analysing the concept of *visual stereotype*, relating it to the concept of *visual prototype*, and discussing visual stereotyping with respect to *visual naturalism* versus *visual stylisation*.

2. The Concept of Visual Stereotype

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.

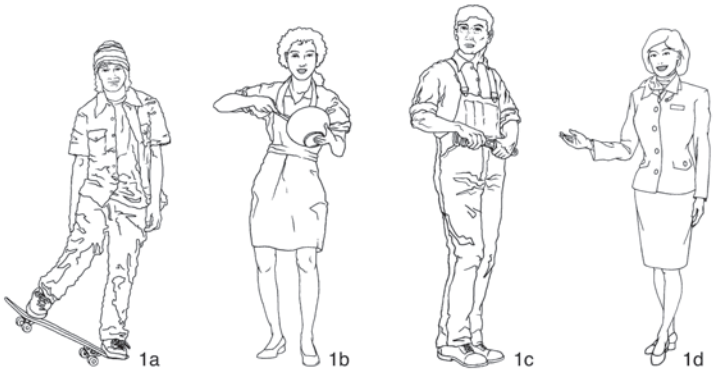


Figure 1. Examples of visual stereotypes.

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.

3. 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 be-

ing prototypical is the frequency of occurrences in human perceptions and experiences, and in that sense a prototype relates to something "really there". A stereotype, on the other hand, is often associated with something that culture, media, etc. has constructed out of dubious starting points but where there is not something "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 Khaki 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 (or men) on the average are stupider than non-blondes.

4. 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. a cartoonish style or other artistic style that modifies and often simplifies a person's appearance (Gulz & Haake, 2006b).

Figure 2. Visual stereotypes: naturalism versus stylization.

Stylisation makes is 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 movies. 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 stage can, furthermore, reinforce the appearance of the artists on the stage as visually stereotyped. In Commedia dell'Arte and classical Chinese opera, for example, dresses and make-up as well as gestures are pronouncedly stylised (see Figure 3). Also movies 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 movies to predict the roles of characters and who is going to die or survive on the basis of the visual stereotyping.

Figure 3. Left: *Harlequin* (Commedia dell'Arte), middle: Chinese opera actor, and right: *Indiana Jones* (actor Harrison Ford).

5. Visual Stereotypes in Traditional Media Characters as well as in Virtual Characters

5.1. 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 part of our social autopilot as an essential navigation tool in a social environment that would otherwise be overwhelmingly complex and demand a practically insurmountable burden of processing (Smith & Medin, 1981). Notably we speak here of very quick and largely unconscious processes.

If we start out to consider traditional visual media – theatre, film, comics – it is indeed essential that a reader or a spectator is 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 character-based digital learning environments may provide starting points both for the interpretation of the 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). Inconsistencies engender disturbance and distraction, which increase the demand on cognitive resources. On the other hand, with a visual appearance that corresponds to behaviour and personality predictions, users get their expectations acknowledged and leveraged, and interaction becomes smooth and efficient.

However, choosing an *adequate visual stereotype* can be difficult as exemplified by the following three examples. In the first example, provided by de Rosis et

al. (2004), a virtual character was to be designed for a digital legal information system in Italy. Initially the virtual character was modelled upon 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 realizing that the lawyer's (female) secretary was the one who most frequently used the system, the designers became aware that the appearance and behaviour of the virtual 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 instead disturbed and distracted them. Furthermore, *visual* redesign was indeed required. It would not have been sufficient to redesign the dialogue and behaviour of the character but leave the visual appearance intact.

Figure 4. The virtual city guide *Sara* (http://www.malmo.se/).

A second example involves two different design cases of virtual assistants for city home pages. In Botkyrka, a Swedish community with a high percentage of immigrants, a stereotypically "Swedish-looking" light-blond female character was introduced as a virtual assistant on the website. However, due to negative comments from site visitors the character was removed and redesigned. This virtual character turned out *not* to be an adequate visual starting point for a visitor entering the site and its content. (In this case the content with its structure and pedagogical design may be perfectly well-designed but nevertheless of no benefit for visitors.) In contrast, in the Swedish city of Malmö, equally 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 underlying the visual design of the Sara character were extensive.

A third example is provided by Baylor's research group at *Research of innovative technology for learning* (RITL). Within a programme for gender equity in science and engineering, Baylor directs the project *Challenging Stereotypes toward Engineering with Pedagogical Agents*. The project investigates, among other things, the influence of character appearance on female choice of engineering subjects. Several empirical studies within the project pointed in the same direction. Young

female students were more strongly affected, as to their motivation and self-efficacy regarding engineering subjects, by virtual coaches that were similar to themselves or similar to how they would like to be – in this case female, young and cool (Figure 5, left). Using such characters as coaches and instructors in tutorials on technology also seemed to increase the willingness of female students to select courses with technical content (Baylor, 2005). However, in an additional study (Baylor & Rosenberg-Kima, 2006) it was seen that even though, again, female students who interacted with a peer model character (female, young and cool) showed a more positive view of, and attitude towards, "an engineer", the outcome was different as to influences on their attitudes regarding the importance and utility of engineering-related fields. Here the young female students were significantly more influenced by a male, older, un-cool "stereotypical engineer" character (Figure 5, right). The researchers tentative conclusion is that perhaps the "… most effective approach would be to use multiple agents (e.g., have a stereotypical engineer and a peer model both interact with participants)." (ibid., p. 6).

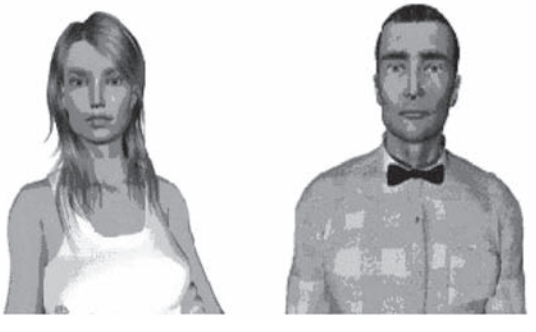


Figure 5. Virtual coaches from (Baylor & Rosenberg-Kima, 2006).

However, even if visual design choices can be difficult to make, we think it is important that educational systems developers acknowledge the influences of visual design, and in particular visual stereotypes. This may not come easily. Traditionally, computer science related domains have *not* acknowledged visual design as important in relation to the "real thing" which is the computer systems behind the – visually designed – interface. Of course the system must be wrapped up and presented in some way, but basically the system stands for itself. For virtual pedagogical agents, specifically, the argument goes that what really matters is the behaviour, the dialogue, the movements, the facial expressions and the pedagogical role of such an agent (Gulz & Haake, 2006a). The quality of these aspects determine to what extent the pedagogical goals set for the agent will be fulfilled.

Then, indeed, the character needs to have a visual look or appearance – a skin colour a hair-cut, a body (shape), some clothing, etc. – but these are considered surface aspects with no real impact on the fulfilment pedagogical goals of the agent or of the learning environment that it inhabits.

This way of reasoning, we argue, is mistaken. On the one hand, it certainly is the case that without the development and refinement of algorithms and modules of behaviour, dialogue and pedagogy there would be *no virtual pedagogical agents at all*. On the other hand, no matter how adequate and well designed these fundamental aspects are – if the agent's visual appearance is inadequate, the pedagogical benefits may decrease considerably. By now there is substantial evidence that learners' expectations, attitudes, understanding and motivation in various ways are affected by the visual design of a virtual pedagogical agent. It may influence the following: learners' beliefs in their own competence in approaching a certain subject matter, their willingness to pay attention to a presentation or tutorial, the extent to which they find something trustworthy or relevant, how hard they try to understand a material, and so on (Baylor & Plant, 2005; Graesser et al., 2004; Gulz et al., 2007b; Massaro, 2004; Moreno et al., 2001).

Furthermore, this does not refer only to "certain rare cases", which involve explicit or apparent visual stereotypes. Stereotypes and stereotypical elements are pervasive in human cognition, and features such as gender, age, ethnicity, clothing, etc., must be decided on in almost all cases of a virtual character. After a series of studies involving such visual elements, Moreno & Flowerday (2006) argue that the choice of an agent's visual appearance is practically always psychologically loaded, and put this empirically based conclusion in contrast to how "… the vast majority of instructional interfaces assign arbitrary animated pedagogical agents assuming that the choice of an agent representation is psychologically neutral." (ibid., p. 191).

Even *subtle* cues as to visual stereotypes can influence learners' experiences and the way they assimilate a given content in a digital context. This is shown, for instance, in (Gulz et al., 2007a) that presents a parallel study to one by Voelker (1994) and also described in Reeves & Nass (1996). The Voelker (1994) study compared user evaluations of two female presenters, where one spoke in a more stereotypically *feminine* voice than the other. Results were that the presenter with the more feminine voice, and thus also the content of her presentation, was evaluated significantly lower on trustworthiness and intelligence, but significantly higher on warmth and empathy. In other words, relatively subtle voice cues evoked evaluations in line with well-known gender stereotypes. The present study (Gulz et al., 2007a) instead manipulated the degree of femininity

of female virtual characters via *visual* cues. One character was designed as more (stereo)typically feminine and the other character as less (stereo)typically feminine by varying visual cues as to degree of femininity (shape of head, hairstyle and makeup) (Figure 6). The characters were also pre-validated in order to make sure that they were indeed perceived in this way. Regarding non-visual cues, both characters were identical as to their professional role as medical doctors, their voices and their lecturing on shift work and health. Again, the visual cues as to the degree of femininity influenced users' evaluation of the characters, and thus the content of their lectures, in accordance with gender stereotypes.

Figure 6. Virtual characters (presenters) from Gulz et al. (2007a).

It is important to remember that when *asked*, most people deny that one or another visual stereotype cue could make any difference for how they experience a presented material or for their attentiveness, etc. When told that this is the case, it can still be hard to believe or admit, possibly because it does not fit with the concept of human beings as rational and capable of identifying good content regardless of its form. Nevertheless, empirical evidence shows that we are all sensitive to visual stereotypes in the sense that on an *unconscious* cognitive level they influence our judgments and interpretations of informational settings that we encounter (Cook, 1979). Those involved in designing educational material and in teaching can profit from acknowledging and learning more about the mechanisms and processes involved.

5.2. Visual stereotypes – drawbacks

A visual stereotype may – by its nature – activate misleading expectations. Even if based on some kind of frequency distributions of property-clusters in peoples' experiences and thus corresponding to "actuality" in a statistical sense (i.e. a prototype), a visual stereotype can in a given instance be inadequate and misleading. In real life an example could be a youngster whose street fashion look signals "tough, rebellious and cheeky" but who actually is very kind and helpful.

Correspondingly, a badly casted visual stereotype 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 fuss around the choice of James Bond actors, where some last for only one production, while others reappear again and again and become more or less synonymous to the role. Even though there are many different variables involved, the visual appearances play an important 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 (which was to welcome and invite everyone to use it, and to boost inhabitants' feelings of belonging to the city in question.)

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 spectacled and somewhat thin craftsman is "no real craftsman". In this way, visual stereotypes can hide or suppress nuances and an existing manifold. There is also the aspect of self-reproduction and self-reinforcing of stereotypes due to the close interactions between media and "real life. 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." (Next Generation, 1998, 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 the use of visual stereotypes, where what is known from traditional media reappears in the virtual world. We now proceed towards what virtual characters bring in as novel.

6. Novel Affordances and Risks with Virtual Characters and Visual Stereotypes

6.1. 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 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 modelling. 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 this out 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).

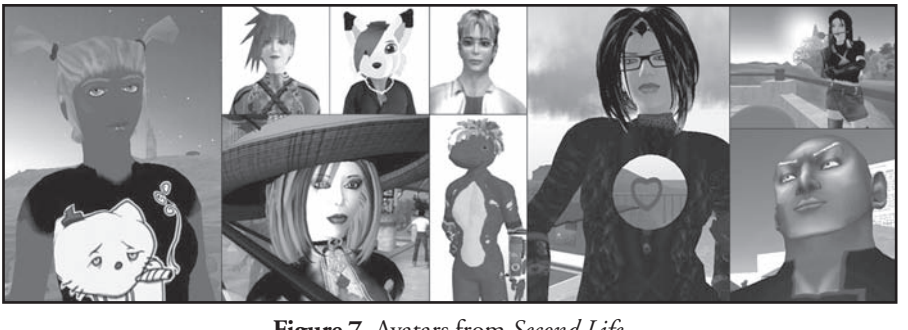


Figure 7. Avatars from *Second Life*.

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, professional roles, social classes, gender, and so on.

Figure 7 shows some avatars from the on-line world *Second Life* (http://secondlife.com), of which some are used to explore alternative gender and personality.

Offering a broader range of styles and identities may, furthermore, enable social identification and role modelling for a larger number of students. In turn, identification and opportunities for role modelling are known to strengthen development in pedagogical terms (Bandura, 1977). Also note the *specific* potentials of the digital virtual world. Here we can have dynamic, interactive situations, involving exploration and feedback that are not possible in traditional, non-interactive media. An example outside of education in a narrower sense, but belonging to the broader pedagogical domain, is a virtual character system for young women with eating disorders. The virtual coach character in question will start out looking really thin – in order for the clients to identify with, find trustworthy and be inclined to interact with – but then over time transform visually towards a more normal weight young female. In this sense, by slowly manipulating he stereotype, it may be possible to help the client to get away from an unhealthy mental visual stereotype.

But apart from the situation where *developers* may design virtual characters visually in a knowledgeable way in order to reach certain educational and pedagogical goals, one can also imagine the situation where learners themselves get to design their virtual instructors or learning companions. Where they, themselves or in a group, decide on the ethnicity, gender, body shape, clothing style, etc. for a virtual pedagogical character inhabiting a certain digital learning environment. "What is this instructor going to look like? Who is it going to be?" – Such a situation will be familiar for many young people used to games such as *the Sims*, avatars in on-line chats, etc. Furthermore, this situation can be set up both through commercial educational systems that provide character design kits, and through digital learning materials put together from scratch by teachers and students themselves, using character toolkits already becoming available (e.g. *PeoplePutty* at http://www.haptek.com/peopleputty and *Meez* at http://www.meez.com).

In both kinds of situations, a human pedagogue with knowledge of the impact of visual stereotypes will have an opportunity to use the situation as a basis for reflection and discussion: Why do we choose this character in this role – coach, learning companion, instructor, for different subject domains? Whose appearance shall be exhibited, in terms of gender, age, ethnicity, class, regional subgroup, etc? (Voice, in terms of gender, dialect and sociolect is another design feature that one will probably be able to choose). Which visual features do we find easy to combine and which not? What can be lost and what can be gained by choosing/designing a less naturalistic character?

Various alternatives may be suggested and explored. We believe that this kind of active and dynamic situation can provide a natural and powerful basis for reflection and discussion – more so than a standard and often more disconnected classroom discussion on stereotypes. This could be a rich and flexible tool for a pedagogue interested in challenging prejudices and proposing reconstruction of roles. Since humans are perceptual creatures and are powerfully affected by perceptual input and materials, this can be a vigorous complement to attempts at verbal reconstruction.

6.2. Dilemmas with novel possibilities

The idea of *breaking with* visual stereotypes for pedagogical purposes that we have repeatedly lifted forth can, however, be in conflict with the pedagogical *exploiting of* visual stereotypes to facilitate smooth and efficient interaction. As observed earlier in the text, the use of visual stereotypes may enable learners to interact more smoothly with a character by acknowledging and leveraging learners' expectations. In this way, the learner can focus on the learning activities and materials in question, rather than being 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 agents" and the question of the "… role that stereotypic information [in the sense of visual stereotypes] plays in facilitating or inhibiting learning from animated agents." (ibid., p. 4). Nevertheless, there is a striking lack of problematizing the issues and the results of their study, 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 this outcome goes that "… the female tutor broke with rules of etiquette about who should teach at a college level by not conforming to the stereotype of males as professors." (ibid., p. 4), and is then left without further comment. That is, there is no mentioning of a conflict between, on 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 reinforce the notion of the male professor as the norm by using this stereotype. Likewise, the authors, without further comments, pose the question: "… do people learn more effectively about car repair from 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?" (ibid., p. 4). Compare Figure 8

showing an extract from the *Joe Doe* instruction series for US Army, drawn during the Second World War by Will Eisner.

Figure 8. *Joe Dope* (by Will Eisner)

In contrast to the Moreno et al. (2002) study, the complexity and hidden dilemmas in the use of visual stereotypes come forth clearly in the work of Baylor and her group at *RITL*. As related above, Baylor and collaborators have demonstrated (Baylor, 2005) that the use of virtual pedagogical coaches portrayed as young and attractive females can increase the willingness of female students to apply for technical education and to help increase their self-efficacy thanks to pedagogical processes such as role modelling and identification (cf. Bandura, 1977). The students find it easier to match these coaches compared to virtual coaches that are "typical, male, engineers", with their own personal identity. However a detailed analysis of the results indicate that the increase in self-efficacy at least partly stems from a conception of *such an engineer* – female, feminine, young and attractive – as less competent than a "real", prototypical, male engineer. What, according to the author, seems to occur is that the prejudice of females, and most of all feminine 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!". Now, this implies a potential conflict between a *short-term* pedagogical goal of recruitment and boosted self-efficacy in female students, and a *long-term* pedagogical goal of changing rather than reproducing gender prejudices and stereotypes. As Baylor (2005) remarks, prejudices about less competent female – especially feminine female – engineers are not ones that one would like to reinforce and disseminate.

Summing up, 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 long term pedagogical perspective. Likewise, there is a need to handle dilemmas where the breaking of a visual stereotype may have positive effects on attitudes as well as learning, but at the same time produce or reinforce questionable conceptions, such as the "female engineers are less capable" conception reported above.

6.3. Introducing novel risks

As to the detrimental *normative* function of visual stereotypes discussed in a previous section, the extended degrees of design freedom offered in the virtual world brings about additional risks. The construction and promotion of *idealized super people* with "perfect" bodies and looks (and even lives) has long since been abounding in non-interactive media such as television, video and magazines. Figure 9 (right) shows the femme fatal *P'Gell* in Will Eisner's comic *Spirit*. 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, such as the big-breasted, wasp-waisted action heroine *Lara Croft* (Figure 9, left): a biological contradiction whose tiny abdomen could hardly house all her vital organs, particularly if she also is to perform spectacular stunts and engage in violent fighting rather than fainting like corseted females of the 19th century.

Figure 9. Left: *Lara Croft* and right: *P'Gell* (by Will Eisner)

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, thus establishing a dynamic, mutual social relation. A possible effect of this is that the distance between users or learners and these "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 9, right). If we are also to actively 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".

Another related risk is the fact that virtual worlds involve users in a more active way than traditional movies and printed material. You can participate in various activities, including simulated everyday activities (c.f. *Second Life* at http://secondlife.com or *Entropia Universe* at http://www.entropiauniverse.com), and 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 much of reality behind. The addiction risk is apparent. Extensive and absorbing interaction with stereotypical characters may have negative consequences for peoples' conceptions of real people and real social life.

7. Conclusion

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

But there are also *novel* possibilities and risks introduced by visual stereotypes in virtual agents. We have emphasized the extended possibilities to challenge visual stereotypes for educational purposes. Specifically we have highlighted the possibilities to provide social identification and role modelling for larger groups of learners. Rightly used, we think, virtual pedagogical agents can be a tool for supporting the exploration and formation of identity in (young) learners while problemizing the reproduction of "undesired" (visual) stereotypes. Furthermore, it can be a pedagogical tool for initiating discussion and reflection on the role and the effects of visual stereotypes.

As to risks introduced by visual stereotypes in virtual agents, we have pointed at the replenished risks that follow from users *interacting* with, and perhaps being absorbed with, (visually) idealized stereotypes. The greater the knowledge about these issues among designers of digital learning material, the better the chance to counter these risks.

Summing up, we hold that the degrees of freedom as to visual design in digital virtual media, compared both to real life and to traditional media, imply an increased need for knowledge in order to navigate the design space in a thoughtful way. In products that are not directly commercial one may certainly wish that the visual designs of virtual pedagogical agents be based on informed design decisions.

7.1. Design guidelines

This is a natural point to ask for *design guidelines* as a support for designers to produce adequate visual appearances for virtual pedagogical characters, and we indeed think that research results within the domain should be used to contribute to a "visual design guidelines project". Such guidelines cannot, however, be step-by-step recipes on "how to visually design a virtual pedagogical agent". The reason for this is that whether a visual design decision is adequate and appropriate will always depend also on the learning context, the learning goals and the group of learners in question. Yet guidelines in the form of pointers or topics and considerations to reflect upon in combination with good and bad examples are certainly both possible and desirable. Pointers suggesting that "this is a question that needs to be answered before doing a choice on this or that visual parameter" or that "these variables relate to each other" can support a designer's navigation through the visual design space of virtual pedagogical agents.

Regarding visual stereotypes and expectations, there is also an overall design consideration and trade-off discussed earlier in the text that should be kept in mind. On the one hand, there are times when pedagogical benefits are gained by *challenging* pre-conceptions and pre-knowledge. On the other hand, there are times when there are pedagogical benefits in *exploiting* existing conceptions, expectations and preferences in students.

7.2. In sum

We began by analysing the use of visual stereotypes in traditional, non-interactive, media, and indeed, there is interesting research in the case of traditional media relating to this topic. However, we hold that the issues must be separately approached for interactive pedagogical media, since there are additional affordances, in positive and negative senses, in the case of interactive media. By this standpoint we disagree to some extent with Reeves & Nass (1996) who hold that traditional media, such as television, and new media, such as computers "… afford the same problems and opportunities of stereotyping." (ibid., p. 170). We believe that certain novel problems as well as opportunities enter the scene

with digital interactive media. And if academic research in the domain of virtual pedagogical agents keeps up with the technological and commercial development, there is a potential to take on some responsibility and to be proactive in channelling the development.

Acknowledgements

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

Copyrights

References

Bandura, A. (1977). *Social Learning Theory*, Englewood Cliffs, NJ: Prentice-Hall.
Baylor, A. (2005). The impact of pedagogical agent image on affective outcomes. *Paper presented at the 2005 International Conference on Intelligent User Interfaces (IUI 2005) Workshop on Affective Interactions: Computers in the Affective Loop*, January 9-12, 2005, San Diego, CA. Retrieved June 11, 2007 from http://garnet.acns.fsu.edu/~abaylor/PDF/image_iui.pdf.

Baylor, A. & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education, 15* (5), 95–115.

Baylor, A. & Plant, A. (2005). Pedagogical agents as social models for engineering: The influence of appearance on female choice. In Looi, C-K., McCalla, G., Bredeweg, B. & Breuker J. (Eds.), *Artificial intelligence in education: Supporting Learning through intelligent and socially informed technology, 125*, Amsterdam, The Netherlands: IOS Press, 65–72.

Baylor, A. & Rosenberg-Kima, R. (2006). Interface agents to alleviate online frustration. *Paper presented at the 7th International Conference of the Learning Sciences*, June 27 - July 1, 2006, Bloomington, IN.
Retrieved March 5, 2008 from http://garnet.acns.fsu.edu/~abaylor/PDF/ICLS_frustration.pdf.

Berscheid, E. & Walster, E. (1974). Physical attractiveness. In Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, *7*, New York, NY: Academic Press, 158–216.

Bickmore, T. & Cassell, J. (2005). Social dialogue with embodied conversational agents. In van Kuppevelt, J., Dybkjaer, L. & Bernsen, N. (Eds.), *Advances in Natural, Multimodal Dialogue Systems*, New York, NY: Kluwer Academic / The Netherlands: Springer, 23-54.

Brahnam, S. (2001). Creating physical personalities for agents with faces: Modelling trait impressions of the face. *Paper presented at the User Modelling 2001 (UM'01) Workshop on Attitudes, Personality and Emotions in User-Adapted Interactions*, July 13-17, 2001, Sonthofen, Germany.
Retrieved March 23, 2006 from http://aos2.uniba.it:8080/sonthofen/brahnam-f.pdf.

Brewer, M. (1988). A dual process model of impression formation. In Wyer, R. & Srull, T. (Eds.), *Advances in Social Cognition*, *1*, Hillsdale, NJ: Erlbaum, 1–36.

Cook, M. (1979). *Perceiving others – The psychology of Interpersonal Perception*, London, England: Methuen.

de Rosis, F., Pelachaud, C. & Poggi, I. (2004). Transcultural believability in embodied agents: A matter of consistent adaptation. In Payr, S. & Trappl, R. (Eds.), *Agent Culture: Human-Agent Interaction in a Multicultural World*, Mahwah, NJ: Laurence Erlbaum Associates.

GameGirlAdvance (2004). Comments: It's time for androgyny, it's just Vaan! *Comment on weblog post at GameGirlAdvance by pseudonym kyoujin*, May 20, 2004.
Retrieved March 5, 2008 from http://www.gamegirladvance.com/cgi-bin/mt-comments.cgi?entry_id=532.

Graesser, A., Lu, S., Jackson, J., Mitchell, H., Vetnura, M. & Oleny, A. (2004). Autotutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers, 36* (2), 180–192.

Gulz, A. (2005). Social enrichment by virtual characters – differential benefits. *Journal of Computer Assisted Learning, 21* (6), 405-418.

Gulz, A. & Haake, M. (2006a). Design of animated pedagogical agents – a look at their look. *International Journal of Human-Computer Studies, 64* (4), 322–339.

Gulz, A. & Haake, M. (2006b). Visual design of virtual pedagogical agents: Naturalism versus stylization in static appearance. *Paper presented at the 3rd International Design and Engagability Conference @ NordiCHI 2006 (iDec3)*, October 15, 2006, Oslo, Norway.

Gulz, A., Ahlner, F. & Haake, M. (2007a). Visual femininity and masculinity in synthetic characters & patterns of affect. In Paiva, A., Prada. R. & Picard, R. (Eds.), *Affective Computing and Intelligent Interaction*, Berlin/Heidelberg, Germany: Springer, 654–665.

Gulz, A., Haake, M. & Tärning, B. (2007b). Visual gender and its motivational and cognitive effects – a user study. *Lund University Cognitive Studies, 137.*

Hall, L., Woods, S., Dautenhahn, K., Sobral, D., Paiva, A., Wolke, D. & Newall, L. (2004). Designing emphatic agents: Adults versus kids. In Lester, J., Vicari, R. M. & Paraguacu, F. (Eds.), *Intelligent Tutoring Systems,* Berlin/Heidelberg, Germany: Springer, 604–613.

Johnson, W. (2003). Interaction tactics for socially intelligent pedagogical agents. *Proceedings of the 8th International Conference on Intelligent User Interfaces (IUI 2003)*, New York, NY: ACM Press, 251–253.

Kalick, S. (1988). Physical attractiveness as a status cue. *Journal of Experimental Social Psychology, 24*, 469–489.

Laurel, B. (1993). *Computers as Theatre*, Reading, MA: Addison-Wesley.

Massaro, D. (2004). A framework for evaluating multimodal integration by humans and a role for embodied conversational agents. *Proceedings of the 6th International Conference on Multimodal Interfaces (ICMI 2004)*, New York, NY: ACM Press, 24–31.

McArthur, L. (1982). Judging a book by its cover: A cognitive analysis of the relationship between physical appearance and stereotyping. In Hastorf, A. & Isen, A. (Eds.) *Cognitive Social Psychology*, New York, NY: Elsevier, 149–209.

Moreno, R. & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology, 31*, 186–207.

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

Moreno, K., Person, N., Adcock, A., Van Eck, R., Jackson, G. & Marineau, J. (2002). Etiquette and efficacy in animated pedagogical agents: The role of stereotypes. In *Working Notes of the 2002 AAAI Fall Symposium on Etiquette for Human Computer Work (Technical Report FS-02-02)*, Menlo Park, CA: AAAI Press, retrieved June 11, 2007 from http://web.media.mit.edu/~bkort/AAAI/Moreno-paper.pdf.

Nass, C., Isbister, K. & Lee, E-J. (2000). Designing and evaluating conversational interfaces with animated characters. In Cassell, J., Sullivan, J., Prevost, S. & Churchill, E. (Eds.), *Embodied Conversational Agents,* Cambridge, MA: MIT Press, 374–402.

Next Generation (1998). Girl trouble. *Next Generation Magazine*, January 1998, 37. Retrieved June 11, 2007 from http://www.electric-escape.net/node/542?PHPSESSID=50a6bdb50478cae8b6d3424d2c79dd57.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobreperez, P., Woods, S., Zoll, C. & Hall, L. (2004). Caring for agents and agents that care: Building emphatic relations with synthetic agents. *Proceedings of the 3rd International Joint Conference on Autonomous Agents & Multi Agent Systems (AAMAS 2004)*, New York, NY: ACM Press, 194–201.

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

Schleiner, A-M. (2000). *Mutation.fem*. Text for the "Alien Intelligence" Exhibit, Kiasma, Museum of Contemporary Art, Helsinki, Finland, retrieved June 11, 2007 from http://www.opensorcery.net/mutation/.

Schneider, D. (2003). *The Psychology of Stereotyping*, New York, NY: Guilford Press.

Smith, E. & Medin, D. (1981). *Categories and Concepts*, Cambridge, MA: Harvard University Press.

Voelker, D. (1994). *The Effects of Image Size and Voice Volume on the Evaluation of Represented Faces*. Doctoral dissertation, Stanford University, Stanford, CA.

Paper III

A Look at the Roles of Look & Roles in Embodied Pedagogical Agents – a User Preference Perspective

Haake, M. & Gulz, A. (2009)
International Journal of Artificial Intelligence in Education
vol. 19, no. 1, pp. 39-71

www.ijaied.org

A Look at the Roles of Look & Roles in Embodied Pedagogical Agents – A User Preference Perspective

Magnus Haake[a], Agneta Gulz[b]

[a] *Department of Design Sciences, Lund University, P.O.Box 118, 221 00 Lund, Sweden*

[b] *Division of Cognitive Science, Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden*

Original paper published in:
Int. J. of Artificial Intelligence in Education, vol. 19, no. 1, (2009), pp. 39-71.

Abstract: The paper presents a theoretical framework addressing three aspects of embodied pedagogical agents: visual static appearance, pedagogical role, and communicative style. The framework is then applied to a user study where 90 school children (age 12-15) in a dummy multimedia program were presented with either an instructor or a learning companion condition. They were then to choose between eight visually different embodied pedagogical agents: four more naturalistic (detailed & 3D-rendered) and four more stylized (simplified & cartoonish). Finally the participants were to choose between a strictly task-oriented versus a task- and relation-oriented pedagogical agent.

The goal of the study was to explore possible relations between the three aspects mentioned above with respect to user preferences. Results were: (i) When the agent was introduced as a learning companion, female students displayed a significant tendency to choose a more stylized visual character, (ii) When the agent was introduced as a learning companion, female students displayed a significant tendency to choose a task- and relation-oriented agent, (iii) In the case when students had chosen a more stylized character, there was a significant tendency to choose a task- and relational-oriented agent.

The paper also discusses limitations and strengths of the study, and advocates careful descriptions in studies of this kind – especially regarding the notion of visual realism.

Keywords. Embodied pedagogical agent, visual style, communicative style, pedagogical role, realism, naturalism, stylization.

1. Introduction

Embodied pedagogical agents – visually represented, computer generated characters in pedagogical roles, such as virtual instructors, mentors and learning companions – populate the digital society in increasing numbers. They are found in educational programmes from preschool to university, as well as in broader educational contexts in the form of virtual health coaches, information guides, etc.

Behind these embodied pedagogical agents, we find more or less sophisticated computational algorithms that handle communication, behaviour, animation, and so on.

Other aspects of *embodiment*, such as the physical appearance of agents, often fall outside this computational practice. Yet these are the aspects of design that form our first impressions of an agent and they do influence the way humans interact with them (e.g. Baylor & Kim, 2005). Therefore, it is necessary to address also these aspects of embodied pedagogical agents.

One of the more elaborated discussions on the topic can be found in an article by Ruttkay, Dormann, and Noot (2004), where a large number of design features are listed. One of the four main categories in the list is *embodiment* which designates 'all low-level aspects which contribute to the physical appearance of the character' (ibid., p. 32). Embodiment is further divided into two sub-categories: *communication modalities* and *look*. Communication modalities can be regarded as the *dynamic* characteristics of face and body that relate to expressivity via facial displays, body and hand gestures, postures, motion generation, etc. Look refers to underlying *static* visual characteristics of the face and body – shapes, colours, attributes, hair, clothes, graphical style, etc. – that convey gender, age, personality, and so forth.

Compared to other properties of embodied agents, such as movements, gestures, facial expressions and speech, their look or, in our terminology, their *visual static appearance* has (as we have already touched upon above) received relatively sparse attention both in embodied agent research in general and in the domain of embodied pedagogical agents, henceforth abbreviated as *EPAs*, (Gulz & Haake, 2006). There are exceptions, of course. Baylor and collaborators are among the (few) research groups that for a long time have approached several aspects of visual appearance, such as visual gender, attractiveness and 'coolness' as well

as visual realism (Baylor 2005a; Baylor 2005b; Baylor & Kim 2004; Baylor & Plant, 2005; Baylor, Rosenberg-Kima, & Plant, 2006).

Visualizations of gender, attractiveness as well as visual realism have also been examined by Nowak and collaborators (Nowak & Biocca, 2003; Nowak & Rauh, 2005; Nowak & Rauh, 2008). Yet another influential researcher is Bailenson, who together with collaborators has investigated user effects of visual realism as well as of height and attractiveness in avatars (Bailenson et al., 2003; Bailenson et al., 2005; Yee & Bailenson, 2007).

In relative terms, static visual appearance is nevertheless an unexplored domain. Furthermore, *when* visual appearance is researched, it is rather often approached in a simplified manner losing sight of the complexity of the notion. For example 'visually realistic' can be used in a sense that mixes up distinctions such as: fantasy or non-fantasy, human-like or not-human-like, 3D or not 3D, detailed or non-detailed, cartoonish or non-cartoonish. This tendency to speak of 'visual realism' in a simplified manner easily leads to unwarranted over-generalizations about 'visually realistic' agents.

From a pedagogical perspective, the relative lack of detailed research on visual appearance aspects of embodied pedagogical agents is unfortunate since there is reason to believe that the visual appearance of an agent may considerably affect learner expectations, attitudes, understanding and motivation. Such influences are well known and documented in related domains such as theatre, animated film, advertisements, as well as in research on interaction between human beings (Gard, 2000; Haake, 2006; Kalick, 1988; Lassetter, 1987; Schneider, Hastorf, & Ellsworth, 1979).

This is not to say that visual appearance is *more* important than other aspects of embodied agents. For one thing, without the development and refinement of computer hardware, algorithms, and computational models of behaviour, dialogue and intelligence, there would be no embodied agents whatsoever. Nevertheless, even with an ingeniously designed EPA from a computational perspective, an inadequate visual appearance can *decrease* the pedagogical benefits considerably. For instance, Baylor and collaborators have shown that a carefully chosen visual appearance of an EPA can influence both students' transfer of learning (Baylor & Kim, 2005) and their belief in their own competence in approaching a certain subject matter (Baylor, 2005b).

If one is interested in how learners are affected by different visual appearances of EPAs, this should not be approached as a separate issue in a 'vacuum' but in relation to contextual variables and/or other relevant characteristics of EPAs (as is also generally done in the studies referred to above).

In the present article we will approach aspects of *visual static appearance* of EPAs in relation to aspects of their *communicative style* and *pedagogical role*. First these three aspects will be presented and discussed. Then we discuss in greater detail the sub-aspects of EPAs that were manipulated in the user study presented in the second half of the article: (i) different aspects of graphic expressions (visual static appearance) (ii) strict task-orientation vs. task- and relation-orientation (communicative style), and (iii) authoritative instructor vs. non-authoritative learning companion (pedagogical role). Overall, the topic of visual static appearance is given prominence throughout the article.

Beside from generating results, the study exemplifies how the theoretical framework of the article can be applied. In particular, we go into more detail than is common regarding visual aspects of EPAs, and scrutinize the design choices we have made. Furthermore, in discussing the results of the study that indicate that the three sub-aspects or variables involved are related from the perspective of user/learner responses, we emphasize that this result strictly only applies to these specific design aspects and variables. On the other hand, the projection into the future is that by putting together pieces of evidence from different studies, *that are detailed and specific in what is actually being varied and evaluated*, we will be able to come up with tentative guidelines and generalizations regarding how different aspects of EPA relate to each other from the perspective of learners.

2. Visual Static Appearance

2.1. Design considerations for visual static appearance

From a positivistic approach visual static appearance may be regarded as a multidimensional design space that, although immensely complex, could be analytically approached. From the perspective of design practices, visual static appearance is better described as a holistic set of more or less elusive and continuously changing, context-dependent, qualities. But this perspective can hardly be subdued to any analytical deconstruction and, as a consequence, any kind of academic guidance can only hope for high-level *design topics* or *design considerations* to be addressed during the design process.

Following this line of reasoning, we propose an outline with three basic high-level design considerations concerning visual static design in EPAs: basic model, physical properties and graphical style. These design considerations are in line with and elaborated from the discussions around a more integrated design approach for virtual characters reported in Gratch et al. (2004) and further developed in Ruttkay, Dormann and Noot (2004).

2.1.1. Basic model

A first topic to consider concerning the visual static appearance of an EPA is a *basic model* or constitution for the character. The reason to employ an EPA is, in most cases, to add or strengthen social and communicative features in a pedagogical system (Johnson, 2003; Johnson et al., 2005; McQuiggan, Mott, & Lester, 2008; Moreno et al., 2001).

In practice this means that an EPA, as a minimum requirement, will need eye(s) and a mouth.[1] The eye(s) and the mouth will, in the general case, be contextualized within a face, often together with a body or part of a body. To model these features – a face with eye(s) and mouth and a potential body – we can draw upon the basic constitution of: a human, an animal or creature, an inanimate, non-living object, a fantasy concept or a combination of these entities. In Figure 1 the basic model is exemplified by the *MS Agent Package* (included in Microsoft Office until 2003) and the 'late' *MS Office Assistant*.

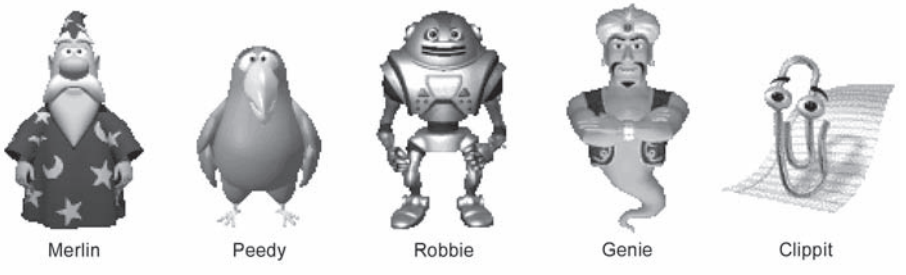


Figure 1. The characters from the *MS Agent Package* and the *MS Office Assistant* are well suited to exemplify the proposed basic model: *Merlin* (a human), *Peedy* (animal/creature), *Robbie* (a combination of a general human constitution, inanimate parts, and the fantasy concept of the robot), *Genie* (a combination of a human and a fantasy concept), *Clippit* (an inanimate object).

Two things should be emphasized here: First, the possibility to model an embodied agent upon different creatures and inanimate objects relies heavily on the power of anthropomorphism. Second, the qualities of the basic model are only discussed in relation to static visual qualities, and it should be observed that dynamic visual qualities mediated by animation in general reinforce the experience of anthropomorphism by mimicking schemas for human behavioural and communicative strategies.

[1] Hypothetically, one could for example imagine a virtual pedagogical agent in the form of a plain cube – but such a case is rather strained and better dismissed as an odd exemption.

2.1.2. Physical properties

A second design consideration regards what we chose to denote as *physical properties* of an EPA such as: body type, face shape, skin colour, hair cut and hair colour, clothes and various accessories (Figure 2). Visual representations of gender, age, ethnicity, profession and so on, can emerge through (combinations of) such properties.

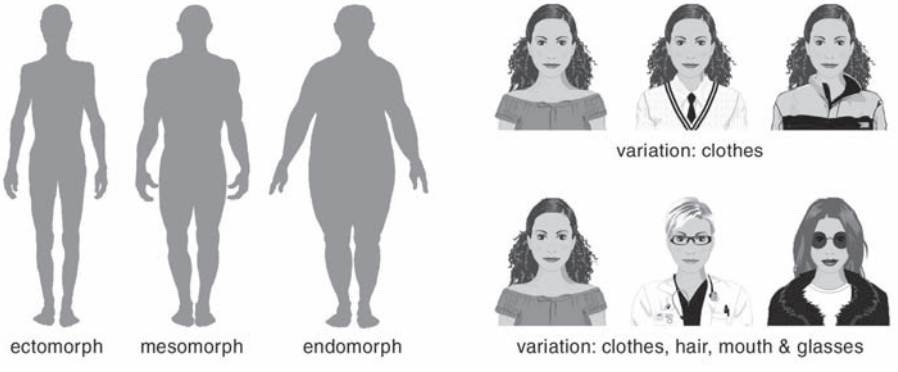


Figure 2. Examples of *physical properties*. Left: the three body types of ectomorph, mesomorpf, and endomorph (Sheldon, Stevens, & Tucker, 1940). Upper right: three different outfits transforming the experience of the one and same character as to age, personality, social position, education, etc. Lower right: variations of the same character by means of outfit, hair cut, hair colour, and use of lipstick and glasses, dramatically changing the experience of the character and characteristics attributed. (The six characters to the right were put together by means of the *SitePal Demo Tool*, www.sitepal.com)

The knowledge about different effects of these physical properties is not entirely impenetrable, and there is much empirical data in academic disciplines such as social psychology and behavioural science. In addition there is a large body of experience based knowledge in areas such as film, theatre, graphic design, and advertising (Gulz & Haake, 2006).

It is important to be aware that *whatever choices* one makes regarding the physical properties of an EPA, these properties will carry social, cultural, psychological and affective baggage. In other words, there is no such thing as a visually *neutral* character (Gulz, Ahlner, & Haake, 2007; Haake & Gulz, accepted; Isbister, 2006) – and nevertheless, to quote Moreno and Flowerday (2006), 'the vast majority of instructional interfaces assign arbitrary animated pedagogical agents assuming that the choice of an agent representation is psychologically neutral' (ibid., p. 191). This unawareness may result in non-optimal and restrained pedagogical adequacy and efficiency of EPAs as well as hamper scientific studies. As

a general recommendation, one ought to be observant whenever a certain EPA is claimed to have been evaluated against a presumable neutral EPA (taken to be some kind of baseline). Consequently, it is important to pay close attention to all the physical properties of EPAs involved in such comparisons, since all aspects that are *not focused on* need to be *comparable* between the evaluated characters. As an example, it would be a flaw to set up a comparison with a focus on gender that involves a male versus a female character, where one of them is substantially fatter than the other, or where one but not the other wears high-fashion clothing.[2]

2.1.3. Graphical style

The two first design considerations described above can to some extent be regarded as *conceptual design topics* in the sense that they may be verbally and analytically identified, described, discussed and decided upon in the design process.

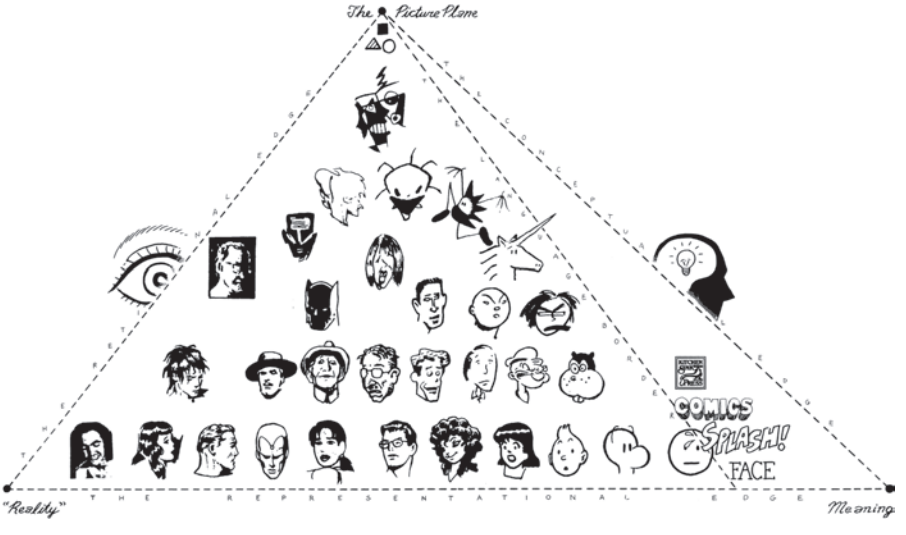


Figure 3. An edited reproduction of the design space of visual iconography (*The Big Triangle*) described by Scott McCloud in his book *Understanding Comics: The Invisible Art* (McCloud, 1993). While this is an ingenious way to set up the design space of graphical style, there is no predictable power to correlate a certain graphical style to a well defined response.

[2] On the other hand, one may have the goal of comparing two different complete *personas* that do differ in a large number of aspects. This is of course legitimate and can be highly valuable, yet in this case the comparisons in questions are carried out on a more holistic level and should be described, evaluated and discussed as such.

This is not as easy with design considerations concerning graphical style (Figure 3). Here we enter a highly complex design space where, for example, small changes in the qualities of the line (Figure 4), shadings, and proportions – relating to artistic and aesthetical properties of the medium as well as the skill of the graphical designer – may completely change the visual experience in diverse and unpredictable directions (Gulz & Haake, 2006; McCloud, 1993).

Figure 4. The inked line has almost unlimited possibilities with a potential of tremendous visual power (McCloud, 1993, p. 126).

For the purpose of embodied agents, the vast and complex design space of graphical style may seem impossible to navigate (c.f. Figure 3). In order to nevertheless relate to this space, one should bear two things in mind: (i) it takes a skilled graphical designer to manoeuvre in (at least) parts of this immense design space, and (ii) the graphical style may have significant but unpredictable effects on the perceived visual experience. In other words: one should acknowledge and pay attention to the qualities of graphical style.

With this said, some aspects of graphical style may be both possible and worth the effort to explore. With regard to EPAs, *degree of detailedness* and *degree of naturalism* may be possible and meaningful to analyze and evaluate – and since we address these two aspects in our study, we will present them both in some detail.

2.1.4. Graphical style: degree of detailedness

Beginning with a standard greyscale photography of a face this can be referred to as highly detailed. By reduction, the photo may be converted into a contour line representation or a two level posterized representation of the very same face (Figure 5, left). These two examples illustrate a straightforward algorithm based reduction of the information contained in the original greyscale photo where the different representations can be described as more and less detailed. However, in

design practice, a 'reduction of details' normally goes hand in hand with complex changes and deviations as to the graphical expression or style (Figure 5, right).

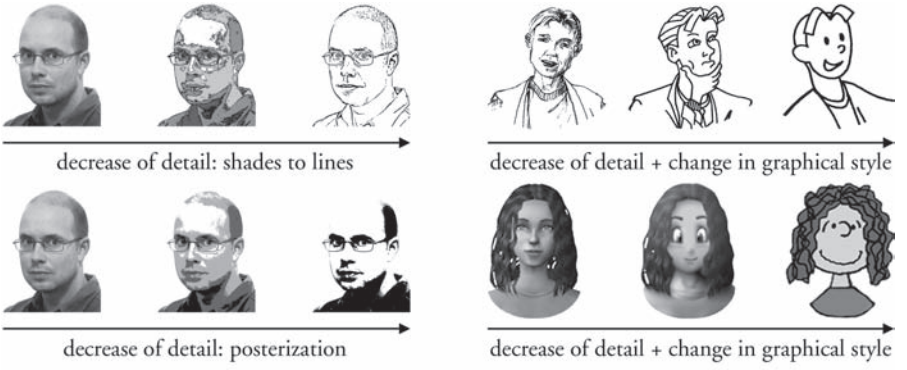


Figure 5. Graphical style: degree of detailedness. Left: two examples of straightforward reductions of details. Right: two more complex examples where reduction of details are combined with divergent, non-linear changes as to graphical style.

From a pedagogical point of view, the degree of detailedness is of interest since the amount of details may correspond to differences in cognitive processing. For example, the reduction of details promotes increased distinctness of facial expressions which may support a more rapid and accurate processing and interpretation (Cook, 1979; Isbister, 2006). The simplification may furthermore facilitate subjective self-identification which is a fundamental strategy to engage the reader of comics (Gulz & Haake, 2006; McCloud, 1993).[3]

2.1.5. Graphical style: naturalism – stylization

According to our proposed design space of graphical style, the same conceptual character with respect to a 'basic model' and 'physical properties' may be visually represented in numerous variations. By manipulating the qualities of line, shape and colour, each single variation (i.e. each graphical style) conveys its own non-linear, complex, dynamic, cultural and context dependent impact on the individual interpretation processes of the receiver.

[3] Herge's Tintin (http://en.wikipedia.org/wiki/The_Adventures_of_Tintin) is maybe one of the most splendid examples of this self-identification phenomenon. The Tintin character is identified by the tuft and the clothes, whereas the face is extraordinary simple (and the personality diffuse and elusive) – and actually 'Tintin' associates with 'nothing' in French. Correspondingly, the rest of the characters are more visually detailed and/or have more prominent personalities. Furthermore, Herge uses the *masking effect* contrasting the more simplified and stylized characters against a rather naturalistic background – making further use of the 'subject vs. object' phenomenon (Gulz & Haake, 2006; McCloud, 1993).

This may be a disappointing conclusion if the aim is to gain some kind of control or analytical power over aesthetic qualities. However, in the more specific domain of embodied pedagogical agents there is useful knowledge to gain by focussing on certain aspects of the design space of visual iconography (*The Big Triangle*) described by McCloud (1993) (see Figure 3). By collapsing the two dimensions of 'meaning' and 'picture plane' in the *pictorial plane* into a single (though complex) dimension of 'stylization', we can construct a dichotomy of *naturalism* versus *stylization*. Naturalism will here constitute a well-defined end point of an immense and diverging design space of different stylized graphical expression since a representation cannot be more naturalistic than naturalistic.

Figure 6. A design space of *naturalism* vs. *stylization*. The two figures to the left (the 'naturalistic' corner) differ in degree of detailedness (cf. Figure 5) but both are to be referred to as 'naturalistic'. This design space is derived out of the *pictorial plane* by Scott McCloud (1993) but is simplified in order to emphasize the dichotomy of *naturalism* vs. *stylization* (thus ignoring the complexity of different stylized expressions).

In Figure 6, this dichotomy is exemplified by the same conceptual character (a young female) visualized in two more or less 'naturalistic' versions and four different 'stylized' versions.

As to the two 'naturalistic' representations (based upon the style of *The Sims 2*) they differ in degree of detailedness, but are both to be regarded as more or less

naturalistic. The leftmost character corresponds to a standard photo while the other character is reduced as to detailedness (cf. Figure 5); a spotlight will produce the same effect as the rightmost of the two naturalistic representations.

The four 'stylized' representations vary in expressive style inspired by: *Peanuts* (simplified, whimsy and humorous), *Manga/Dragonball* (cute, emotional and friendly), *Hernandez* (underground, rebellish), *Picasso* (abstract and intellectual). Also observe that the design space of 'stylization' in Figure 6 does not convey any actual information about the relation between different kinds of stylization. Any attempt to fix coordinates is meaningless since the effects on the user of different graphical styles are everything but well-defined and linear.

It should furthermore be noted that a *3D-rendered* representation does not automatically equal visual naturalism, but is basically an aspect of graphical style that can be more or less detailed as well as more or less stylized.[4] To exemplify, a 19th century (2-dimensional) black-and-white photo is a far more naturalistic representation than the low polygon 3D-rendering of the embodied agent *REA* (Bickmore, 2003) or 3D-rendered cartoon characters like *Mr Incredible* in the animated film *The Incredibles* (Walt Disney Pictures & Pixar Animation Studios, 2004).

2.2. What about 'realism'?

At this stage we are ready to approach the concept of visual 'realism' with respect to visual static appearance. Considering the embodied agent literature, the term 'visual realism' or 'visually realistic' is used in different ways that can be mapped to one or some of the design considerations proposed in this article. In an attempt to grasp a common concept of a 'realistic embodied agent', we propose the following definition: (i) modelled upon some prototypical basic model of a human, (ii) adequate and relevant as to physical properties of the chosen basic model, (iii) detailed, and (iv) naturalistic.

It is problematic that many studies, in presenting their material and their conclusions, lack in specificity as to what is actually meant by the term realism/realistic. Sometimes comparisons regard a character modelled upon a human versus a character modelled upon a fantasy concept, but where both are actually perfectly naturalistic in relation to their basic models and thus equally realistic given a fictive context of e.g. a science fiction movie.

Sometimes the difference is in the degree of detailedness, but combined with a difference as to naturalism-stylization that is not mentioned or problematized,

[4] When it comes to spatial interaction, on the other hand, the 3D experience must be considered a fundamental aspect for the simulation of a 'virtual reality'.

and so on. Nevertheless, results are often generalized in terms of 'visual realism' versus 'visual non-realism'. An illustrative example is reported by Gustavsson and Czarniawska (2004). At a conference, there was a discussion on the development of the interactive assistant *Olga* (Figure 7). The linguists in the development team opted for a 'more realistic' Olga, since they insisted that in order for Olga's speech to be understood, Olga must be made as humanlike as possible. Olga's lip movements, in particular, had to correspond to those of a living human. The designers in the development team, on the other hand, were of the opinion that the comic styled Olga was actually more humanlike as well as attractive than the 'more realistic' 3D-Olga, who they thought looked like Frankenstein's monster (an opinion shared by most of the conference audience).

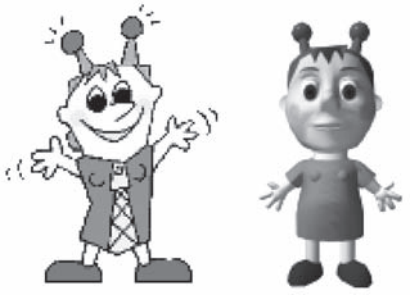


Figure 7. The interactive information assistant Olga

Our point with this example is that none of the Olgas (Figure 7) are 'realistic' but instead constitute two different visually 'stylized' representations where both could be referred to as: (i) modelled upon the same combination of a human and a fantasy concept of an alien from outer space, (ii) comparable as to physical properties (with the exception of the neck and the tie), (iii) rather comparable as to a low degree of detailedness, (iv) differences as to graphical style: 2D rendered, a bit squarish/angular shapes, some visual freedom as to postures *versus* 3D-rendered, a bit softer/rounder shapes, constrained (stiff) as to postures. This example then indicates that the notion of 'realistic' sometimes is more of a non-reflected idea than an actual and explicit visual quality ('*If it's 3D, it's realistic!*').

2.3. Realism and pedagogical effects

The number of empirical studies on user responses towards more or less visually realistic EPAs is actually rather limited (Gulz & Haake, 2006). Furthermore, as discussed above, the visual variables in these studies are often not sufficiently controlled and described or even considered. Regardless of this, there is a relatively extensive theory based discussion pro and con visual realism in virtual characters that exposes a number of arguments *for* as well as *against* realism in visual static appearance in EPAs. (For an overview of this, see Haake & Gulz, 2006.)

In view of the study to be presented in this paper, it is relevant to consider the arguments for and against visual realism regarding the potentials of virtual characters to establish some kind of (social) relations with learners.

If we start with the pro side, it has been proposed that a high degree of visual realism will increase user/learner *involvement* and sense of presence in a digital environment, and therefore even be a condition for human co-operation with embodied agents (Welch et al., 1996). Another argument states (Wilson, 1997) that a high degree of visual realism is important in order to elicit users' curiosity about the personality of a character, which in turn can be regarded as an important condition for the development of a relationship between a learner and a EPA. Furthermore it is sometimes argued that visual realism will facilitate learners' self-identification with an EPA (Baylor & Kim, 2005).

On the other side, we find the approach (mentioned earlier in the section on 'degree of detailedness') arguing that a 'non-realistic' representation may facilitate self-identification and involvement. In these arguments 'non-realistic' is probably best described as a character with a low degree of visual detailedness in combination with a simplified, stylized graphical representation related to the lower right base of *the design space of visual iconography* (Figure 3), e.g. the graphical style of Charles Schulz's *Peanuts*.

According to McCloud (1993), people will more easily get *involved* with and also be likely to *project themselves* into a visually simplified character than in a highly detailed and naturalistic character. The highly detailed and naturalistic character is more of a visual and socio-emotional fact, which does not leave much for a user to elaborate on and fill in, whereas a stylized character invites elaboration by the user, who may fill in from his or her own personal and subjective experiences (Haake & Gulz, 2006). In McCloud's wordings, the stylized character is 'an empty shell that we inhabit' (McCloud, 1993, p. 36).

We will return to these different framings and arguments in the discussion of our study.

3. Pedagogical Roles

In our everyday world, we can find a large variety of pedagogical roles. The following list, as well as the overall categories of more and less authoritative roles, is based on a paper by Chou and collaborators (Chou, Chan, & Lin, 2003), to

which we have added some pedagogical roles as well as some examples of EPA correspondences for certain roles.[5]

More authoritative roles:

- the tutor, e.g. the agent Autotutor (d'Mello et al., 2006)
- the coach, e.g. the agent Laura (Bickmore, 2003)
- the guide, e.g. the agent PPP Persona (André, Rist, & Müller, 1998)
- the instructor, e.g. the agents Steve (Shaw, Johnson, & Ganeshan, 1999) and Adele (Johnson, Rickel, & Lester, 2000)
- the mentor (e.g. Baylor & Kim, 2005)
- the expert (e.g. Baylor & Kim, 2005)

Less authoritative roles:

- the competing co-learner (e.g. Chan et al., 1992)
- the collaborating learning companion, e.g. the agent Sam (Ryokai, Vaucelle, & Cassell, 2003)
- the tutee, e.g. the teachable agents Bill and Betty (Biswas et al., 2005)
- the peer tutor that can alternate in the role of being a tutor and a tutee (e.g. Chan & Chou, 1997)
- the troublemaker (e.g. Aïmeur & Frasson, 1996)
- the critic, the criticizing co-learner (e.g. Hietala & Niemirepo, 1998)
- the clone (e.g. Chang et al., 1999)

We believe that this classification into more and less authoritative roles together with the listed examples gives a satisfying overview of the different pedagogical roles of interest for EPAs. Focusing on the individual specified pedagogical roles, it is not obvious that another kind of classification would add much of interest. Nonetheless, in parallel to our arguments on the visual static appearances of EPAs, we argue that pedagogical roles, when they are involved in user evaluations of EPAs, should also be spelled out in detail. The reason is simply that these role concepts listed above are not clearly defined. The boundaries between them are often diffuse and change between cultures and sub-cultures. A notion such as authoritativeness certainly varies between cultures, and also within cultures we find variations. For instance, many of us have experienced mentors, coaches and teachers that have varied more or less with regard to their authoritative style.

[5] Note that Chou et al.'s (2003) classification also applies to pedagogical agents that are not visually embodied, e.g. classical intelligent tutoring systems.

As an instructive example of a study that provides this specificity, we would like to hold forth the paper *Simulating Instructional Roles* by Baylor and Kim (2005). The study presented in this paper involved the three pedagogical roles of *Expert*, *Mentor* and *Motivator*.[6] For each of the roles the following was designed and specified: (i) a script where the relation between information and encouragement is varied between roles (primarily informative for the Expert, primarily encouraging for the Motivator, and both informative and encouraging for the Mentor), (ii) an emotional or affective style varying between the roles (no emotional expressions in the Expert versus a variety of emotional expressions for the two other roles), (iii) animation gesture schemes (deictic gestures in the Expert, emotional expressive and highly-animated gestures in the Motivator, and a mixture of these gestures in the Mentor), and (iv) visual static appearances (older, serious looking, formally dressed Expert, a young, smiling, causally dressed Motivator, and the Mentor being in between).

In the study, these particular role instantiations of EPAs were empirically validated in the sense that it was established that learners perceived the embodied agents as having these intended pedagogical roles. Furthermore, it was confirmed that the three agents had divergent effects on learning and motivation according to their roles. The Expert agent led to increased information acquisition, the Motivator led to increased self-efficacy, and the Mentor led to overall improved learning and motivation (*ibid.*). The possible effect of visual appearance on these results was not addressed in this particular study.

4. Communicative Style

The notion communicative style is used to refer to a number of different dimensions such as degree of argumentativeness (Myers, 1998); degree of assertiveness (Myers, 1998), descriptive vs. narrative (Kemper et al., 1990; Berman & Slobin, 1994), extrovert vs. introvert (Janvier & Ghaoui, 2003), separate vs. connected (Galotti et al., 2001).

It is probably evident to all who have been to school or education of some kind that human pedagogues differ on such aspects of communicative style. Given that most EPAs also communicate verbally, in text and/or speech, it can be relevant to consider dimensions of communicative style in the development and evaluation of EPAs. Here again, the most important for an evaluation is, in our

[6] Note that motivator is not included in the list above, but is close to a collaborating learning companion.

view, that the communicative style dimension used and evaluated is described in sufficient detail. The work of Bickmore and MIT Media Lab provides exemplary specificity on the aspects of communicative style implemented and studied in the embodied agents *REA* and *Laura*, where the former is a virtual real estate agent who interviews potential home buyers and shows them around houses and the latter a virtual coach or advisor for individuals who want to increase their physical activity.

The communicative style aspect used in our study is largely inspired from Bickmore's work. The definitions we provide below of *strict task-orientation* vs. *task- and relation-orientation* is largely based on Bickmore's categories: (i) *task condition* vs. *social condition* (Bickmore, 2003), and (ii) *relational* vs. *non-relational* (Bickmore, 2003).

Given a pedagogical context, our definition of *strict task-orientation* vs. *task- and relation-orientation* is as follows:

- A pedagogue (instructor, mentor, collaborative learning companion, etc.) who is *strictly task-oriented* sticks closely to the learning task, provides information in a succinct and objective way and focuses on matters of fact. He or she is strictly professional and does not bring in additional elaborative issues that are not clearly related to the knowledge domain and learning task at hand.
- A pedagogue (instructor, mentor, collaborative learning companion, etc.) who is *task- and relation-oriented* does, apart from contributing to the solving of the learning task, also focus on the developing of a social relationship with the learner; he/she is more subjective, personalizes the task and focuses less strictly on the task in the dialogue. The dialogue may also contain small-talk, conversational storytelling, getting-acquainted-talk, joke-telling, sharing of personal experiences, preferences and opinions, etc.

Returning to REA and Laura, both agents have been developed in two versions that can be mapped on the two communicative styles just described, even though the pedagogical context is more prominent in the case of Laura. The task condition REA can be mapped on a strictly task-oriented virtual estate agent, and the social condition REA on a task- and relation-oriented virtual estate agent. The non-relational condition Laura can be mapped on a strictly task-oriented virtual coach, and the relational condition Laura on a task- and relation-oriented virtual coach.

Additionally the two Lauras also differ in non-verbal aspects. In the relational version there is more of forward lean, body and facial orientation towards the partner, and more smiling, nodding, gazing, gesturing, etc. Certain voice features

are also more often present such as greater warmth and expressiveness, reinforcing interjections and more variation in pitch, amplitude, duration and tempo (cf. Bickmore, 2003). We have not included these non-verbal aspects in our definition above, primarily since the possibilities and resources for implementation in our study set-up were limited, and to only describe them would hardly be meaningful. In a fuller developed implementation of strict task-orientation vs. task- and relation-orientation, they should be included since there is empirical support for this from studies of communicative styles in human beings (Bickmore, 2003).

Also Baylor and Kim (2005) made use of a similar pair of communicative styles when implementing their *Expert* and *Mentor* agents mentioned previously. The *Expert* agent was task-oriented in the sense of (only) providing accurate information in a succinct way. He also spoke in a formal and professional manner with authoritative speech, showed not affect and was in the animation limited to deictic gestures with little expressiveness. The *Mentor* agent was task- and relation-oriented in the sense of working more closely together with the learner and with the goal of demonstrating competence to the learner while simultaneously developing a social relationship to motivate the learner. Gestures incorporated both deictic and emotional expressions, and the agent showed various affects such as confusion, approval, excitement and pleasure.

5. Variables of the Study

After this exposition of visual style, pedagogical role, and communicative style, we now provide a description of the variables used in our empirical study.

5.1. Visual appearance: detailed & 3D rendered vs. simplified & cartoonish

The initial approach to the design of the characters used in the study was to vary 'degree of visual realism' by using two sets of characters. Character set 1 made use of a 'detailed' and 'naturalistic' representational form (e.g. the style of *The Sims 2*), and character set 2 made use of a 'semi-detailed' and 'semi-naturalistic/semi-stylized' representational form (e.g. the style of *Marwin Comics*), (Figure 8).

Based upon our proposed approach to visual static appearance, we will below describe the two sets of characters used in the study in more detail. The notion 'neutral' should here be interpreted as: not visually deviating and/or conspicuous for the group of people in the study.

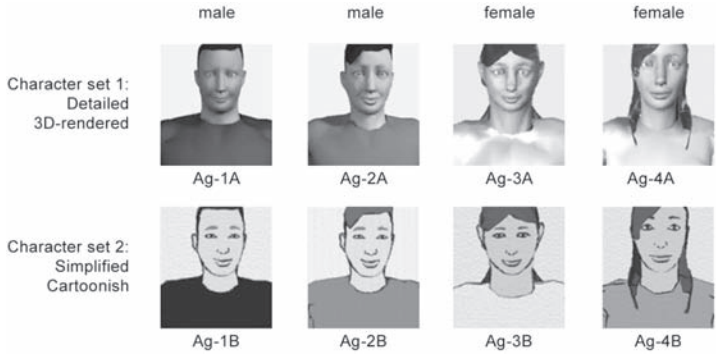


Figure 8. The two sets of characters (top row and bottom row respectively) used in the study.

Character Set 1

a) Basic model: A human.
b) Physical properties: Age 20-25, averaged/neutral body related aspects, neutral hairstyles, neutral clothes (t-shirts), Caucasian ethnicity, overall similarity except for gender.
c) Graphical style: Striving for 'realism', the characters can be described as detailed and naturalistic with a 3D-shading effect (cf. *The Sims 2*). Henceforth the graphical style of character set 1 will be referred to as 'naturalistic' and/or 'detailed & 3D-rendered'.

Character Set 2

Same as character set 1, except for:

c) Graphical style: These characters are simplified (semi-detailed) and less naturalistic (semi-naturalistic/semi-stylized) with a cartoonish style (cf. *Marwin Comics*). Henceforth the graphical style of character set 2 will be referred to as 'stylized' and/or 'simplified & cartoonish'.

5.1.1. Additional design comments

The facial forms of all characters were modelled to be as similar as possible given the condition that one should perceive them as different individuals. Age related features were held constant. Body shape and eye colour on the male agents were identical. The female agents had small differences in body shape, the same hair length, and the same eye colour. Clothing (t-shirts) was simple and discrete in all agents. The underlying design rationale was to make all characters relatively 'neutral' with respect to the actual group of participants, as well as comparable between the four characters. By this we mean that we strived to avoid extremes

and avoid any character(s) sticking out with respect to visual stereotypes (attractiveness stereotypes, personality stereotypes, gender stereotypes, etc.). Pre-tests of the characters confirmed/validated this goal. However, we are not suggesting that the visual characters are neutral in any absolute sense. They certainly carry social and cultural connotations as to gender, age, ethnicity, etc. (Gulz, Ahlner, & Haake, 2007).

Another rationale behind the visual static design decisions was to adhere to the association of 'computer characters' as being 'realistic' (naturalistic, relatively detailed and 3D-rendered) and the association of 'comic characters' as being simplified, cartoonish and 'flat'.[7]

Summing up, the visual variation between the two character sets can be described as differences in graphical style: detailed & 3D-rendered vs. simplified & cartoonish.

5.2. Pedagogical role: expert/instructor vs. learning companion

The two pedagogical roles varied in the study were an authoritative *instructor* versus a non-authoritative, collaborating *learning companion*.

More specifically the pedagogical roles were presented in the form of *a chief editor at a magazine* and *a companion journalist*, with the student being herself a journalist who is going to be sent to various European countries to do article research on different topics.

These roles were not implemented in any actual interaction between learners and agents but were provided within the multimedia program in the form of verbal scenarios. (These scenarios had been developed together with teachers who work with students of the actual age group, in order to ensure that the linguistic style would be adequate and easy to understand.)

In the *instructor version* of the program, the student was presented with the narrative that there was a chief editor in London who would be her or his instructor. The chief editor would formulate the missions, orient the journalist (the student) and provide necessary information at critical stages. The journalist (the student) should continuously report back to the chief editor who would evaluate the reports and tell what is well done and what should be improved and further worked on. Furthermore, it was implied that the chief in question had high standards and demanded much from the journalists.

In the *companion version* of the program the student was presented with the narrative that there would be a companion journalist with whom s/he would

[7] This association was also confirmed in interviews.

conduct the missions. The student was also told that it would be important to co-operate with the companion, who might not be completely reliable when it comes to knowledge but would provide some of the keys necessary to complete the missions.

5.3. Communicative style: strictly task-oriented vs. task- and relation-oriented

The two styles we chose to include in our study are based on the definitions given above in the section 'Communicative Style': (i) a *strictly task-oriented* communicative style versus (ii) a combined *task- and relation-oriented* communicative style.

As in the case of pedagogical role, these were not implemented in the form of an interactive session between user and agent, but were provided by an experimental leader in the form of two verbal scenarios. (Again, the scenarios had been developed together with teachers who work with students of the actual age group in order to ensure that the linguistic style would be adequate and easy to understand. The exact wordings of the two scenarios also varied somewhat depending upon the pedagogical version of instructor or companion.)

i) One scenario described an instructor/companion that focuses on the mission and sticks to this, deals with matters of fact in an efficient way, is strictly professional and does not invest time in the establishing of any more personal relation with the user (the student).
ii) The other scenario described an instructor/companion that is more social and apart from working on the professional tasks also supplies information about him or herself, tells about former missions, her/his family and friends, interests, etc. and seeks to share stories, opinions on various things and personal experiences with the user.

6. Method

6.1. Goal of the study

The goal of the study was to explore users' choices of EPAs as to their *static visual appearance* (detailed & 3D-rendered vs. simplified & cartoonish) and communicative style (task-oriented vs. task- and relation-oriented), in relation to their pedagogical role (instructor vs. learning companion). In particular, we wanted to explore possible relations between these variables as to user responses.

6.2. Experimental design

The participants were presented with one of two alternative versions of a pedagogical multimedia program with regard to the *pedagogical role* of the embodied agent: an *instructor* version or a *companion* version. The participants were then to make two choices.

i) First regarding the *static visual appearance* of the pedagogical agent: two male and two female characters, each presented in two *graphical styles*: detailed & 3D-rendered vs. simplified & cartoonish (c.f. Figure 5), altogether eight alternatives (Figure 8).
ii) Second regarding the *communicative style* of the agent (*task-oriented* vs. *task- and relation-oriented).*

This yields an experimental set-up, where *pedagogical role* is the independent variable, and *static visual appearance* and *communicative style* are the dependent variables (Figure 9).

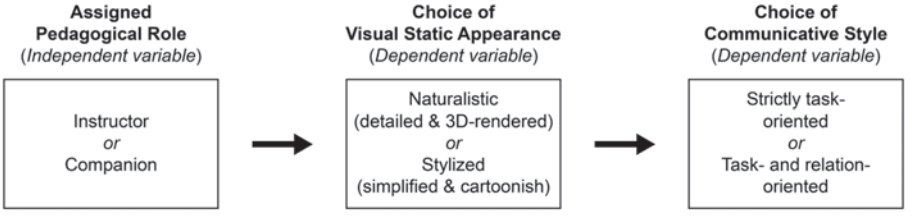


Figure 9. Experiment flow.

6.3. Hypotheses

Three hypotheses were formulated.

– *Hypothesis 1:* The *visual style* preference will depend on the assigned *pedagogical role* of the embodied agent in the following manner: the *instructor* version will yield a preference for a *detailed & 3D-rendered (naturalistic)* agent, and the *companion* version will yield a preference for a *simplified & cartoonish (stylized)* embodied agent.
– *Hypothesis 2:* The *communicative style* preference will depend on the assigned *pedagogical role* of the embodied agent in the following manner: the *instructor* version will yield a preference for a *strictly task-oriented* embodied agent, and the companion version will yield a preference for a *task- and relation-oriented* embodied agent.
– *Hypothesis 3:* The *communicative style* preference will depend on the *visual style* preference in the following manner: a preference for a *detailed & 3D-rendered (naturalistic)* embodied agent will yield a preference for a *task-ori-*

ented agent, and a preference for a *simplified & cartoonish (stylized)* embodied agent will yield a preference for a *task- and relation-oriented* embodied agent.

6.3.1. Comments on the hypotheses

Thus, the first and third hypotheses concern the choice of visual style in relation to the pedagogical role of the agent, and in relation to the choice of communicative style. These hypotheses are based on pilot studies (described in Gulz & Haake, 2005) as well as on theoretical considerations (Gulz & Haake, 2005; McCloud, 1993).

According to the theoretical framework proposed by McCloud (*ibid.*) cartooning and simplifying is framed as a way to *amplify the meaning* of an image. For instance, the meaning of facial expressions may be amplified, and thereby afford more powerful socio-emotional communication. This, in turn, is a fundamental element in companionships and friendships but not so in boss-subordinate relationships. Similarly socio-emotional communication is integral to relation-oriented but not to strictly task-oriented communication.

Another theoretical consideration presented by McCloud (*ibid.*) is the *universality* of stylized and simplified representations. Whereas a photorealistic picture represents only one individual, an abstract cartoon image might represent millions of people. An abstract cartoon image is therefore more likely to afford feelings of 'being related to' and of identification. These, again, are fundamental and inherent aspects in companionships and friendships, but not in boss-subordinate relations. Likewise they are aspects integral for relation-oriented but not for strictly task-oriented communication.

It should be observed, however, that there are *other theoretical considerations* that do not lead to the hypotheses formulated in this article, but rather to their reverse. *Media richness theory* instead predicts that emotionally engaged communication, in contrast to strictly task-oriented communication, is facilitated by *visual realism* in the representation of actors (Daft & Lengel, 1986; Dennis & Kinney, 1998; Kang, Watt, & Ala, 2008). It predicts, that it is more likely that a more visually naturalistic character compared to a more stylized one will evoke associations such as: 'this could be a friend of mine', 'this is someone I can imagine having a friendly chat with', and 'this is someone that I could feel close to', etc. (Kang, Watt, & Ala, *ibid.*). The naturalistic representation, being richer in visual cues and providing higher visual fidelity, increases social presence, which involves emotionally engaged communication and feelings of closeness and relatedness to one's communicative partner. In other words, more natural-

istic character representations facilitate aspects integral to companionships (in contrast to relations with a more authoritative and distanced person) as well as aspects integral to relation-oriented communication (in contrast to strictly task-oriented communication).

Since there are, thus, divergent predictions related to different theoretical considerations, hypothesis 1 and hypothesis 3 correspond to open research questions.

The second hypothesis regards pedagogical role and communicative style and is derived from 'real world' experiences and conceptions, where an instructor or boss is more readily associated with a task-oriented communicative style, and a companion is more readily associated with a task- and relation-oriented communicative style.

6.4. Participants

Ninety 12-15 year-olds adolescents (48 girls and 42 boys) from a Swedish secondary school participated in the study, which was organized in the context of their regular arts lessons. The students came from nine different teaching groups. Nearly all students in all groups participated. An overall observation was that the students seemed enthusiastic to participate and a couple of times the students spontaneously organized a queuing system among themselves. All students had at least some familiarity with pedagogical programs making use of embodied computer characters.

6.5. Materials

For the study, two dummy versions of a scenario based multimedia program for elementary school were developed (Figure 10a). In both versions the student is to take the role of a journalist at a magazine, being sent to various European countries to do article research. In the *instructor version* the student is presented with the scenario that she is to be guided by a virtual *chief editor*, and in the *companion version* accompanied by a virtual *journalist companion*. The multimedia presentation dummies were created in *Macromedia Director* and include: (i) an introduction where the program and a first mission is presented, and (ii) a module where the student is asked to choose one out of eight animated characters presented as their virtual instructor (in the instructor version) or as their virtual companion (in the companion version), see Figure 10.

The presentation of the mission includes illustrations from Istanbul and traditional Turkish music (Figure 10c). A male speaker voice tells about the mission and presents the student with her or his role as a journalist. As described in the

section 'Variables of the study' the student also gets information about the role of the chief editor (in the instructor version) or the companion journalist (in the companion version). All the material had been pre-validated by teachers who work with the age group in question.

Figure 10. The dummy multimedia program used in the study: a) Data collection of participants and setting for instructor/companion version, b) Start page, c) Introduction page, d) Preference choice page.

The eight embodied agents are the same in the instructor and companion versions: two males and two females with two different visual representations (detailed & 3D-rendered and simplified & cartoonish respectively). They are presented in an elliptical layout on the page (Figure 10d), where the individual positions of the agents are determined from a predefined randomized table. The layout was also pre-tested to minimize experimental artefacts of position effects.

6.5.1. The agents

The eight embodied agents were developed out from four basic figures, two males and two females. They were created as 3D-models in the program *3d Studio Max 5* and their faces created with the plug-in module *FacialStudio*. The four basic figures were then rendered into: (i) a detailed & 3D-rendered character set using the *3d Studio Max 5* default renderer, and (ii) a simplified & cartoonish character set using the *finalToon* rendering tool.

The agents were furthermore slightly animated. The animation was parsimonious and included only subtle eye blinks and breathing – there were no sudden movements that would attract attention. The movement patterns were similar between agents but with a displacement in time between the different agents in order to avoid movements from different agents from coinciding. Each animation lasted a few seconds and was repeated in a loop so that the agents seemed to move continuously.

Had the animations been more extensive and explicit, the animation patterns ought to have been randomized between trials and agents, but for the present study there was no indication that the animation patterns had any effect on the choice patterns.

6.6. Study procedure

During the experiment, three experimental leaders alternated between the following roles: *Experimenter A* initializing the study program and entering participant data into the computer (Figure 10a), *Experimenter B* conducting the interviews including taking interview notes, *Experimenter C* making back-up notes about the participant, double-checking the experiment, and assisting in various ways when needed. The interview scripts had been developed in collaboration with a teacher who works with the age group in question. The procedure of the experiment was as follows.

1) The three experimental leaders presented themselves to the class as coming from the university and doing research on educational media. Students were told they were welcome to participate in a study. The students went, one at a time, to a small room behind the classroom. In most cases all students in the class participated. It was emphasized that the main purpose was to listen to students' opinions on various aspects of a program that they would be asked to try out and that they were fully anonymous. (The experimental leaders did not mention who created the program.)
2) Before each new student, Experimenter A initialized the next experiment session by entering the session number and program version (instructor/companion) according to a random scheme. Each student was randomly assigned to one of two conditions: (i) the companion version, or (ii) the instructor version. There were 45 participants in each condition.
3) After welcoming a student, Experimenter A asked what grade she was in, whereupon she was asked to sit down at the computer. Meanwhile Experimenter A entered codes for grade and gender into the computer (Figure 10a) and initiated the session.

4) The student was told to press 'start' (Figure 10b) and then listen to further instructions provided by the program (Figure 10c).
5) After the program introduction, the student reached the module for the choice of visual representation (Figure 10d) and was instructed by the text on the screen to choose one of the eight embodied agents. Right after the choice, the chosen agent was enlarged and centred on the screen, with the other agents simultaneously disappearing.
6) Experimenter B then asked the student the open-ended question: 'Why did you choose the instructor/companion you did?'
7) Upon answering, Experimenter A backed the program to bring forth all eight embodied agents again. Experimenter B asked the student whether there were any that she would definitely not have chosen as her instructor/companion. Thereafter Experimenter B asked the student what she thought to be the important differences, if any, between the agents.
8) Experimenter B then verbally presented the student with the two scenarios on: (i) a strictly task-oriented agent (chief editor/journalist companion) and (ii) a task- and relation-oriented agent (chief editor/journalist companion). The student was then asked by Experimenter B which of the two agents she would prefer and asked to motivate her choice. The scenarios as well as the questions were based on a standard script. When required they were rephrased or clarified in order to secure that all students had an adequate understanding of the tasks.
9) Finally, two cognitive style inventories were completed by Experimenter B (Gulz & Haake, 2006). After the completion of this, the student was offered refreshment, was debriefed and thanked for valuable help.

For each participant the following was logged by the program: the program version (instructor/companion), the positions of the embodied 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 arguments for and articulations of their choices were noted down manually during the session and transferred to a computer transcript within a few hours.

7. Results

The test data can be presented as frequencies in a three dimensional contingency table categorized by the variables: pedagogical role (*P*), visual style (*V*), and communicative style (*C*) as shown in Table 1.

Table 1. The test data categorized by the three variables: Pedagogical Role (*P*), Visual Style (*V*), and Communicative Style (*C*).

		Communicative Style		
Pedagogical Role	**Visual Style**	*Task*	*Relation*	**Total**
Instructor	*Naturalistic*	12	6	18
	Stylized	9	20	29
	Total	21	26	47
Companion	*Naturalistic*	7	8	15
	Stylized	7	21	28
	Total	14	29	43
	Column Total	35	55	90

To control for higher order interactions in multi-dimensional contingency tables, interactions can be evaluated using a log-linear analysis. In our case, the experimental data were evaluated by log-linear analyses (using *R for Windows*) in the saturated model:

$$\ln(F_{ijk}) = \lambda + \lambda^{P} + \lambda^{V} + \lambda^{C} + \lambda^{PV} + \lambda^{PC} + \lambda^{VC} + \lambda^{PVC}$$

The log-linear analyses used were (c.f. Howell, 2002):

1) Simultaneous test of: model effects + marginal and partial tests of association.
2) Test of all possible models.
3) Stepwise solution for an optimal model.

All analyses suggested that the distribution of frequencies can be explained by the final model:

$$\ln(F_{ijk}) = \lambda + \lambda^{V} + \lambda^{C} + \lambda^{VC}$$

The final model (*VC*) is nonsignificant at $p = 0{,}755$ and is nonsignificantly different from more complex models. The model does not take pedagogical role (*P*) into account and consequently excludes the two-ways interactions λ^{PV} and λ^{PC} (Pedagogical Role × Visual Style and Pedagogical Role × Visual Style) as well as the three-way interaction λ^{PVC} (Pedagogical Role × Visual Style × Communicative Style).

Knowing that the three-way interaction (λ^{PVC}) could be excluded from the model, the three hypotheses can be evaluated as they correspond to the two-way interactions λ^{PV}, λ^{PC}, and λ^{VC}.

From the log-linear analysis above we have seen that the contribution of the two-way interactions λ^{PV} and λ^{PC} can be excluded, i.e. hypotheses 1 and 2 can be rejected. Likewise the contribution of the two-way interaction λ^{VC} cannot be excluded, i.e. hypothesis 3 cannot be rejected.

Even if the hypotheses 1 and 2 are rejected, a separation of pedagogical role into the two different pedagogical roles of instructor and learning companion, together with a separation of the data according to participant gender, may provide a more detailed understanding. Thus we will set out to examine the three hypotheses one by one by evaluating their corresponding 2 × 2 contingency tables in three steps:

Step 1: Standard Pearson chi-square test on the 2 × 2 contingency table (corresponding to the tests of the two-way interactions in log-linear analyses above).

Step 2: Standard Pearson chi-square tests on the separated levels of the 'independent variable' (Pedagogical Role: Instructor / Companion, Visual Style: Naturalistic / Stylized).

Step 3: Repetition of the standard Pearson chi-square tests in Steps 1 and 2 above separated as to participant gender (Male / Female).

In addition, all the standard Pearson chi-square tests in Steps 1 to 3 have been paralleled with Yates' correction for continuity (in the cases of 2 × 2 contingency tables) and likelihood ratio tests. In all cases the different tests agreed on the results.

7.1. Choice of visual style as dependent on pedagogical role (hypothesis 1)

To evaluate hypothesis 1, the data were initially collapsed across participant gender and communicative style.

Step 1: The standard Pearson chi-square test on the two dimensional contingency table of pedagogical role and visual style gave no support for the rejection of H_0 ($\chi^2_{\text{total}} = 0{,}113$, $p = 0{,}737$), i.e. it could not be concluded that the chosen visual style is dependent on the assigned pedagogical role (Figure 11 & Table 2).

Step 2: With the data separated in consistence with the two different pedagogical roles of instructor and learning companion, the standard Pearson chi-square tests showed a dependency of the visual style choice in the learning companion condition with $\chi^2_{\text{comp}} = 3.930$ significant at $\alpha = 0.05$ (Figure 11 & Table 2).

Step 3: With the data separated according to participant gender and collapsed across communicative style, the standard Pearson chi-square tests yield-

ed the conclusion that the dependence found above originates from the female group, i.e. for the female group, the choice of visual style depends on the assigned pedagogical role in the learning companion condition with $\chi^2_{comp} = 6.545$ significant at $\alpha = 0.05$ (Figure 11 & Table 2).

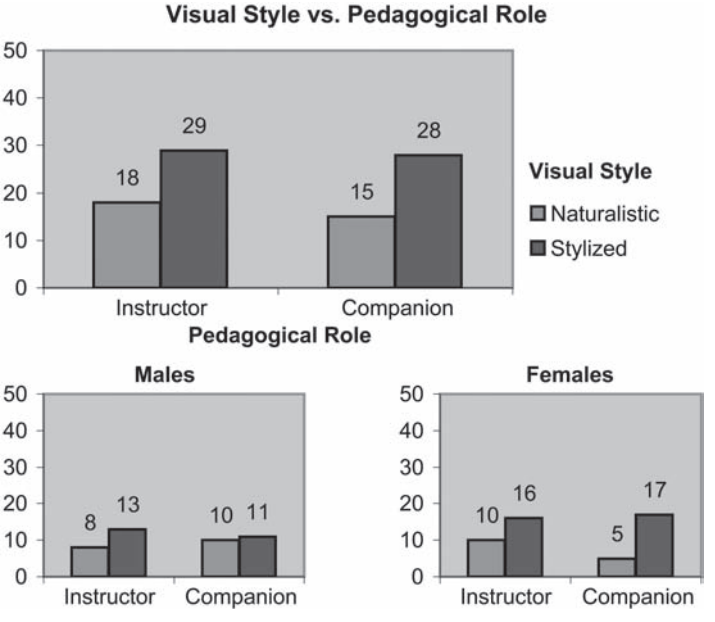


Figure 11. Frequency distributions of the visual style choice against the pedagogical role (of the embodied pedagogical agent). 'Naturalistic' stands for the detailed & 3D-rendered visual expression, 'Stylized' stands for the simplified & cartoonish visual expression.

Table 2. The standard Pearson chi-square tests on the data in Figure 11, evaluating whether the visual style choice depends on the assigned pedagogical role.

	χ^2_{total}	χ^2_{instr}	χ^2_{comp}
Both gender	0.113	2.574	3.930 *
Males	0.389	1.190	0.048
Females	1.373	1.385	6.545 *

* $p < 0.05$

7.2. Choice of communicative style as dependent on pedagogical role (hyp. 2)
Following the same procedure as above in the evaluation of hypothesis 2, the data were in this condition initially collapsed across participant gender and visual style.

Step 1: The standard Pearson chi-square test on the two dimensional contingency table of pedagogical role and communicative style showed no support for the rejection of H_0 (χ^2_{total} = 1,389, p = 0,239), i.e. it could not be concluded that the chosen communicative style is dependent on the assigned pedagogical role (Figure 12 & Table 3).

Step 2: With the data separated in consistence with the two different pedagogical roles of instructor and learning companion, the standard Pearson chi-square tests showed a dependency of the choice of communicative style in the learning companion condition with χ^2_{comp} = 5.233 significant at α = 0.05 (Figure 12 & Table 3).

Step 3: With the data separated according to participant gender and collapsed across visual style, the standard Pearson chi-square tests yielded the conclusion that the dependence found above in 'Step 2' originates from the female group, i.e. for the female group, the choice of communicative style depends on the assigned pedagogical role in the learning companion condition with χ^2_{comp} = 6.545 significant at α = 0.05 (Figure 12 & Table 3).

7.3. Relations between visual style and communicative style (hypothesis 3)
For an evaluation of hypothesis 3, the same procedure as in the two previous evaluations was once again performed. In this evaluation the data were initially collapsed across participant gender and pedagogical role.

Step 1: The standard Pearson chi-square test on the two dimensional contingency table of visual style and communicative supports the rejection of H_0 (α^2_{total} = 7,656, p = 0,006), i.e. there is a significant dependence at α = 0.01 between the chosen visual style and the chosen communicative style (Figure 13 & Table 4). This result is also in accordance with the model suggested by the log-linear analysis above.

Step 2: With the data separated in consistence with the two different visual variables of naturalistic (detailed & 3D-rendered) and stylized (simplified & cartoonish), the standard Pearson chi-square tests gave support for the conclusion that the dependence between visual style and communicative style can be attributed to the stylized condition, i.e. there is a very strong support for the conclusion that a person who chose a stylized embodied agent will also choose a task- and relation-oriented agent with χ^2_{style} = 10.965 significant at (Figure 13 & Table 4).

Step 3: With the data separated according to participant gender and collapsed across pedagogical role, the standard Pearson chi-square tests in this condition provide a more complex picture than in the two previous evaluations. A possible conclusion is that the strong support for the result of Step 2 can be attributable to both the male and the female group with $\chi^2_{style} = 4.167$ significant at $\alpha = 0.05$ and $\chi^2_{style} = 6.818$ significant at $\alpha = 0.01$ respectively (Figure 13 & Table 4).

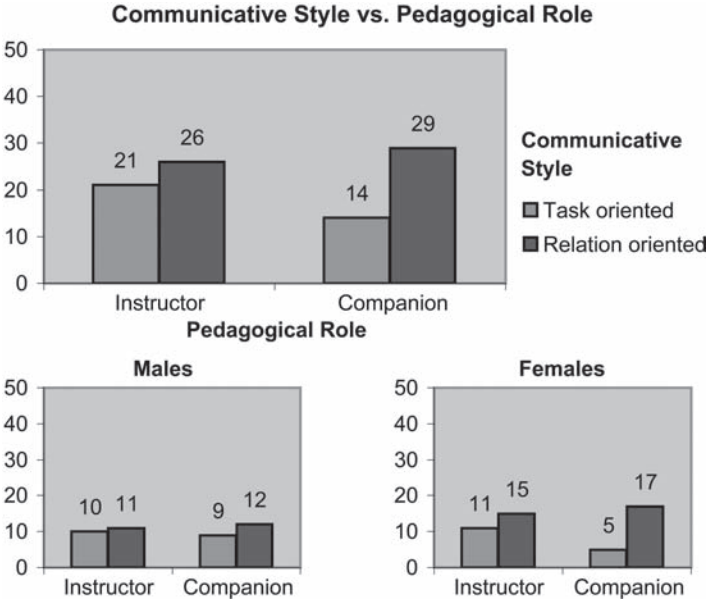


Figure 12. Frequency distributions of the communicative style choice against the pedagogical role (of the pedagogical agent).

Table 3. The standard Pearson chi-square tests on the data in Figure 12, evaluating whether the communicative style choice depends on the pedagogical role asigned to the agent.

	χ^2_{total}	χ^2_{instr}	χ^2_{comp}
Both gender	1.389	0.532	5.233 *
Males	0.096	0.048	0.429
Females	2.056	0.615	6.545 *

* $p < 0.05$

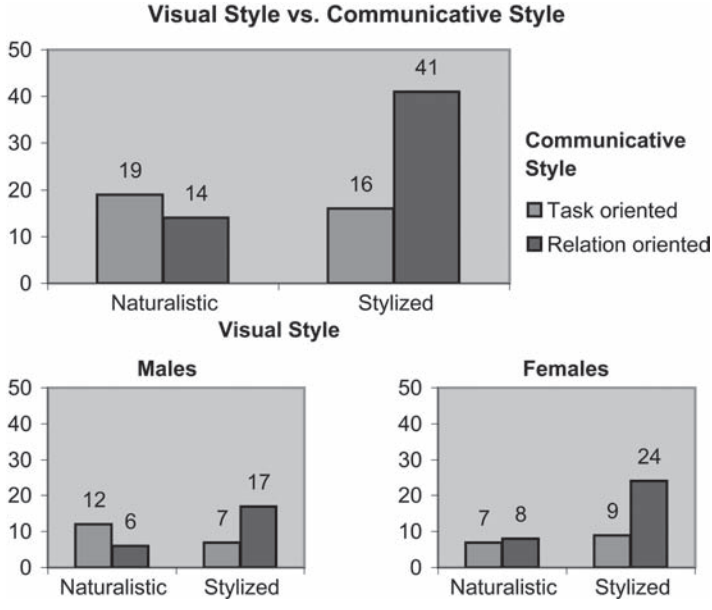


Figure 13. Frequency distributions of the visual style choice against the communicative style choice (of the pedagogical agent). 'Naturalistic' stands for the detailed & 3D-rendered visual expression, 'Stylized' stands for the simplified & cartoonish visual expression.

Table 4. The standard Pearson chi-square tests on the data in Figure 13, evaluating whether the visual style choice is related to the communicative style choice.

	χ^2	χ^2_{natur}	χ^2_{style}
Both gender	7.656 **	0.758	10.965 ***
Males	5.839 *	2.000	4.167 *
Females	1.745	0.067	6.818 **

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Summing up, the results are as follows:

- Participants who chose a naturalistic representation of the embodied pedagogical agent, showed no specific preference in the choice between a strictly task-oriented agent and a task- and relation-oriented agent.
- Participants who chose a stylized representation of the embodied pedagogical agent, showed a very strong preference for a task- and relation-oriented agent (Figure 13 and Table 4).

8. Analysis and Discussion

In what follows we first reflect on some limitations of the study. Then we analyze and discuss the three hypotheses and corresponding results. We conclude with a general discussion.

8.1. Limitations of the study

8.1.1. Implementation

One obvious limitation of the study is that the variables *pedagogical role* and *communicative style* of the agents were not implemented as actual algorithm driven agent behaviour and encountered and evaluated as such by the students.

Here we share the goal and desire of many other EPA researchers to develop and work with fully developed educational programs where learners can act and interact over time in order to increase the ecological validity of the studies. At the same time, it is not always the case that an exploratory study with a focus on certain variables and their relation is best carried out in a fully implemented running program. In the present case we may gain additional knowledge on a more generalized conceptual and articulated level concerning preferences and relations with regard to pedagogical role, visual style, and communicative style. This kind of information would be blurred and even inaccessible in an actual implementation where the information certainly would be of higher ecological value, but at the same time more case specific.

Continuing this line of argument, a desirable and logical follow-up on the present study would be to use brief implementations with interactive human-agent scenarios, systematically varying the communicative styles in relation to the pedagogical role.

It is furthermore possible to look on the limitation constituted by our 'meagre' stimuli material (as to pedagogical role and communicative style), from still another angle. Given that these stimuli could be considered 'quite weak', results in terms of relations between the variables may be even more notable.

8.1.2. The context of study and participant motivation

Another ecological validity issue relating to the context of this study is the question of whether the results would have been different if the test was embedded into a history or social science lesson, in which case the participants' need or motivation to use the program might have been more obvious.

It is possible that participants would have paid even more attention to and cared about their choices in such a case. Note, however, our use of the notion

'even more'. A common observation made by all the three experimenters was that also in the present study participants showed a commitment, focus and engagement in the tasks. We may speculate that this has to do with the fact that their regular arts teacher expressed a commitment and engagement in what was going on.

Another open question is whether students might have been somewhat more inclined to choose a strictly task-oriented pedagogical agent than in our study if the context had been a more regularly scheduled activity that they wanted to 'get done with' in order to continue with next activity.

8.1.3. The age group

One should furthermore be careful not to generalize from the actual group of participants to users or learners in general. The participants were 12-15 year old Swedish adolescents. A replication of the study in a group of younger school children or in a group of university students, and/or in another culture might – or might not – produce different results. We simply cannot tell.

8.2. Pedagogical role and visual style (hypothesis 1)

The evaluation of hypothesis 1 indicated that for the female group the pedagogical role of the agent in the 'learning companion condition' conveyed a preference for a visually stylized embodied agent. For the male group there were no results indicating any form of relation between pedagogical roles and visual style preferences.

As already observed the pedagogical role of the agent was a quite weak stimulus in the sense that participants were simply being *told* that they were to interact with either a chief editor (instructor condition) or a companion journalist (learning companion condition) together with a description of what this involved. This method can be criticized for low ecological validity, but the fact that such a weak stimulus could still yield a strong (even if partial) effect is notable.

Furthermore, although this is not a straightforward and unambiguous support for hypothesis 1, one can note the shift from a non-significant distribution (instructor condition) to a significant distribution (learning companion condition) regarding the choice of visual style. In the learning companion condition, a stylized character is clearly preferred, in coherence with hypothesis 1.

Why would participants prefer a simplified and cartoonish learning companion agent? Based on previous discussions in this article, a possible explanation relates to the long tradition in comics to use simplified and stylized characters to emphasize socio-emotional expression and facilitate subjective identification and immersion into the characters and the story (McCloud, 1993). If partici-

pants associate a learning companion with a 'friend' with whom they can have a personal relationship (in contrast to a strictly task-oriented, authoritative boss or instructor) it may then be close at hand for them to select a more simplified and cartoonish (stylized) character rather than a more detailed and naturalistic one (Gulz & Haake, 2005).

But it should be emphasized that such a choice pattern, of course, can be overturned if other important visual aspects of the alternative characters are not equivalent, for instance age, clothing and friendliness, since such variables may affect the choice as well. For the characters in our study we explicitly sought to keep such aspects equivalent (see section 'Additional design comments'), and pre-tests in the age-group in question were conducted to ensure that none of the characters would 'stand out' in relation to the others. Of particular importance for the present study was that no characters in their appearance would stand out with regard to the aspect of 'friend' v. 'boss'. Thus the visual aspect of 'friendliness' vs. 'non-friendliness' was of particular importance. In retrospect, we believe it would have been a good idea with an additional and more targeted focus on this in the pre-tests. However, the pre-tests as carried out gave no indications of deviations with respect to 'friendliness' vs. 'non-friendliness'.

To illustrate why choice patterns like the one we have described cannot be generalized in an unqualified way, consider the following hypothetical situation: Your task is to assign two given visual characters – a naturalistic version of Calvin (in Calvin and Hobbes) and a cartoon character in Dick Tracy style (c.f. Figure 4) – to two different given roles, namely your young companion/pal and a tough boss of a criminal gang. In such a case we would not predict that people would choose the Dick Tracy style character to be the companion and the naturalistic version of Calvin to be the criminal boss.

8.3. Pedagogical role and communicative style (hypothesis 2)

Also here the move from the 'instructor condition' to the 'learning companion condition' brings about a change in the preference distribution in favour of a task- and relation-oriented agent, and significantly so for the female participants. Thus, the results may be interpreted as partially in coherence with hypothesis 2.

Again one should consider the relatively weak stimuli as to the 'pedagogical roles' in combination with the same weakness in the stimuli of the 'communicative roles'. This makes the interpretation of the results more intriguing and at the same time the strong partial significant dependence even more conspicuous.

One interpretation of the results – in line with *the Media Equation* of Reeves and Nass (1996) – is that social strategies and responses that we apply in the real

world are being transferred into the virtual world, in this case real world relational strategies. This interpretation is supported by the observation that when the pedagogical agent in our study shifts from a more *formal* social role (instructor) to a more *relational* social role (learning companion), the tendency to turn to real life relational strategies and transfer them into the virtual world becomes more distinct.

If we also pay attention to the fact that the significance of the result arises from the female group, we can speculate on gender-specific behaviours and social strategies. In the real world girls/females are considered more competent in using subjective relational strategies, whereas boys/males rely more on formal, objective strategies for social interaction (Walker, 2005). Perhaps this pattern is seen reproduced here in a virtual setting.

8.4. Visual style and communicative style (hypothesis 3)

The results concerning the relation between the choices of visual style and of communicative style were statistically more distinct than for the two previous hypotheses. Given that a participant had chosen a stylized (simplified & cartoonish) embodied agent, he or she showed a significant tendency to choose a 'task- and relation-oriented' agent. The significance persisted both when collapsed across gender and when separated as to gender. For the male group, furthermore, there was also a tendency for a participant who had chosen a naturalistic (detailed & 3D-rendered) embodied agent to choose a strictly task-oriented agent.

Our interpretations of these relations between preferences for visual and preferences for communicative style of the embodied agent, as presented in the section 'Realism and pedagogical effects', propose that *subjective identification* may be facilitated with simplified and cartoonish (stylized) characters. A visually realistic embodied agent, on the other hand, promotes objectivity. Consequently a stylized character is more easily associated with a subjective and personal relation, as expressed in relation-oriented communication, and a realistic character is more easily associated with strictly task-oriented communication. Support in this direction comes from Nowak and Biocca's (2003) study, involving relatively realistic versus heavily stylized characters in VR-environments. Here users rated the stylized characters significantly higher than the naturalistic ones as to their experience of a psychological connection with the character, in terms of co-presence and social presence.

As a rule of thumb or design consideration to reflect upon, we propose that an embodied non-authoritative learning companion agent can in many cases be

favourably designed to be visually simplified and cartoonish (stylized) as well as to be communicatively task- and relation-oriented.

8.5. General discussion

There is an overall need for more studies in order to assemble evidence from which guidelines about different aspects of EPAs can be developed. To accomplish this, it is of importance that studies are detailed and specific in the description of the variables they examine.

In this article, we have proposed a theoretical framework and a method to specifically address different aspects or variables of *visual static appearance* in evaluations of EPAs. The purpose is to augment the granularity in discussions and evaluations of characteristics of embodied agents in general, and EPAs specifically.

From a broader perspective the article may be seen as a contribution to a 'visual design guidelines project' in the sense suggested above. It should be emphasized that we are not referring to generalized step-by-step recipes (guidelines) on 'how to visually design an EPA'. We believe this would be a futile objective since the adequacy of a visual design depends on such a complex set of variables including the learning context, the learning goals, and the group of learners, in combination with – as indicated by the study presented in this article – the pedagogical role and communicative style of an EPA. What we believe is desirable and possible, though, is to delimit and encompass the visual design space by means of guidelines in the form of high-level *design topics* or *design considerations*: 'These variables may relate to each other in the following ways…', 'Given equivalence in other relevant visual design aspects, these different design alternatives can have the following consequences….' and 'This is a topic that ought to be reflected upon before making a choice on a certain visual parameter'. We are not implying that designers of EPAs could thereby refrain from user evaluations, but believe that these kinds of guidelines can be a support for designers and developers, as well as researchers, in order to navigate in the immense visual design space that is there at the outset, and as a support through the design process.

Such a 'visual design guidelines project' would also be a useful counterbalance to the seeming elusiveness of visual appearance aspects in the area of embodied agents. Even though there are, admittedly, important ingredients of subjective evaluation regarding visual aspects, this is not all there is to it. The visual appearance of EPAs is a topic that should *not* be neglected by scientific and pedagogical perspectives and left to the opinions and gut feelings of individual designers or 'the market'. For this, it is far too important and influential.

Acknowledgements

Parts of this work were supported by the Swedish Research Council (VR). Thanks to Martin Jonasson for collaboration in the study, and also thanks to anonymous reviewers for highly valuable comments on previous versions of this paper. (One of them being so helpful that he/she would deserve to be included as contributor, had this been possible.)
And not least, thanks to all student participants in the study, and to helpful teachers.

Copyrights

Figure 1. *MS Agents* © Microsoft Corporation.
Figure 2. The six characters to the right where designed by means of the *SitePal Demo Tool* (www.sitepal.com) © Oddcast Inc.
Figure 3. © Scott McCloud. Reproduced with permission from HarperCollins Publishers, Inc. The portrayed figures are all copies (by Scott McCloud) of the original drawings. All copyrights held by the creators.
Figure 4. From left to right: Chester Gould's *Dick Tracy* © Chicago Tribune – New York Syndicate, Carl Barks' *Scrooge McDuck* © Walt Disney Production, Reed Richards and Cable © Marvel Entertainment Group, © José Munoz and Carlos Sampayo, © Art Spiegelman, © Will Eisner.
Text by Scott McCloud (McCloud, 1993, p. 126). Reproduced with permission from HarperCollins Publishers, Inc.
Figure 7. The interactive information assistant *Olga* was developed by: CID NADA KTH, TMH KTH, Lingvistik SU, SICS & Nordvis AB. http://www.nada.kth.se/~osu/olga/

References

Aïmeur, E., & Frasson, C. (1996). Analyzing a new learning strategy according to different knowledge levels. *Computers and Education*, 27 (2), 115-127.

André, E., Rist, T., & Müller, J. (1998). Integrating reactive and scripted behaviors in a life-like presentation agent. In K. Sycara & M. Woolridge (Eds.) *Proc. of the Second Int. Conf. on Autonomous Agents* (pp. 261-268). New York, NY: ACM.

Bailenson, J., Blascovich, J., Beall, A., & Loomis, J. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin*, 29, 1-15.

Bailenson, J., Swinth, K., Hoyt, C., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 14, 379-393.

Baylor, A. (2005a). Preliminary design guidelines for pedagogical agent interface image. In J. Riedl, A. Jameson, D. Billsus & T. Lau (Eds.) *Proc. of the 10th Int. Conf. on Intelligent User Interfaces* (pp. 249-250). New York, NY: ACM.

Baylor, A. (2005b). The impact of pedagogical agent image on affective outcomes. In *Proc. of the Workshop on Affective Interactions: Computers in the Affective Loop*. 10th Int. Conf. on Intelligent User Interfaces 2005, San Diego, CA.

Baylor, A., & Kim, Y. (2004). Pedagogical agent design: the impact of agent realism, gender, ethnicity, and instructional role. In J. Lester, R-M. Vicari & F. Paraguaçu (Eds.) *LNCS 3220: Intelligent Tutoring Systems* (pp. 592-603). Berlin/Heidelberg: Springer.

Baylor, A., & Kim, Y. (2005). Simulating instructional roles through pedagogical agents. *Int. J. of Artificial Intelligence in Education*, 15 (1), 95-115.

Baylor, A., & Plant, E. (2005). Pedagogical agents as social models for engineering: the influence of appearance on female choice. In C-K. Looi, G. McCalla, B. Bredeweg & J. Breuker (Eds.) *Artificial Intelligence in Education – Supporting Learning through Intelligent and Socially Informed Technology* (pp. 65-72). Amsterdam: IOS Press.

Baylor, A., Rosenberg-Kima, R., & Plant, E. (2006). Interface agents as social models: the impact of appearance on females' attitude toward engineering. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems* (pp. 526-531). CHI 2006: Conf. on Human Factors in Computing Systems, Montréal, Canada.

Berman, R., & Slobin, D. (1994). Different ways of relating events: introduction to the study. In R. Berman & D. Slobin (Eds.) *Relating Events in Narrative. A Crosslinguistic Study* (pp. 1-16). Hillsdale, NJ: Lawrence Erlbaum.

Bickmore, T. (2003). *Relational Agents: Effecting Change through Human-Computer Relationships*. Ph.D. Thesis, Media Arts & Sciences, Massachusetts Inst. of Technology, Cambridge, MA.

Biswas, G., Schwartz, D., Leelawong, K., Vye, N., & Teachable Agents Group at Vanderbilt (2005). Learning by teaching: a new agent paradigm for educational software. *Applied Artificial Intelligence*, 19 (3-4), 363-392.

Chan, T-W., & Chou, C-Y. (1997). Exploring the design of computer supports for reciprocal tutoring. *Int. J. of Artificial Intelligence in Education*, 8, 1-29.

Chan, T-W., Chung, Y-L., Ho, R-G., Hou, W-J., & Lin, G-L. (1992). Distributed learning companion systems – WEST revisited. In C. Frasson, G. Gauthier & G. McCalla (Eds.) *LNCS 608: Intelligent Tutoring Systems* (pp. 643-650). Berlin/Heidelberg: Springer.

Chang, L-J., Wang, J-C., Hus, B-Y., & Chan T-W. (1999). Four applications of student modeling – my animal companions. In *Proc. of the 3rd Global Chinese Conf. on Computers in Education 1999* (p. 366-370). GCCCE'99, Macau, China.

Chou, C-Y., Chan, T-W., & Lin, C-J. (2000). Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, 40, 255-269.

Cook, M. (1979). *Perceiving others: the psychology of interpersonal perception*. London: Methuen.

Daft. R., & Lengel, R. (1986). Organizational information requirement, mediarichness and structural design. *Management Science*, 32, 554-571.

Dennis, A., & Kinney, S. (1998). Testing media richness theory information systems research, the effects of cues, feedback, and task equivocality. *Information Systems Research*, 9 (3), 256-274.

D'Mello, S., Craig, S., Sullins, J., & Graesser, A. (2006). Predicting affective states expressed through an emote-aloud procedure from Autotutor's mixed-initiative dialogue. *Int. J. of Artificial Intelligence in Education*, 16, 3-28.

Galotti, K., Clinchy, B., Ainsworth, K., Lavin, B., & Mansfield, A. (2001). A new way of assessing ways of knowing: the attitudes towards thinking and learning survey (ATTLS). *Sex Roles*, 40 (9/10), 745-766.

Gard, T. (2000). Building Character. *Gama Network: Gamasutra.com*. http://www.gamasutra.com/features/20000720/gard_pfv.htm (2008-07-17).

Gratch, J., Marsella, S., Egges, A., Eliëns, A., Isbister, K., Paiva, A., Rist, T., & ten Hagen, P. (2004). Design criteria, techniques and case studies for creating and evaluating interactive experiences for virtual humans. In *Report (Paper 04121): Dagstuhl seminar 04121 on Evaluating Embodied Conversational Agents*. Schloss Dagstuhl, Germany. http://wwwhome.cs.utwente.nl/~zsofi//eeca/WG2article.pdf (2008-07-15)

Gulz, A., & Haake, M. (2005). Social and visual style in virtual pedagogical agents. In *Proc. of the Workshop on Adapting the Interaction Style to Affective Factors*. 10th Int. Conf. on User Modeling 2005, Edinburgh, Scotland.

Gulz, A., & Haake, M. (2006). Design of animated pedagogical agents – a look at their look and visual form. *Int. J. of Human-Computer Studies*, 64 (4), 322-339.

Gulz, A., Ahlner, F., & Haake, M. (2007). Visual femininity and masculinity in synthetic characters & patterns of affect. In A. Paiva, R. Prada & R. Picard (Eds.) *LNCS 4738: Affective Computing and Intelligent Interaction* (pp. 654-665). Berlin/Heidelberg: Springer.

Gustavsson, E., & Czarniawska, B. (2004). Web Woman: the on-line construction of corporate and gender images. *Organization*, 11 (5), 651-670.

Haake, M. (2006). *Virtual Pedagogical Agents – Beyond the Constraints of the Computational Approach*. Licentiate Thesis, Lund University, Lund, Sweden.

Haake M., & Gulz, A. (2006). Visual design of virtual pedagogical agents: naturalism versus stylization in static appearance. In *Proc. of the 3rd Int. Design and Engagability Conf.*. NordiCHI 2006, Oslo, Norway.

Haake, M., & Gulz, A. (accepted). Visual stereotypes and virtual pedagogical agents. To appear in *Educational Technology & Society*.

Hietala, P., & Niemirepo, T. (1998). The competence of learning companion agents. *Int. J. of Artificial Intelligence in Education*, 9, 178-192.

Howell, D. (2002). *Statistical Methods for Psychology (5th Ed.)*. Pacific Grove, CA: Duxbury.

Isbister, K. (2006). *Better Game Characters by Design – A Psychological Approach*. San Francisco, CA: Morgan Kaufmann.

Janvier, W., & Ghaoui, C. (2003). Using Communication Preference and Mapping Learning Styles to Teaching Styles in the Distance Learning Intelligent Tutoring System – WISDeM. In V. Palade, R. Howlett & L. Jain (Eds.) *LNCS 2774: Knowledge-Based Intelligent Information and Engineering Systems* (pp. 185-192). Berlin/Heidelberg: Springer.

Johnson, L. (2003). Interaction tactics for socially intelligent pedagogical agents. In L. Johnson, E. André & J. Domingue (Eds.) *Proc. of the 8th Int. Conf. on Intelligent user interfaces* (pp. 251-253). New York, NY: ACM.

Johnson, L., Rickel, J., & Lester, J. (2000). Animated pedagogical agents: face-to-face interaction in interactive learning environments. *Int. J. of Artificial Intelligence in Education*, 11, 47-78.

Johnson, L., Mayer, R., André, E., & Rehm, M. (2005). Cross-cultural evaluation of politeness in tactics for pedagogical agents. In C-K. Looi, G. McCalla, B. Bredeweg & J. Breuker (Eds.) *Artificial Intelligence in Education – Supporting Learning through Intelligent and Socially Informed Technology* (pp. 298-305). Amsterdam: IOS Press.

Kalick, S. (1988). Physical attractiveness as a status cue. *J. of Experimental Social Psychology*, 24, 469-489.

Kang, S., Watt, J., & Ala, S. (2008). Communicators' perceptions of social presence as a function of avatar realism in small display mobile communication devices. In *Proc. of the 41st Annual Hawaii Int. Conf. on System Sciences 2008*. Washington. DC: IEEE Computer Society.

Kemper, S., Rash, S., Kynette, D., & Norman, S. (1990). Telling stories: the structure of adults' narratives. *European J. of Cognitive Psychology*, 2 (3), 205-228.

Lassetter, J. (1987). Principles of traditional animation applied to 3D computer animation. In M. Stone (Ed.) *Proc. of the 14th Annual Conf. on Computer Graphics and interactive Techniques* (pp. 35-44). New York, NY: ACM.

McCloud, S. (1993). *Understanding Comics: The Invisible Art*. New York, NY: HarperPerennial.

McQuiggan, S., Mott, B., & Lester, J. (2008). Modeling self-efficacy in intelligent tutoring systems: an inductive approach. *User Modeling and User-Adapted Interaction*, 18 (1-2), 81-123.

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: a test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology*, 31, 186-207.

Moreno, R., Mayer, R., Spires, H., & Lester, J. (2001). The case for social agency in computer-based teaching: do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19 (2), 177-213.

Myers, S. (1998). Instructor socio-communicative style, argumentativeness, and verbal aggressiveness in the college classroom. *Communication Research Reports*, 15, 141-150.

Nowak, K., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. Presence, 12 (5), 481-494.

Nowak, K., & Rauh, C. (2005). The Influence of the Avatar on Online Perceptions of Anthropomorphism, Androgyny, Credibility, Homophily, and Attraction. *J. of Computer Mediated Communication*, 11 (1), article 8. http://jcmc.indiana.edu/vol11/issue1/ (2008-07-15)

Nowak, K., & Rauh, C. (2008). Choose your 'buddy icon' carefully: The influence of avatar androgyny, anthropomorphism and credibility in online interactions. *Computers in Human Behavior*, 24 (4), 1473-1493.

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

Ruttkay, Z., Dormann, C., & Noot, H. (2004). Embodied conversational agents on a common ground. In Z. Ruttkay & C. Pelachaud (Eds.) *From Brows to Trust: Evaluating Embodied Conversational Agents* (pp. 27-66). Norwell, MA: Kluver Academic Publishers.

Ryokai, K., Vaucelle, C., & Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *J. of Computer Assisted Learning*, 19, 195-208.

Schneider, D., Hastorf, A., & Ellsworth, P. (1979). *Person Perception*. Reading, MA: Addison-Wesley.

Shaw, E., Johnson, W., & Ganeshan, R. (1999). Pedagogical agents on the web. In O. Etzioni, J. Müller & J. Bradshaw (Eds.) *Proc. of the third annual Conf. on Autonomous Agents* (pp. 283-290). New York, NY: ACM.

Sheldon, W., Stevens, S., & Tucker, W. (1940). *The Varieties of Human Physique: An Introduction to Constitutional Psychology*. New York, NY: Harper & Brothers.

Walker, S. (2005). Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. *The J. of Genetic Psychology*, 166 (3), 297-312.

Welch, R., Blackmon, T., Liu, A., Mellers, B., & Stark, L. (1996). The effects of pictorial naturalism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence: Teleoperators and Virtual Environments*, 5 (3), 263-273.

Wilson, M. (1997). Metaphor to personality: the role of animation in intelligent interface agents. In *Proc. of the Workshop on Animated Interface Agents: Making them Intelligent. 15th Int. Joint Conf. on Artificial Intelligence 1997*, Nagoya, Japan.

Yee, N., & Bailenson, J. (2007). The Proteus Effect: self transformations in virtual reality. *Human Communication Research*, 33. 271-290.

Paper IV

Visual femininity and masculinity in synthetic characters & patterns of affect

Gulz, A. Ahlnér, F. & Haake, M. (2007)
In: A. Paiva, R. Prada & R. Picard (Eds.)
Lecture Notes in Computer Science (LNCS) 4738, Proceedings of the Second International Conference on Affective Computing and Intelligent Interaction
Lisbon, Portugal, September 2007, pp. 654-665
Berlin/Heidelberg, Germany: Springer

doi: 10.1007/978-3-540-74889-2_57
www.springerlink.com

Visual femininity and masculinity in synthetic characters & patterns of affect

Agneta Gulz[a], Felix Ahlner[a], Magnus Haake[b]

[a] *Division of Cognitive Science, Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden*

[b] *Department of Design Sciences, Lund University, P.O.Box 118, 221 00 Lund, Sweden*

Original paper published in:
LNCS 4738: Affective Computing and Intelligent Interaction, Proceedings of ACII 2007, pp. 654-665
Berlin/Heidelberg, Germany: Springer.

Abstract: It has been shown that users of a digital system perceive a more 'masculine-sounding' female voice as more persuasive and intelligent than a corresponding but more 'feminine-sounding' female voice. Our study explores whether a parallel pattern of affectively coloured evaluations can be elicited when *femininity and masculinity* are manipulated via *visual* cues instead of via voice. 80 participants encountered synthetic characters, visually manipulated in terms of femininity and masculinity but with voice, spoken content, linguistic style and role of characters held constant. Evaluations of the two female characters differed in accordance with stereotype predictions – with the exception of competence-related traits; for the two male characters evaluations differed very little. The pattern for *male* versus *female* characters was slightly in opposite to stereotype predictions. Possible explanations for these results are proposed. In conclusion we discuss the value of being aware of how different traits in synthetic characters may interact.

1 Background

As Reeves, Nass and collaborators [16] have shown, many of the *affective responses* that human beings evoke in one another are likewise evoked by different kinds of media, as soon as the media involves a social cue such as natural language interaction, voice, something face-like, etc. In response to these social cues, humans

tend to experience, for instance, various degrees of trust, irritation and empathy, feeling flattered, humiliated, distanced, and so on. Furthermore, [16] documents how certain *response patterns*, known from social psychology, are paralleled in responses towards media. As an example, extroverted people tend to respond more positively towards others that speak relatively loud and with varied pitch than those who speak quietly and with little variation in pitch, and the opposite is the case for introverted people. This pattern can be reproduced towards a computer application with a voice interface [16].

An explicitly anthropomorphic form of digital media is that of synthetic characters. These characters populate the digital society in increasing number, and there is no doubt that they do evoke affective responses. But to what extent does the embodiment in the sense of a *visualized* humanlike character influence the affective responses that *appear already* towards the same system without this visualized character? And can affective response patterns known from social psychology be elicited by manipulating only the *static visual appearance* of the character, while controlling for voice, content and visual dynamics such as movements and facial expressions?

1.1 The role of static visual characteristics

In line with Ruttkay and collaborators' [17] analysis of the *physical appearance* of a synthetic character we distinguish between:

- *static visual characteristics,* including: (i) degrees of anthropomorphism and degrees of visual naturalism/realism; (ii) basic physical properties such as body-type, face properties, skin, hair and hair-cut, clothes and attributes – characteristics that, furthermore can make up representations of age, gender and ethnicity
- *dynamic visual characteristics* such as facial expressions, body and hand gestures, postures, movements, etc.

Whereas visual dynamics have been extensively researched, relatively little attention has been paid to the static visual appearance or *look* of synthetic characters [7, 8], even though there are important exceptions [3, 4, 11, 14, 22, 24][1]. It is, however, well established in the case of *human-human* interaction that static visual appearance plays an important role for the eliciting of affective responses and evaluations. In an encounter with another person, visual cues as to the persons

[1] Furthermore, in [3] and [24], just like in the study to be presented, aspects of visual femininity are addressed and manipulated. Baylor and collaborators [3] compare responses towards a feminine-looking and a "geek"-looking (virtual) female engineer. Waern and collaborators [24] vary body shape and clothing in terms of 'hypersexuality' in female avatars.

face, body shape, height, ethnicity, hair, dressing, etc. are immediately picked up and exploited to guide our approach to this person [18]. Among those very quick, and affectively colored, evaluations are, for instance, estimates of how friendly, how trustworthy, how dominant, and how similar to oneself somebody is [7, 14]. Our focus is on an aspect that is crucial for socio-affective responses and interaction in a human-human context, namely *gender*.

1.2 Gender and affective responses

One of the very first things we do when we encounter someone is to determine whether the person is female or male. This categorization, in turn, colors several consequent responses and evaluations. Analyses reveal, that peoples' evaluations of an identical proposal, narrative, message etc. as presented by a man versus by a woman (for instance, by a female or a male author; in a woman's or in a man's voice) often differ in accordance with gender stereotypes [1, 5]. A content mediated by a man is perceived as more credible than the same content mediated by a woman, and a male mediator is perceived as more competent than a female mediator[2].

That such evaluative patterns can also be replicated towards computer-based systems has been shown by Reeves, Nass and collaborators. In one study [16], participants perceived a (visually non-present) digital tutor as significantly more competent and credible when it had a male voice than when it hade a female voice. Participants were more satisfied and found the praise from the tutor more valuable and trustworthy when the tutor had a male voice than when it had a female voice.

Other gender aspects, applicable both to males and females, are degrees of *femininity* and of *masculinity*. Here, Reeves, Nass and collaborators [16, 23] performed a study where screen depictions of six women were accompanied by female voices electronically altered to be either 'feminine-sounding' or 'masculine-sounding'. The former were amplified in the higher and reduced in the lower frequencies and were played at softer volume. For masculine voices, the lower frequencies were amplified, the higher ones dampened, and the voices were made louder than average. Participants encountered three screen depictions with a feminine-sounding voice and three with a masculine-sounding voice. Results were that the evaluative patterns reproduced patterns known from social psychology. Masculine-sounding women were perceived as significantly more persuasive and intelligent than feminine-sounding women.

[2] This is the case for 'neutral' content. Patterns look different with respect to contents that are considered to belong to 'typically female domains' (c.f. [16]).

As to the visual images used in the study, theses are neither included nor described in the thesis [23], and thus we know nothing about femininity/masculinity as *visually* represented in these images. It is, however, precisely this aspect that is the focus of the present study: Can femininity/masculinity *as visually coded* influence affective evaluations regarding persuasiveness, intelligence, empathy, warmth etc. of synthetic characters? In contrast to the [23] study we focus on evaluations of animated characters (even though the *manipulated* aspect is the underlying static visual character).

1.3 Issues addressed in the study

The inspiration to our study came from the above-mentioned study [23] in combination with our own line of research. Its primary goal was to investigate whether a gender stereotypical pattern of affective evaluations could be elicited when degrees of femininity and masculinity were manipulated visually instead of via voice cues. Would evaluations of intelligence, empathy, warmth, trustworthiness, etc. also in this case vary according to gender stereotypes known from human-human interaction?

Table 1. The parameters used in the questionnaire (*translated from Swedish*), categorized into five groups and ordered as opposites with stereotypic feminine traits (to the left) vs. stereotypic masculine traits (to the right). The order also corresponds to the order used in Table 2, 3 & 4.

Trustworthiness / Expertise	**Personality (Communication)**
ignorant vs. knowledgeable (L1)	personal vs. impersonal (L3)
non-expert vs. expert (L4)	subjective vs. objective (L11)
un-experienced vs. experienced (L9)	
undecided vs. decided (L13)	**Intelligence (talent)**
non-persuasive vs. persuasive (L17)	non-intelligent vs. intelligent (L6)
Personality (Empathy)	**Self-confidence**
warm vs. cold (L5)	unsure of oneself vs. self-confident (L7)
empathic vs. non-emphatic (L16)	submissive vs. dominant (L8)
friendly vs. unfriendly (L18)	
pleasant vs. unpleasant (L20)	

In addition to the parameters 'intelligence' and 'persuasiveness' used in [22] we included twelve additional parameters (se above) for which peoples' conceptions and attitudes are known to relate to gender stereotypes [1, 5, 15]. Thus, the question was whether our manipulations of degree of femininity/masculinity

via visual cues would be reflected in evaluations of the synthetic characters with respect to the gender stereotypical traits listed in Table 1.

2 The study

2.1 Study design

To address these issues we created eight conditions, which each consisted of an encounter with one female and one male synthetic character. The female character should be either a feminine-looking character (F+) or a more neutral, slightly masculine-looking, female character (F-). The male character should be either a masculine-looking male character (M+) or and a slightly feminine-looking male character (M-).

Both characters should speak about a similar, gender-neutral, topic divided in two subsequent parts. After listening to each character, the participants should report their impressions of the character by means of a Likert scale questionnaire.

The participants were randomly assigned one of the eight conditions: [M+ F-], [M+ F+], [M- F+], [M- F-], [F+ M-], [F+ M+], [F- M-], [F- M+].

2.2 Materials

Visual characters

A number of images for each character were designed on the basis of the *SitePal* on-line demo design tool [19]. The images were further manipulated in *Photoshop*, to establish an image set consisting of: a body, about 10 different views of the face and about 10 different mouth shapes. The animation (including lip synchronization) was done in *Flash* using frame-by-frame animation. The experiment was run on a web page, with the animation incorporated as a *Flash* movie.

The design of the characters, exploited a number of visual cues, as follows (c.f. Figure 1):

– *Feminine character (F+):* Manipulated with feminine attributes such as: the *baby-face* scheme (rounded head shapes, bigger eyes, smaller nose, narrower shoulders); long (coloured) hair and make up, that pronounces feminine attributes by enlarging the eyes, making them rounder and more distinct and making the lips fuller.

– *Weak feminine character (F-):* Manipulated with masculine attributes such as: broader head, a more angular and pronounced jaw, a high for-head; paler colours as to eyes, mouth and hair, which weakens the impact of these

female attributes; overall paler colour scheme reducing the number of distinct features and thus weakening any categorization of gender – whether feminine or masculine.

Figure 1. Screen shots of the four characters used in the study. Upper left: (F+); upper right: (F); lower left: (M-); lower right: (M+).

– *Weak masculine character (M-):* Manipulated with feminine attributes such as: rounder and less pronounced shapes of head, jaw and nose; narrower shoulders; slightly red lips in combination with an overall paler colour scheme, that weakens any distinct categorization of gender.

– *Masculine character (M+):* Manipulated with masculine attributes such as: broader, angular and more pronounced head shapes; broader shoulders, a distinct Adam's apple, pronounced, dark eye brows; neatly done hair; a more prominent colour scheme which produces distinct features and strengthens the categorization with respect to gender.

Since we were interested in unreflective and affectively mediated evaluations, we wanted to avoid the use of over-explicit visual stereotypes – i.e. *extremely* feminine-, non-feminine-, masculine- or non-masculine-looking characters – since these are liable to initiate conscious deliberation as to what is an 'appropriate' answer, etc.

The characters were pre-evaluated in order to ensure that they were perceived as intended in terms of femininity and masculinity. Forty subjects were presented

with pictures of one male and one female character, without voice, animation and background. They were asked whether they thought the male character looked masculine, neutral or non-masculine, and whether they thought the female character looked feminine, neutral or non-feminine. Results were that 14 out of 20 categorized M+ as masculine; 1 did so with M-. For the female characters, 11 out of 20 categorized K+ and 4 categorized K- as feminine. In this pre-study, note, no subject saw both females or both males. In a brief follow-up 13 additional subjects were shown all four characters and were asked which female character they thought looked more feminine and which male character they thought looked more masculine. Twelve unhesitatingly categorized the characters according to our prediction. One categorized K- as more feminine, explaining that "she looks more like a mother". A limitation of the pre-validation was that (for practical reasons) all subjects were 20 years and older, whereas half of the participants in the main study were teenagers.

An issue brought up in the pre-validating process was this: What counts as an appropriate and satisfactory pre-validation of a character design when the target is not conscious and deliberate responses towards it, in this case specifically towards visual femininity/masculinity – but instead unreflective and non-deliberate responses? These kinds of responses, note, are known to differ considerably from one another [6, 18].

Animations

Frame-by-frame animations makes it difficult and time consuming to produce exact and identical movements in shifts of head position, lip movements (synchronization), eye blinking, and breathing – nevertheless these parameters have been considered and are more or less the same for all four characters (see demo [25]).

Voice

Since we wanted to keep femininity/masculinity aspects as mediated through voice constant, one and the same female voice was to be used for the F- and F+ characters, and one and the same male voice for the M- and M+ characters. However, *inconsistencies* in virtual characters can disturb and irritate users. It was therefore important to choose a female voice that would fit *both* female characters, and a male voice that would fit *both* male characters (especially since we were interested in affectively coloured evaluations and did not want to over-layer those with affect in terms of irritation or confusion due to a mismatch between voice and visual design). The challenge was thus to find: (i) a female voice that

was neither pronouncedly feminine nor pronouncedly non-feminine; and (ii) a male voice that was neither too masculine nor too non-masculine. We consulted two speech therapists (whereof one specialized in voice and gender) that arranged a contact with two people, with voices suitable according to our criteria.

Script

The verbal script had two parts, and the aim was to have them as equivalent as possible with respect to masculinity/femininity. The subject domain chosen was medicine, which within Swedish culture is a relatively gender-neutral subject domain. The original text was taken from a Swedish popular science magazine [21], and was adapted in order to be suitable for oral presentation. Furthermore, the text was modified in order to obtain equivalence in the two parts in terms of length, number of facts and of linguistic style, especially with respect to differences in male versus female linguistic styles [2, 15]. The script parts were then pre-tested and validated in the sense that a number of readers and listeners could not decide whether they were written by a man or a woman. The final script parts were each about 2 minutes in length.

Questionnaire

The questionnaire consisted of a series of twenty 7-point Likert-type scales with opposite traits on the two poles. Since some of the trait pairs include one more negative and one more positive trait, the positive traits were evenly placed to the left and to the right, respectively. *Fourteen* of the twenty scales concerned the gender stereotype related traits that we were interested in – those presented in table 1.

The six scales omitted here relate to a parallel study on perceived attractiveness, which is outside the scope of the present article. For reference, see [12].

2.3 Participants

Forty adolescents (aged 13-18) and 40 adults (aged 25-65) participated in the study. Within each age group there were an equal number of females and males.

2.4 Procedure

A laptop was used, and the sessions took place at different, but always calm and quiet locations with only the participant and the experimental leader present. Sessions lasted on average 25 minutes. The participants were informed that they would get to listen to two computer characters speaking about diurnal rhythm and on what happens in the body during shift work. They were told that they

would first listen to doctor Elm and thereafter to doctor Ask, and that after having listened to each doctor character they would be asked questions about their impressions. Participants were also told that there would be no memory or knowledge tests, but that we were only interested in their impressions and opinions. Furthermore, they were told that they could interrupt their participation at any time. Which of the 8 conditions (order and combination of characters) was assigned to a participant was decided according to a random table. Observe that no participant encountered the two female, or the two male characters but always one male and one female character. Upon each of the two presentations, the participants filled out the (identical) questionnaire. When the second and last questionnaire had been filled out, a brief interview was carried out with the participant on masculinity and femininity. At this stage, all four visual characters were shown to the participant, and finally she/he was debriefed about the aim of the study.

3 Results

3.1 Female characters - more or less feminine

Table 2. Evaluation of the feminine character (F+) vs. the weak feminine character (F-). The mean values correspond to the bipolar Likert type scales presented in Table 1.

	L1	L3	L4	L5	L6	L7	L8	L9	L11	L13	L16	L17	L18	L20
m(F+)	5,75	4,05	5,15	3,45	5,35	5,23	3,28	5,38	5,48	5,63	3,48	5,65	2,38	2,28
m(F-)	5,95	5,23	5,30	4,83	5,05	5,60	3,65	5,28	6,05	5,73	3,90	5,78	2,88	3,30
p	0,46	0,00	0,60	0,00	0,41	0,30	0,29	0,76	0,10	0,75	0,25	0,66	0,13	0,00
		a		b					d				e	c

If we instead look at the traits with the *smallest* differences in evaluations between the two characters, these are: decisiveness, experience, persuasiveness, expertise and knowledge – which all belong to the heading of expertise/trustworthiness[3] (Table 1). Intelligence, furthermore, is one of the *two* traits that does not follow gender stereotype predictions. The more feminine-looking woman is evaluated as more intelligent than the more masculine-looking one.

[3] This, furthermore, is the only of the five larger categories with clearly non-significant differences between evaluations of the two female characters.

3.2 Male characters - more or less masculine

Table 3. Evaluation of the masculine character (M+) vs. the weak masculine character (M-). The mean values correspond to the bipolar Likert type scales presented in Table 1.

	L1	L3	L4	L5	L6	L7	L8	L9	L11	L13	L16	L17	L18	L20
m(M+)	6,00	4,60	5,05	4,00	5,35	5,20	3,75	5,10	5,40	5,45	3,73	5,28	2,75	2,70
m(M-)	5,90	4,75	4,73	4,25	5,10	5,30	3,48	5,20	5,15	5,03	3,80	5,18	2,65	2,55
p	0,72	0,71	0,32	0,49	0,48	0,79	0,42	0,78	0,45	0,23	0,84	0,79	0,76	0,61

In the evaluations of the two male characters we find much less divergence. They are very evenly evaluated, with no significant differences. If we, nevertheless, look at the four *somewhat* larger differences, they are all consistent with gender stereotypes. The more masculine-looking male is evaluated as more decisive, more expert, more dominant and more objective (less subjective) than the less masculine-looking male.

3.3 And how about male vs. female?

Table 4. Evaluation of the female characters (F+ & F-) vs. the male characters (M+ & M-). The mean values corresponds to the bipolar Likert type scales presented in Table 1.

	L1	L3	L4	L5	L6	L7	L8	L9	L11	L13	L16	L17	L18	L20
m(F)	5,85	4,64	5,23	4,14	5,20	5,41	3,46	5,33	5,76	5,68	3,69	5,71	2,63	2,79
m(M)	5,95	4,68	4,89	4,13	5,23	5,25	3,61	5,15	5,28	5,24	3,76	5,23	2,70	2,63
p	0,61	0,89	0,12	0,96	0,92	0,53	0,54	0,46	0,04	0,06	0,77	0,04	0,75	0,45
			d						a	c		b		

Our focus was on *masculinity and femininity within* the categories of males and females, respectively. However, since we had the data, we also wanted to analyze the evaluations of *male* versus *female* characters. As described in section 1.2., studies have shown that when a female and a male voice are exchanged for one another in a digital system, patterns of evaluations tend to correspond to gender stereotype patterns known from social psychology. For example in the study [16] where participants found praise from a digital tutor with a male voice more credible and valuable.

In our study, as well, we had a female versus a male voice, but additionally also visual representations of male versus female characters. Given this, we expected similar results as in [16] – i.e. results corresponding to gender stereotypes. However, quite to the contrary to our expectation, evaluations of male and female characters were very *uniform* (see table 4). They were perceived as *equally* warm, intelligent, emphatic, friendly, knowledgeable, personal and pleasant. The four significant or close to significant differences found – in expertise, objectiveness, decisiveness and persuasiveness – were all *contrary* to gender stereotypes, in the sense that the female characters were held to be more objective, decisive, experts and persuasive.

4 Discussion

4.1 More and less feminine-looking female characters

The result indicates that manipulation of visual appearance of female virtual characters as to how feminine-looking they are may influence users' evaluation of the characters in accordance with gender stereotypes. Furthermore, this can, apparently, be achieved through *relatively subtle* visual cues.

The latter statement may seem depressing. Are humans really that sensitive to stereotype mechanisms? Do they really associate a *slightly* more feminine-looking female synthetic character with higher degrees of warmth, subjectivity and empathy than a *slightly* less feminine-looking one? It does seem so. One should remember, though, that stereotypes often operate non-consciously in evaluations and expectations [6, 10, 18]. When *asked,* many people deny that one or another gender stereotype cues could make any difference to them or that they would even notice them.

In the actual study, there was, however, also one cluster of traits where the evaluations of the female characters were not in line with gender stereotypes, namely those concerning the characters' decisiveness, experience, persuasiveness, expertise and knowledge. The more and the less feminine-looking characters scored more or less even on these traits. In fact, *all four* characters were very similarly, and highly, rated on these traits. A possible explanation, we suggest, is that *gender* stereotypes are here overridden by a *medical doctor stereotype*[4]. In the Swedish society this is a high status profession, with its practicians ascribed expertise, knowledge and intelligence.

[4] Such an interaction between role-of-character effects and visual-appearance-of-character effects would fall in line with other studies [9] indicating interdependence of (i) user responses to visual appearance and (ii) the roles of synthetic characters.

4.2 More and less masculine-looking male characters

The evaluations of the two male characters followed one another quite closely. In other words, the two characters did not elicit differences in accordance with gender stereotype patterns. The reason for this can only be speculated about at this stage. Perhaps more explicit visual cues in the degree of masculinity are required in order to elicit gender stereotypes for male characters. Perhaps there are inadequacies in the visual designs of the two male characters that level away evaluative differences (*e.g.* too large differences in the color schemes, potentially influencing evaluations of warmth and empathy). Note that such possible explanations are compatible with a pre-validation of the characters with M+ being categorized as more masculine than M-, since conscious evaluations as answers to explicit questions about gender may not necessarily match unconscious interpretations and evaluative effects.

4.3 Male characters versus female characters

The results of the comparison between male and female characters at first seemed puzzling. Here we had *two unmistaken females*, as to their voices as well as looks (even if one slightly more and one slightly less feminine-looking). Similarly, we had *two unmistaken males* as to their voices and looks (even if one slightly more and one slightly less masculine-looking). Yet, the differences in how the male versus the female characters were evaluated did not at all correspond to gender stereotypes and evaluative patterns known from social psychology. Instead, the evaluations were very *uniform* – with the exception of four (statistically or nearly statistically significant, yet not very large) differences. These, though, all went in the *opposite direction* than the direction predicted by gender stereotypes.

In order to seek an explanation of this, in our view, puzzling result, we decided to repeat the study with a number of (new) participants, but this time using only the voices and no visually appearing synthetic characters[5]. Eighteen adults took part in this follow-up-study. Again, the male and the female (now only mediated via voices) were very uniformly evaluated, and for the few differences, these were *contrary* to gender stereotypes. In other words, again we had a result quite contrary to social psychology studies and to studies reported in [16]. At this stage one of the article authors suggested that the puzzling results might be due to the somewhat differing *dialects* in the male and the female voices – the woman speak-

[5] Regarding the role of the virtual characters as medical doctors, participants were, just as in the original study, *told* that they would listen to medical doctor Elm and medical doctor Ask.

ing something like standard Swedish, and the man speaking a regionally colored dialect from southern Sweden (yet quite refined or mild).

According to dialect research [20], standard Swedish is associated with high status and authority. It is held to be neutral and clear, and is often used by politicians, company directors, actors and newscasters in national media. It is also perceived as credible and trustworthy, yet somewhat machinelike and cold. This, then, was the dialect of the female character(s) – in the original as well as the follow-up study. A participant in the follow-up-study indeed commented that the female voice sounded "monotonous but informative". Another participant said, that "she sounds like a newscaster or a spokesman for something".

Turning to regional dialects, someone who speaks a regional dialect is often perceived as kinder, duller, slower and more naïve than someone who speaks standard Swedish [20]. Some participants in the follow-up-study described the male voice, in contrast to the female voice, as "less professional", "milder" and "more easy-going".

In sum, in our experimental setup a *woman's* voice with high-status dialect, associated with professionalism, neutrality, facts, distance, credibility etc. (i.e. 'male' traits), had been compared to a *man's voice* with a low-status dialect, associated with naïveté, niceness, non-professionalism, non-intelligence, subjectivity and commonplaceness (i.e. 'female' traits). In light of this, the 'out-leveled' results in the original as well as in the follow-up study – in the sense that males and females were valued very similarly on a number of gender stereotypical traits – may be due to the fact that dialect and gender stereotypes in this case counteracted one another.

5 Conclusion

In an empirical study like this, one often strives to focus on and isolate specific parameters, while controlling the rest. In this case the focus has been on *femininity and masculinity in the visual static appearance* of synthetic characters. The results support the proposal that this is, indeed, a design aspect of these characters that plays a role for users' perceptions and evaluations of them. More importantly, influences seem to appear also when no obvious or extreme visual categories or stereotypes are involved, indicating that, perhaps, the static visual representation of synthetic characters is *never* socio-affectively neutral.

However, for any complex digital product, a number of factors are there to (potentially) explain users' affective responses (cf. [13, 17]). As already pointed out, a synthetic character has, of course, more to it than its visual gender. There is

an inevitable complexity in which any given design parameter will interact with others. In the present study, for instance, *visually mediated gender aspects* seemed to interact in particular with dialects and roles of characters[6], in an unforeseen way that interfered with the set-up where we had aimed at controlling critical parameters.

A possible lesson to draw is that user studies, such as this one, from the outset should be framed as *iterative studies* to be repeated in different versions. *Which* aspects – in settings, tasks and various character traits – that should be modified and manipulated throughout the series of studies is however not to be settled at the start, but instead be allowed to emerge during the process. Within a more pragmatic design tradition such an explorative and iterative approach is self-evident. We hold that it can also be useful within an academic context [cf. 13]. Furthermore, we think that results from collected sets of user studies can be used as design heuristics – not that designers of synthetic characters can thereby refrain from (iterative) user evaluations, but as a means to start the navigation in the immense design space at the outset, and influence the design of the initial materials supplied for the first user evaluations.

With respect to the issue of *gender perception*, specifically, we believe it is useful to realize that a character is here being loaded by several parameters. Furthermore, we think that insight into how dialect stereotypes, gender stereotypes and role stereotypes can be combined as to strengthen, weaken or override each other, can be useful for designers of synthetic characters. In particular while these are artifacts with a seemingly enormous potential to elicit affective responses in users.

Acknowledgements

Our thanks to Dan Holmér and to ACII reviewers for valuable comments.

References

[1] Adelsvärd, V.: *Kvinnospråk och fruntimmersprat – Forskning och fördomar under 100 år*. Brombergs, Stockholm (1999)

[2] Argamon S., Koppel M., Fine J., Shimoni A.: Gender, Genre, and Writing Style in Formal Written Texts. *Text*, 23, 3 (2003) 321–346

[3] Baylor, A. Rosenberg-Kima, R., Plant, E.: Interface Agents as Social Models: The Impact of Appearance on Females' Attitude Toward Engineering. *CHI 2006*, (2006) 526–531

[4] Baylor, A.: Preliminary Design Guidelines for Pedagogical Agent Interface Image. *IUI 2005*, (2005) 249–250

[6] Most certainly there is also interaction with user characteristics.

[5] Bem S.: *Lenses of Gender: Transforming the Debate on Sexual Inequality*. Yale University Press, New Haven, CT (1993)
[6] Cook M.: *Perceiving Others – The Psychology of Interpersonal Perception*. Methuen & Co Ltd, London/New York (1979)
[7] Gulz, A., Haake, M.: Design of Animated Pedagogical Agents – A Look at Their Look. *Int. J. of Human-Computer Studies*, 64, 6 (2006a) 322–339
[8] Gulz, A., Haake, M.: Visual Design of Virtual Pedagogical Agents: Naturalism versus Stylization in Static Appearance. *3rd Int. Design and Engagability Conference @ NordiChi 2006*, (2006b)
[9] Haake, M., Gulz, A.: *A Look at the Roles of Look & Roles in Virtual Pedagogical Agents*. (Submitted)
[10] Haake, M., Gulz, A.: *Visual Stereotypes and Virtual Pedagogical Agents*. (Submitted)
[11] Hall, L., Woods, S.: Empathic Interaction with Synthetic Characters: The Importance of Similarity. In: C. Ghaoui (ed.): *Encyclopaedia of HCI*, IDEA Group. (2005)
[12] Holmér, D.: *Stereotyper i den virtuella världen: Hur visuella intryck påverkar deltagares uppfattningar av virtuella agenter*. Master Thesis, Lund university, Lund (2006)
[13] Höök, K.: User-centered Design and Evaluation of Affective Interfaces. In: Ruttkay, Z., Pelachaud, C. (eds.): *From Brows to Trust*. (2004) 127-160
[14] Isbister, K.: *Better Game Characters by Design*. Elsevier (2006)
[15] Koppel M., Argamon, S., Shimoni, A.: Automatically Categorizing Written Texts by Author Gender. *Literary and Linguistic Computing*, 17, 4 (2003) 401–412
[16] Reeves, B., Nass, C.: *The Media Equation*. Cambridge University Press, New York, NY (1996)
[17] Ruttkay, Z., Doorman, C., Noot, H.: Embodied Conversational Agents on a Common Ground. In: Ruttkay, Z. Pelachaud, C. (eds.): *From Brows to Trust*. (2004) 27-66
[18] Schneider, D.: *The Psychology of Stereotyping*. Guildord Press, New York, NY (2003)
[19] *SitePal*: http://www.oddcast.com/home/sitepal
[20] Skogsberg, Y.: *Styr fördomar om andra människor vår uppfattningar om dialekter?* Dept. of Lingustics, Stockholm University, Stockholm (2004)
[21] Snaprud, P.: Få klarar att vända på dygnet. *Forskning & Framsteg*, 7 (2005) 37–39
[22] Van Vugt, H., Hoorn, J., Konijn, E., de Bie Dimitriadou, A.: Affective Affordances: Improving Interface Character Engagement through Interaction. *Int. J. of Human-Computer Studies*, 64, 9 (2006) 874-888
[23] Voelker. D.: *The Effects of Image Size and Voice Volume on the Evaluation of Represented Faces*. PhD Thesis, Stanford University, Palo Alto, CA (1994)
[24] Waern, A., Larsson, A., Néren, C.: Hypersexual Avatars: Who Wants Them? *Advances in Computer Entertainment Conference*, Valencia, Spain. (2005)
[25] *Demo*: http://wwwold.eat.lth.se/Personal/Magnus/default_eng.htm

Paper V

Challenging Gender Stereotypes using Virtual Pedagogical Characters

Gulz, A. & Haake, M. (in press)
S. Booth, S. Goodman, G. Kirkup (Eds.)
Gender Issues in Learning and Working with Information Technology

Challenging Gender Stereotypes using Virtual Pedagogical Characters

Agneta Gulz[a], Magnus Haake[b]

[a] *Division of Cognitive Science, Lund University, Kungshuset, Lundagård, 222 22 Lund, Sweden*

[b] *Department of Design Sciences, Lund University, P.O.Box 118, 221 00 Lund, Sweden*

Manuscript to be published in:
S. Booth, S. Goodman, G. Kirkup (Eds.)
Gender Issues in Learning and Working with Information Technology

Abstract: This paper explores motivational and cognitive effects of more neutral or androgynous-looking versus more feminine-looking and masculine-looking virtual characters. A user study is presented, in which 158 students, aged 17-19, encountered four virtual characters that were visually manipulated to represent gender stereotypicality versus androgyny. On the one hand we explored students' attitudes towards the different characters as seen in how they ranked them as preferred presenters and articulated their arguments for doing so. On the other hand we looked for patterns as to which character(s) influenced female and male students most positively with respect to their attitude towards a university level computer engineering programme. Results from the study are presented and discussed. We conclude with pointing towards future research and potentials within the area.

Keywords. Virtual characters, gender, visual design, femininity, masculinity, androgynities, computer engineering, educational choice.

1. Introduction

A long-standing issue in higher education in engineering and other technical fields has been that of recruitment and retention of female students. The arguments for this are many, and here we will restrict our interest to recruitment with the support of virtual coaches. Baylor and collaborators demonstrated (Baylor

and Plant, 2005; Baylor, Rosenberg-Kima and Plant, 2006) that the use of virtual coaches portrayed as young and attractive females can increase the willingness of female students to chose technically oriented courses and help increase their believes in their own ability in technical domains. Processes of role modelling and identification (cf. Bandura 1977, 2000) seem to be involved. The female students could more easily match these coaches with their personal identity compared to a virtual coach portrayed as a 'typical male engineer' (see Figure 1).

Figure 1. Example of two alternative engineering coaches (young, attractive female versus 'typical' male engineer) from Baylor et al. (2006).

When the results of Baylor and collaborators are analysed in detail it appears, however, that the increase in their belief in their own abilities partly stems from a conception of a 'female, feminine, young and attractive' engineer as *less competent* than a 'real, typical male engineer'. The prejudice that females, and in particular females with a strongly feminine appearance, are less competent in technical domains seems to spill over to the virtual area, generating increased self-efficacy of the kind 'If she is able to do it, I can do it!'

Now, this implies a potential conflict between a *short-term* pedagogical goal of recruitment and boosted self-efficacy in female students, and a *long-term* pedagogical goal of changing rather than reproducing gender prejudices and stereotypes. Attempting to avoid this conflict, the present study explores motivational and cognitive effects of more neutral- or androgynous-looking virtual characters versus more typically feminine-looking and masculine-looking ones – in a recruitment context. A multimedia presentation was developed, featuring four different virtual presenters of a university programme in computer engineering. The characters (presenters) were *visually* manipulated – and pre-validated during the design process – to represent a young feminine woman; a more androgynous young woman; a more androgynous young man and a young masculine man. (The characters are depicted and described in detail later on in the article.)

Figure 2. Screenshot from the multimedia presentation with the 'more androgynous young woman' presenting the programme in computer engineering at Lund University.

Participants encountered one of the four characters in the role as presenter (see Figure 2) and were afterwards asked whether and how the presentation had affected their attitude towards the computer engineering programme as well as what they thought of the presenter. Finally they were presented with all four characters and asked to rank them in terms of which one they themselves would prefer as the presenter of the computer engineering programme, and to motivate their ranking.

1.1. Issues in focus

We wanted to:

1a) Explore students' attitudes towards the different characters, as seen in how they *ordered* them as preferred presenters of the computer engineering programme: Would the more neutral, androgynous characters be preferred to the more gender typical characters, or vice versa? Would the rankings of female students differ from those of male students?

1b) Explore how students *articulated* their attitudes towards the four characters: What reasons would they give for their first and last choice?

Additionally, we wanted to:

2) Explore which character(s) *influenced* female and male students most positively with respect to their attitude towards the computer engineering programme.

It should be emphasized that it was not given beforehand whether there would be a concordance between the character(s) that the students *explicitly chose and argued that they preferred* most as presenters, and the character(s) that had the most *positive influence on their attitudes* to the educational programme presented. Many studies have showed that perceptual gender related stimuli can have a considerable impact on peoples' non-conscious cognitive processes that is not necessarily in concordance with what the same people are aware of and consciously report (e.g. Reeves and Nass, 1996; Brave and Nass, 2005).

Therefore the study was designed to collect *both* conscious, articulated responses in the format of preference rankings and arguments, *and* responses that reflect less conscious influences and processes.

2. Study[1]

2.1. Participants

Eighty-six female and 72 male 17-19 year old students at four different high schools in two cities in southern Sweden participated.

2.2. The virtual characters

2.2.1. Visual appearance

The design aspect manipulated in the four virtual characters (Figure 3) was their *visual appearance*. This was done by one of the team members (educated in visual arts) according to *gender schemes* used in design practice.

The two more neutral or androgynous-looking characters, *FA* (female androgynous) and *MA* (male androgynous), were developed out of an identical bust, differing in:

- the length of their hair
- the eye brows (with FA having more regular and slightly plucked eye brows)
- the clothing (somewhat neutral as to fashion but gender specific)
- FA having short eye lashes accentuating the eyes
- MA having a somewhat darker colour scheme, producing slightly bigger and more pronounced shapes

A more *feminine-looking* female character, *FF*, was developed from the bust of MA and FA by using feminine attributes such as rounded head shapes, bigger

[1] For a more detailed account of the study, see Gulz, Haake and Tärning, 2007.

eyes and smaller nose, make-up, long, dark eye lashes, the mouth modelled with fuller lips and the cheeks lifted and slightly more pronounced.

Figure 3. The four virtual presenters used in the study.

A more *masculine-looking* male character, *MM*, was designed by using masculine attributes such as broader, angular and more pronounced head shapes (especially the shape of the jaw), broader shoulders, a distinct Adam's apple and pronounced eyebrows.

2.2.2. Gender stereotypical and androgynous visual appearances

The focus on both conscious and articulated and more non-conscious responses to the characters, with respect to visually represented gender, set some constraints on the visual designs.

Over-explicit visual gender stereotypes had to be avoided. A feminine female character who looks like Pamela Anderson, and a masculine male character who looks like Arnold Schwarzenegger (compare Figure 4), would probably initiate *conscious* reflections on gender and gender stereotypes, the purpose of the study, 'politically correct' answers, and so on, and dominate or rule out more immediate and non-reflected responses. Thus, FF and MM were not designed as *pronouncedly gender stereotypical* in their looks, and in particular, we sought to avoid the bimbo stereotype for FF.

FA and MA, in turn, are not *pronouncedly androgynous* in their looks, even though only a few visual parameters differ between them. In an early design phase, they were actually more similar to one another, and in particular FA appeared as more clearly androgynous. This, however, evoked negative responses in the validation process. In a pre-test a number of participants declared that they at

first were uncertain whether the FA character was female or male, and that they found her – they all finally decided it was a girl – rather unattractive.

Figure 4. Examples of visual gender stereotypes in digital media. Left: two characters from the console game *Ninja Gading Sigma* (Ninja Gaiden ® Sigma, © 2007 Tecmo, LTD) that reflect parts of the computer game domain's action/fantasy-genre. Right: two characters designed by members in the on-line-world *Second Life* (Second Life™ Showcase, © Linden Research, Inc.). The characters can be found in a collection of examples (*showcase*) directly in the main menu of the home page (http://secondlife.com/showcase/).

This was undesired given the aim of the study. We wanted to explore the potential in using more androgynous in comparison to more strongly gender differentiated, visual characters. But such a comparison presupposes that all characters are comparable in the sense that none is perceived as particularly unattractive, irritating, strange or unusual, since this may otherwise take away the focus from the central variables.

However, the need to avoid strange or unusual characters is also a dilemma. Not being able to decide whether someone is a man or a woman is known to induce insecurity and unease in many people (Brave and Nass, 2005), and this constraints the possibility to explore more pronounced androgynous-looking characters. In order to progress here, our next step will be to use less naturalistic-looking, more cartoonish characters. More of this in the section on future research.

2.2.3. Character features other than the visual appearance

Since we wanted to explore possible effects of visual appearances, we strived to keep all other character variables constant, or at least comparable, between the four characters. The slide show behind the characters was identical, as well as the information communicated by the characters. The characters' facial expressions followed the same animation scheme. As to voice, one and the same recording – originally a woman's voice – was digitalized into a female and a male voice.

In this way features such as dialect and tone were better controlled for. It was also important to choose a female voice that would work with *both* the female characters, and a male voice that would work with *both* the male characters since mismatches between look and voice, like other inconsistencies in virtual agents, are known to disturb or irritate people (Nass, Isbister and Lee, 2000).

2.3. Procedure

The computer program was run on four laptops. The 158 participants all used headphones to ensure that they were able to hear well and concentrate.

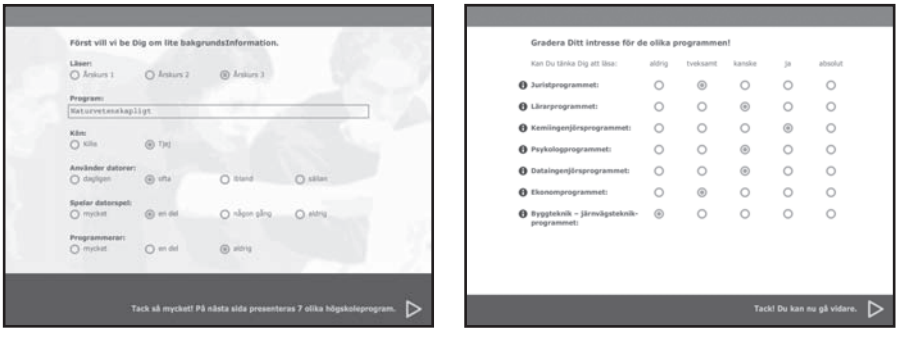


Figure 5a & 5b. Screenshots from the multimedia presentation. 5a (left): Demographic data form; 5b (right): Evaluation of the seven university programmes.

After filling in demographic data on the screen (see Figure 5a), the participants read brief texts about seven university level educational programmes and were asked to what extent they could imagine themselves as students on these different programmes. For each programme they were asked to check one of the following alternatives: 'never', 'unlikely', 'perhaps', 'yes' or 'absolutely' (see Figure 5b).

Thereafter they were told, in the digital environment, that a new presentational media was being developed, which they were invited to help evaluate. It would concern the computer engineering educational programme (from their perspective seemingly randomly chosen among the programmes they had just evaluated).

At this instance the virtual presenter appeared (see Figure 5c) and spoke about the educational programme during 2 minutes, with an accompanying slide show presentation. Parts of the presentation had been pre-validated as to its content and style by other students in the same age group in a previous and related study (Altmejd and Vallinder, 2007).

Figure 5c. Screenshot from the slide show with the virtual character presenting the university programme in computer engineering.

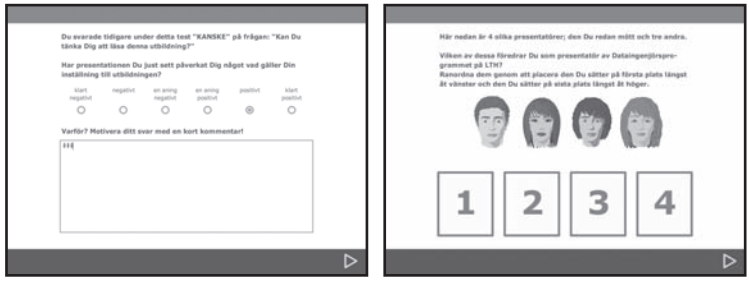


Figure 5d & 5e. Screenshots from the multimedia presentation.

When the presentation was finished and the presenter had thanked the listener, a number of questions followed. All were presented on-screen and filled out on the computer. First, the students were asked to evaluate on a 6-step Likert scale whether the presentation had influenced them in their attitude towards the computer engineering programme: in a positive or negative direction and to what extent. Thereafter they were asked *why* they had been influenced in this way (see Figure 5d).

The next question regarded their view of the presenter they had encountered. Thereafter all four alternative presenters were shown, and the participants were asked to rank them from 1 to 4 in order of preference: Who would they prefer

as presenter of the computer engineering programme? (Figure 5e). Following this, the virtual character they had ranked as number 1 appeared and they were asked why this was their first choice. Thereafter the character they had placed as number 4 appeared and they were asked why this was their last choice. Finally, the participants were thanked on screen for their participation and asked to speak to one of the experiment leaders for debriefing and to receive a lottery ticket for cinema vouchers.

Ten groups of participants, in total about a fifth of the participants, also took part in focus group interviews after they had completed the session just described. These interviews centred around masculinity and femininity in appearance, behaviour, style and in occupations; as well as on the topic of androgyny.

3. Results and analysis

3.1. Character choices

3.1.1. First and fourth places in rank

The female participants most frequently chose MA and FA as favourite presenter with 29 and 28 choices respectively, versus 21 for FF and as few as 8 for MM. As many as 32 female participants put MM last and 24 put FF last. Only 16 and 14 put FA and MA, respectively, in fourth place. In other words, the two androgynous characters were clearly preferred. (See Figure 6).

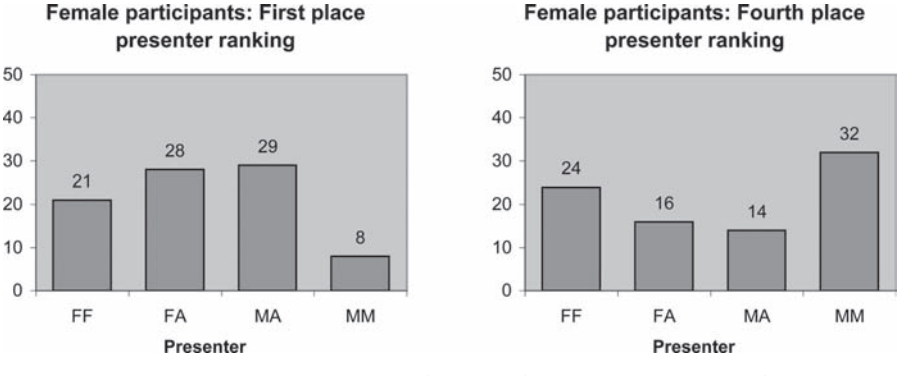


Figure 6. Female participants' first and fourth place ranking of the virtual presenters.

The first choice of the male participants was much more even (see Figure 7). As to the fourth place, the pattern was, on the other hand, very pronounced with

28 votes for FF, 25 for MM, and only 9 and 10 for FA and MA. In other words, also the male participants showed a preference for the androgynous characters, although primarily seen through their fourth place ranking (see Figure 7).

Two other choice patterns are particularly interesting from a gender point of view: (i) the extent to which participants ranked the *two female characters* or the *two male characters* in the two first places, and (ii) the extent to which participants put *the two more androgynous characters* in the two first places.

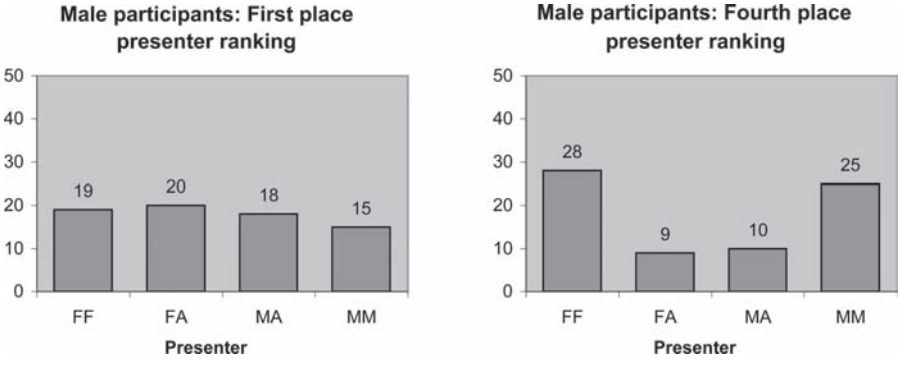


Figure 7. Male participants' first and fourth place ranking of the virtual presenters.

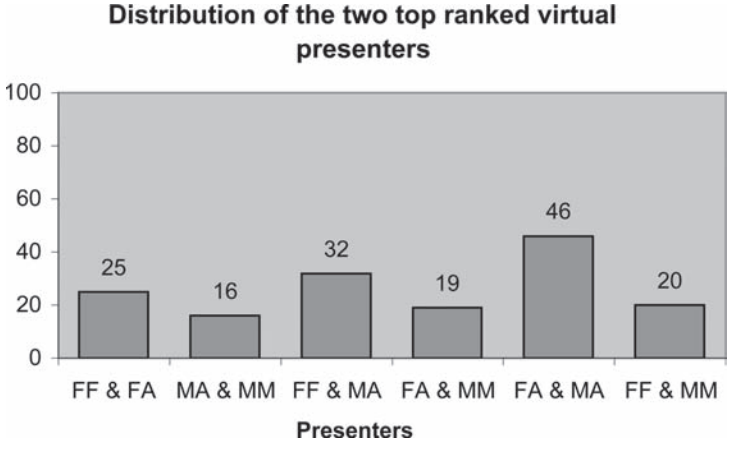


Figure 8. Distribution of the different combinations of the two top ranked (first and second) virtual presenters (χ^2[total distr.] = 23,595; p = 0.000).

3.1.2. Same-gender characters in the first two places

Thirteen female and 12 male students, corresponding to random distribution, ranked the two female characters in first and second place (see Figure 8, leftmost bar). Ranking the two male characters in first and second place occurred much less often. Ten male and only 6 female students clearly correspond to less than random distribution (see Figure 8, second leftmost bar). In other words, even though computer engineering is an education and professional field with strong male dominance, students and in particular female students, tended not to place two male characters in first and second place.

3.1.3. The two androgynous characters in the first two places

As many as 28 female and 18 male students, 46 in total (see Figure 8, second rightmost staple), put *the two more androgynous characters* as their two first choices of presenter. This is significantly above what is expected by chance. Placing the two more gender stereotypical characters in first and second place, on the other hand, was a little less common than what is expected by chance.

3.2. Arguments and reasons for character preferences

We now leave the ranking preference data and move to the participants' arguments and reasons for their character rankings. The issues focused on in analyzing these are: (i) participants' referring to the *gender of the characters* in arguing for a placement of the character, (ii) comments and arguments about the *attractiveness of characters*, and (iii) *the nerd* as appearing in arguments and comments. These issues are presented below in due order.

3.3. Referring to the gender of characters in arguing for ones preferences

On 35 occasions the gender of characters was brought up in motivating a first of last choice of presenter: by female participants in 22 cases and by male participants in 13. This difference in number may reflect a more pronounced gender consciousness in women compared to men (e.g. Hirdman, 2003). Below we look in more detail at the content of participants' comments when referring to characters' gender. This material gives indications both of ways of reasoning, and of how the different characters afford or mediate – or perhaps even trigger – different types of gender related arguments.

3.3.1. MM: Gender related arguments

The character that is *least* – only four times – referred to in gender terms in choice/non-choice arguments is MM. For instance, one male and one female student chose him since:

'it is suitable for a boy to talk about technical things'

'this is a guy, and I think that this [the computer engineering programme] is for guys'

3.3.2. MA: Gender related arguments

Nine participants raised gender related arguments concerning MA. Arguments for choosing MA as presenter were, for example:

'it is a guy, and it is a computer education'

'he looks like a *computer-guy*'

'he looked nice; and in many contexts, many people find it more reliable when it is a man speaking'

An example of an argument for the ranking of MA in the *fourth place* is:

'since he is a guy, and "the typical kind of guy" for this kind of education' [In this female student's further reasoning it was clear that she thought it would be good to break with the 'usual' associations.]

One female student argued for ranking MA first in a way that might be interpreted as a wish to avoid gender typicality:

'Because he did not have typical short "boy's hair"'

3.3.3. FA: Gender related arguments

Turning to the female characters, eight female and two male students came up with gender related arguments concerning FA. All argued *for* their choosing FA in terms of her *being female*. They said for example:

'I think it is good that girls are more visible'

'I think it is important that it is a women speaking since that can make more girls realize this can be for them'

'she seems young and looks nice, and I think she would make more girls interested'

'I like to see that also girls can be profiles for an engineering education, in particular one involving computers, which has many male students'

'because she was good – and a girl'

'I think women too ought to have influence in speaking for such educations, so that girls can see that there are also female students here'

3.3.4. FF: Gender related arguments

FA was, thus, almost entirely described positively in gender related terms: she is chosen/preferred as a girl/woman. The gender related argu-ments about FF, by eight female and five male students, were in contrast more split and ambiguous.

Three females argued in positive terms about their choice of FF:

'when one thinks about computer educations one thinks, at least I do, mostly of males. To hear a woman present is really a good thing'

'because she was the best-looking and seems more conscious of what women want'

'she looks like a focused woman, who knows what she wants =) ha-ha'

The other five females as well as the five males put FF in last place. Sev-eral of them seemed to defend this by saying that they do want women presenters in this context, but not this woman; not FF. The female stu-dents commented:

'it could perhaps be good to have a female presenter, but perhaps not her' [She placed FA as number 2.]

'I cannot say straight on that she would not be good as presenter but she just does not seem competent' [She placed FA first, with the motivation that she finds it important to have a woman as presenter.]

'she is a kind of woman I don't like' [She ranked FA first.]

'I am not against her really but I think it should be mixed between women and men when it comes to influencing' [She placed FA first.]

'males are usually good at presenting these things' [Nevertheless she ranked FA first, saying that FA was good.]

The males reasoned:

'women are, on the whole, less interested in computers, and this one looked less motivated than the other woman'

'she looked the least like someone dealing with technology'

'she was a woman on a computer education and she did not seem to be-long there'

'she doesn't give the feeling of being as serious as the man, doesn't seem to have the same working experience as the man'

'as I said, a woman feels more welcoming than a man, but she looked so styled, which I don't like'

3.3.5. Summing up

In summary there is a considerable difference between *ten* arguments (five by girls and five by boys) *against FF* as presenter in terms of her being a woman – or as being 'this kind of woman' – and *no* argument *against FA* as presenter in terms of her being a woman. On the other hand, *ten* participants (eight female and two male students) argue *for FA* as presenter since she is a woman, but only *three* (three females) argue *for FF* as presenter since she is a woman.

3.4. Arguments involving the attractiveness of characters

In general, it appears to be crucial for young peoples' educational choices that they can find adequate role models (Kessels, 2005; Rommes et al., 2007).

One of the parameters known to influence the strength of a role model is *attractiveness*, If a role model is perceived as attractive the behaviour of the model is more often imitated (Rommes et al., 2007). Therefore, we wanted to analyse our material to see to what extent attractiveness and non-attractiveness was brought up in evaluating the characters as presenters of the computer education.

Indeed, comments on attractiveness were quite frequent in arguments for choosing a character as first place presenter, while comments on non-attractiveness were quite frequently involved in argument for placing a character in fourth place.

There were no large differences in the number of comments on attractiveness made by male students (52) and female students (67). The number of comments was relatively evenly distributed among the four characters with the exception of FF who received one third of the attractiveness/non-attractiveness comments. Notably, there was quite considerable divergence in students' opinions regarding *all four* characters in terms of their attractiveness/non-attractiveness. Firstly, this is positive in view of the study design. If one character was found particularly attractive or non-attractive this might have interfered with the factor we intended to study, namely the influence of gender stereotypicality in visual appearance. Secondly, the divergence in opinions reinforces the notion that taste differs.

Another appearance aspect relatively frequently commented upon was that of looking 'plain' or 'common'. Also here *all* four characters were commented on as looking (most) 'normal'/'plain'. That the perception and evaluation of who looks 'normal' and 'common' differ among participants is interesting since such perceptions as well can play a role in an identification process.

3.5. *The Nerd* in arguments and comments

The topic of attractiveness/non-attractiveness leads us to the topic of *the nerd*. An elaborate analysis of this issue, based on a rich empirical material from the Neth-

erlands, is provided by Rommes et al. (2007). Using several methods, including focus group analyses and pictures drawn by young students, they pinpoint the Dutch cultural image of a computer scientist: male, unsociable, 'married to his computer', wearing unfashionable clothes and glasses, has a bad haircut, is overall unattractive and basically asexual – that is, *a nerd*.

Figure 9. The world famous nerd Dilbert (© Scott Adams, Inc. / Dist. by UFS, Inc.).

Also organizers of computer engineering programs in Sweden mention a problematic image of the computer engineer student: a male student, constantly in front of his computer, drinking large amounts of coca-cola (Kihl, 2003).

Given that the image or prototype of a person studying at a certain education or belonging to a certain professional group seems important for young peoples' educational and career choices (Hannover and Kessels, 2004) a nerd image associated with computer engineering is a major obstacle for young people to apply to such programmes. The nerd is not somebody to identify with or aspire to be, but someone extremely non-attractive and non-glamorous. Furthermore, relationships and sexuality are important during adolescence, and thus the risk of being associated with the 'asexual' 'nerd' image can be extremely threatening (Lippa, 1991; Rommes et al., 2007). Baylor and collaborators as well (Baylor & Plant, 2005; Baylor, Rosenberg-Kima and Plant, 2006) touch upon these issues, in holding up virtual role models, that are knowledgeable in engineering and simultaneously stand out as attractive and as affirmative in their sexuality.

Rommes et al., (2007) suggest that it is the nerd prototype of the computer scientist, rather than ideas of what is actually involved in studying or working in the field, that makes many young Dutch females – and males – refrain from applying to computer related programmes. Thus, not least given the student recruitment context that our study took place in, we were interested to see whether we would find 'the nerd' in our material when subjects motivated why they chose or did not choose a certain presenter.

3.5.1. The nörd

The Swedish language has the word *nörd*, pronounced very similar to *nerd*, along with the specific word *datornörd* (*computer nerd*). In the logged material (we do not include the focus group interviews here) the word *nörd* was used twelve

times, primarily by female participants. Additionally there were fifteen occasions of arguments, that can be associated with the 'nerd' concept, even though the word *nörd* was not used.

Six participants ranked MM last as a presenter because he was – explicitly or implicitly – a *nörd,* writing e.g.: 'because I think he looked like a computer-nerd (*datornörd*) with his very ugly hair'; 'he looks a little dull, and a little nerdy (*nördig*; 'because he looks like a proper computer-person (*äkta datamänniska*), no one you can recognize yourself in'; 'he looks a bit dull and stiff, a typical computer-guy'. The last participant placed MA first since: 'he is cool, and I got the impression that cool people study at this educational programme'.

On the other hand, a student who had previously said that he was very interested in the computer engineering programme and later on declared that he had become even more positive after having heard the presentation by FF, argued *for* MM as his choice 'because he looks nerdy (*nördig*) ☺'. Another male student wrote that he chose MM: 'because he looks like a genuine computer-guy'. (The word *äkta* (genuine) used here has a positive connotation in Swedish in general.)

Such variation in value ascribed to nörd, recurs when turning to MA. Some participants chose MA and argued about the look of 'computer-people' and of 'the nerd' in positive terms: 'he looks like a handsome computer-guy'; 'he looks good, and looks like a computer-nerd (*datanörd)*'. On the other hand, one female ranked MA as her number one presenter, motivating this by MA *not* seeming like a nerd: 'he looks good and not too nerdy *(nördig*)'.

Summing up, both male characters were at some occasions held to be a *nörd*. However, none of them really fits with the description of the nerd offered by Rommes et al., and both were held to be *good-looking* by at least some participants. MM was slightly more often associated with a nörd than MA and more frequently said to look dull and boring. However, MA received many more comments than MM of the sort that 'he fits in with this education'. Together this does not support the notion that *the nörd* is tightly connected to this education.

In connection with the female characters, *nörd* comments occurred more rarely. Yet one female student put FA in the fourth place 'because she looks dull and nerdy (*nördig*)', and put FF first, 'because she is no nerd (*nörd*)'. Another female student chose FA 'because she does not seem nerdy (nördig)'.

Summing up, what emerges from this material is not an equally strong cultural image as the one that emerged of *the nerd* as the typical computer scientist/engineer (Rommes et al., 2007). Perhaps *nörd* has a partly different meaning and use

than the English *nerd*, with a less tight association to an unattractive appearance or look.[2]

3.6. Attitude influences

We have come to the part of the analysis that is not based on participants' explicit rankings and arguments about the characters, but instead on changes in participants' attitudes towards the computer engineering programme after they had listened to the presentation by one of the four virtual presenters.

Table 1. Attitude influences from the multimedia presentation.

Added positive (+) influences only

	FF	*FA*	*MA*	*MM*
Males	54 +	38 +	21 +	58 +
Females	31 +	36 +	36 +	36 +

Added negative (–) influences only

	FF	*FA*	*MA*	*MM*
Males	8 –	11 –	14 –	0 –
Females	0 –	4 –	15 –	5 –

Added total (positive and negative) influences

	FF	*FA*	*MA*	*MM*
Males	46 +	27 +	7 +	58 +
Females	31 +	32 +	21 +	30 +

Before the presentation they had been asked to what extent they could imagine themselves as a student on some different educational programmes. After the presentation of the computer engineering programme they were asked to mark on a Likert scale the extent to which their attitude towards this programme had been influenced: very negatively, negatively, a little negatively, a little positively, positively, or very positively.

Eighty-six participants answered 'a little positively', and four answered 'a little negatively'. In our analysis we do not include these two middle positions but only 'negatively' and 'very negatively', that were given the values 2- and 3-; and

[2] An anecdotic observation in the context is that the Swedish 'James Bond like' master spy Carl Hamilton (from books and films) has an alias as a computer expert.

'positively' and 'very positively', that were given the values value 2+ and 3+. In table 1 the added values character by character are presented.

Overall more male than female participants reported a clearly positive influence from the presentation. The total sum of positive values for the male participants was 171 (+) and for the females 139 (+). However, the values were overall relatively high. The negative influence, measured in negative values, was considerably smaller, with 33 (-) for males and 24 (-) for females.

When it comes to the presenters involved when participants reported a strong positive change in attitude towards the education, the results were as follows. For female students, the positive influence was evenly distributed over the characters. This is not in line with the results of another study (Baylor, Rosenberg-Kima and Plant, 2006) where female students' attitudes towards engineering classes were considerably more positive if the virtual coach was *female*. What must be borne in mind is however the difference in contexts. Baylor investigated students' encounters *over some time* with a pedagogical coach, directing a tutorial. In our study there was *one brief* encounter with a presenter.

Also the very evenly distributed positive influences from the characters on the female participants, does not concord with the strong preferences in the female participant group for the two androgynous characters indicated through their explicit ranking and reasoning about the characters.

With the male participants the contrast between their explicit ranking and reasoning about characters, and the positive and negative influences from characters, was even more striking. In explicit rankings and reasoning, the androgynous characters were preferred but the positive influence values were low for these characters. Furthermore, the character the male participants ranked lowest and argued most negatively about was MM. But the positive influence on attitudes on male participants was clearly strongest from MM, followed by likewise strong positive influences from FF.

3.7. Analyses in the pipeline

When circumstances allow we would, first of all, want to analyse the material from the focus group interviews where participants discuss androgyny, masculine and feminine professions, etc. Furthermore, we would like to pursue a more focused analysis from the perspective of recruitment of young females to computer engineering programmes. It would also be interesting to pursue culture comparative studies. In what ways do countries, such as the Netherlands and Sweden, differ as to whether computer engineering is an unattractive or attractive

discipline; and can differences, if found, be related to different cultural images of the computer engineer and of 'the nerd'?

4. Tentative conclusions

It has long been acknowledged that there are close symbolic associations between information technology, and masculinities and femininities (Cockburn and Ormod, 1993; Faulkner, 2003). Here, we suggest, virtual agents or characters, with their properties of interactivity and human-likeness, constitute a particular form of information technology endowed with a particular (re)constructive power with respect to gender. The presented study used virtual characters for presenting a university programme on computer engineering, but they could be used to present all kinds of educational domains, and may be of particular interest when aiming for less gendered occupational choices. Studies by Baylor and her collaborators (Baylor & Plant, 2005; Baylor, Rosenberg-Kima and Plant, 2006) provided important background and a point of departure for our study in highlighting the importance of images and alternative cultural role models for engineering students. Also other researchers have suggested that more physically attractive and glamorous female role models might change the negative prototypes of computer scientists (Coltrane and Adams, 1997).

But Baylor and collaborators also pointed out problems of stereotype reproduction in using such characters and images. The aim of our study was to look for an alternative to the explicitly feminine female characters used in Baylor's studies (Baylor and Plant, 2005; Baylor, Rosenberg-Kima and Plant, 2006), and thus a means to reconcile the *short-term* pedagogical goal of recruitment and boosted self-efficacy in female students, and the *long-term* pedagogical goal of changing rather than reproducing gender prejudices and stereotypes.

4.1. Using more than one character?

For the female participants in our study, the following could be seen with respect to the two virtual female characters. In explicit rankings and arguments FA, with her in a relative sense more androgynous look, was clearly preferred before FF. In attitude influences no difference between FA and FF was seen on female participants.

A central result comes from the analysis of arguments that refer to the *gender of the character*. Here we found that a whole group of ten participants argued *against FF* as presenter in terms of her being a woman – or as being 'this kind of woman' – whereas *no* participant did so *against FA*. And vice versa, a consider-

ably larger group of participants argued *for FA* as presenter *as being a woman* than correspondingly *for FF.* Our interpretation is that the more androgynous female character has more positive affordances in gender terms. The FA character was more frequently and more consistently used in positive reasoning and arguments about women in this computer technological context. Female students who already have thoughts about a positive role for women in the computer science domain, or in technical domains in general seem more satisfied with picking the FA-character than the FF-character. The FF-character, on the other hand, seems to mediate or afford, or lend herself more easily, to arguments about women not fitting in this context.

But two things concerning divergence and multiplicity should be pointed out. First, it should not be neglected that there was one group of female participants for whom the FF character appeared to be valuable, as reflected in comments such as: 'she looks as if she also knows what a woman wants', 'she has a chic look', 'she looks like a focused woman who knows what she wants'. Likewise, it should be remembered that opinions on attractiveness/non-attractiveness, as well as on commonness/plainness, diverged as to all four characters. In all, this points towards the possibility of using not just one virtual characters, in this case one presenter, but two or several that take turns and interact with each other.

Second, it is important to situate the results of the present study in a cultural context. Virtual characters, which might function well and be adequate in Sweden, are not necessarily the ones that ought to be chosen in another country. For instance, we observed that *the nörd* seems to have less impact and be more modulated than *the nerd* in some other cultures – which might decrease a need to introduce attractive, sexy female and/or male characters as a counter balance.

4.2. Further potentials

On a broader scale the results from this and other studies indicate a possibility to exploit virtual characters to support identification and formation of identities in young people while avoiding the reproduction of undesired stereotypes. Smartly used this kind of technology could, to borrow from Rommes (2007), be developed into tools that may increase the freedom for young people to create their personal 'gender identity cocktails'.

There certainly exist information technological applications that to the contrary reproduce and even reinforce gender stereotypes, (compare Figure 4). But there is reason to focus also on the strong constructive potential for *changing* and *broadening* cultural images. For one thing, we have quite a different room in which to manoeuvre in virtual worlds than in the real world. As Brave and Nass (2005)

reason: 'Rapidly increasing the number of female teachers in stereotypically male disciplines (or vice versa) seems difficult. But technology provides a wonderful opportunity to […] "staff" educational software to counter stereotypes.'

Furthermore, within computer game communities where members continually contribute to game development, we see a growing diversity in characters – not least new kinds of female heroines. In the wordings of Pinckard (2003) 'in MUDs and MOOs, one can often create a third sex and invent a pronoun and refer to oneself always with that pronoun (and insist others do that same). In these science-fiction and fantasy-themed online worlds, it's perfectly plausible that ungendered, ambiguously gendered, or bi-gendered races could exist'. Examples of gender busting characters can also be found in existing ready-made programs, such as: *Nights into Dreams*, with the magical, androgynous character *Nights*, whose identity can be assumed by both the female and male characters; or *The Legend of Zelda* with gender ambiguity around all main characters; or *Metroid*, with the gender bending action heroine *Samus Aran*.

5. Some broader issues & future research

We regard the presented study as a first step in a larger programme of exploring the pedagogical potential of stereotype-challenging virtual characters, with several different paths to follow and to explore. In this section we discuss some that have high priority on our agenda: 1. Exploration of perceptions and attitudes towards androgyny in virtual characters and the pedagogical potentials herein. 2. Education and gender with respect to educational programmes where men are in the minority; 3. Development of dedicated gender pedagogical digital tools involving virtual environments and characters.

5.1. More imaginative androgynities and their pedagogical potentials

The first study would explore perceptions and attitudes towards androgyny and how androgyny in virtual characters can be used pedagogically, focusing on more stylised and imaginative characters. As related earlier in the text, *visually naturalistic* characters like those from the presented study constrain possibilities of exploiting more pronounced androgyny, since naturalistic androgyny induces insecurity in many people. But with more imaginative characters, the design freedom and potentials appears wider (see figure 10) while there is still evidence that *identification processes* with respect to visually less naturalistic characters can function well (GameGirlAdvance, 2004; Gulz and Haake, 2006a, Gulz and Haake, 2006b; Haake and Gulz, 2007; McCloud, 1993).

Figure 10. Androgynous and visually less naturalistic virtual characters: Unknown artist; Weblog user icon (www.catahya.net); Avatar for Shaun Altman (secondlife.com); Cropped illustration (The Nuremberg Chronicle: Strange People: Androgyn XIIr, Morse Library, Beloit College, © 2003 Beloit College); Illustration by Johnny Scharonne (scharonne.wordpress.com); Painting by artist Klaus Hausmann (www.arsvenida.de); Cropped extract from a poster portraying androgynous young Japanese 'host-boys' (tokyolove.blogspot. com); Illustration by Mireille Schermer for NRC Webpagina (© NRC Handelsblad).

Attractiveness, mentioned above, would here be an issue, since it seems of importance for positive and well-functioning role models and cultural images. We will therefore look for and explore *non-gendered*, *attractive characters*. An implication is a need to avoid androgyny in the sense of in-between: greyish, neither feminine nor masculine, average.

In the context that we are dealing with, androgynous attractiveness could be a gain over gender-stereotypical attractiveness, and in particular over *feminine attractiveness* of the kind that Baylor and collaborators work with. By turning towards androgynous attractiveness, one may get away from associations of attractiveness and women and femininity leading to conceptions like 'a womans primary role or function is not competence but to be good-looking and to attract'.

Our approach to androgyny is overall optimistic. We consider it a large space with many possibilities of combinations of characteristics – those classified as feminine and those classified as masculine. Other researchers who express an optimistic view on the boundary widening potential in digital world with respect to gender are Haraway (1991), Turkle (1995), Gilmore (2004) and Chess (2006). They all argue, in various ways, that androgyny in the digital world is a rich continuum with many possibilities for femininity, masculinity, both or neither.

5.2. Educations with male students in minority

The second study would focus at a domain with under-representation of male students, in contrast to the presented study which has focused on images of computer engineering, a domain where women are underrepresented.

The underlying issue regards the possibility to influence such situations by offering examples that differ from culturally dominant prototypes of engineers, nurses, etc. We wish to explore the potential in virtual worlds to offer a broader range of styles and identities than in the real world. Can this be a place where cultural images are constructed, with potentials to countering other cultural representations of gendered technology? A number of researchers discuss or propose this (e.g. Reeves and Nass, 1996; Rommes, 2007), and we find it highly interesting. Not the least given the dynamics and interactivity of these novel digital media, which might make them even more powerful in mediating cultural norms than traditional media.

5.3. Dedicated gender pedagogical tools

Going one step further, there is a potential for a third study, constructing dedicated digital tools based on virtual characters to support gender perspectives in teaching/education. Already we see tools used for experimentation with simulations and virtual role-games, and toolboxes for designing virtual characters, such as *SitePal*, *PeoplePutty* and *Meez*.[3]

The use of such tools imply that a number of design decisions must be made: 'What is this judge, this police officer and this arrested person going to look like and speak like? Which gender, age, ethnicity, clothing, voice, dialect shall we assign to them? Educationalists interested in discussing and challenging prejudices and proposing alternatives now have an opportunity to use such a situation as a basis for reflection and discussion: 'Why did we choose this character in this role? Whose appearance is exposed and whose voice is heard, in terms of gender, age, ethnicity, class, regional subgroup, etc? What features did we combine with one another and which not?'

Notably, *various* alternatives may be suggested and *tried out*. Such an active and dynamic situation can provide a natural basis for reflection and discussion – in contrast to a more disconnected classroom discussion on stereotypes.

This is perhaps particularly important in the academic context, where language and other abstractions often have a strong position compared to a more tangible visual tradition. Yet we know that visual, and other perceptual, stimuli and codes have a subtle but powerful influence on us all (Schneider, 2003). It is therefore unlikely that real success in attempts to reconstruct gender structures can be reached by too much focus on language. Perhaps virtual characters can help in the endeavour.

[3] *SitePal*: www.sitepal.com; *PeoplePutty*: www.haptek.com; *Meez*: www.meez.com

Acknowledgements

Thanks to GLIT, LearnIT and the KK-Foundation for financial support. Also thanks to Adam Altmejd and Aron Vallinder for conducting a pre-study, to Betty Tärning and Thérèse Deutgen for collaboration in the study, and to Else Rommes for sharing valuable manuscripts!

Not the least, thanks to all student participants and to helpful teachers and school leaders at the participating schools.

References

Altmejd, Adam & Vallinder, Aron (2007) *Betydelsen av en virtuell agents sociala förmåga*. [The significance of the social skills of virtual agents]. Student research paper, Lund University, Lund, Sweden.

Bandura, Albert (1977) *Social Learning Theory*. Prentice Hall.

Bandura, Albert (2000) *Self-efficacy: The foundation of agency*. Mahwah, NJ: Lawrence Erlbaum.

Baylor, Amy & Plant, Ashby (2005) 'Pedagogical agents as social models for engineering: The influence of appearance on female choice', *Proceedings of Artificial Intelligence in Education* 2005, Amsterdam, the Netherlands.

Baylor, Amy, Rosenberg-Kima, Rinat & Plant, Ashby (2006). 'Inteface agents as social models: The impact of appearance on females' attitude toward engineering', Proceedings of CHI 2006, Montreal, Canada.

Brave, Scott & Nass, Cliff (2005) *Wired for Speech*. Cambridge, MA: MIT Press.

Chess, Shira (2006) *Working with Sex, Playing with Gender: Locating Androgyny in Video Games*. Available at: http://www.shiraland.com/Work/gender_media.pdf (last accessed December 2007)

Cockburn, Cynthia & Ormod, Susan (1993) *Gender and Technology in the Making*. London, Sage.

Coltrane, Scott & Adams, Michele (1997) Work-family imagery and gender stereotypes: Television and the reproduction of difference. *Journal of Vocational Behavior*, 50, 323-347.

Faulkner, Wendy (2003) "Teknikfrågan i feminismen", [The technology issue in feminism] in. Boel Berner (ed), *Vem tillhör tekniken?[Who owns technology]*, Arkiv förlag, Lund, Sweden.

GameGirlAdvance (2004) *Game+Girl=Advance*. Web-blog available at: http://www.gamegirladvance.com/archives/2003/12/02/its_time_for_androgyny_its_just_vaan.html (last accessed December 2006)

Gilmour, H<<. (2004). What girls want: The intersections of leisure and power in female computer game play. In: C. Carter & L. Steiner (Eds.), *Critical Readings: Media and Gender*, 328-352. Maidenhead, UK & New York, NY: Open University Press/McGraw-Hill.

Gulz, Agneta & Haake, Magnus (2006a) 'Pedagogical agents – design guide lines regarding visual appearance and pedagogical roles', *Proceedings of the IV International Conference on Multimedia and ICT in Education (M-ICTE2006)*, Seville, Spain.

Gulz, Agneta & Haake, Magnus (2006b) 'Visual design of virtual pedagogical agents: Naturalism versus stylization in static appearance', *Proceedings of the 3rd International Design and Engagability Conference @ NordiChi 2006*, Oslo, Norway.

Gulz, Agneta, Ahlner, Felix & Haake, Magnus (2007) 'Visual femininity and masculinity in synthetic characters & patterns of affect', pp. 654-665 in: Ana Paiva, Rui Prada. & Rosalind Picard (eds) *Proceedings of the Second International Conference on Affective Computing and Intelligent Interaction (ACII 2007)*, Lisbon, Portugal, 2007, Berlin/Heidelberg: Springer.

Gulz, Agneta, Haake, Magnus & Tärning, Betty (2007) 'Visual gender and its motivational and cognitive effects – a user study'. *Lund University Cognitive Studies. 137.*

Haake, Magnus & Gulz, Agneta (2007) 'Virtual Pedagogical Agents: Stylisation for Engagement. *Interfaces*, 70: 12-13.

Hannover, Bettina & Kessels, Ursula (2004) 'Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science'. *Learning and instruction*, 14, 51-67.

Haraway, Donna, (1991) A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In (Ed.), Haraway, D: *Simians, Cyborgs and Women: The Reinvention of Nature*, 149-181. New York, NY: Routledge.

Hirdman, Yvonne (2003) *Genus [Gender]*. Liber AB, Stockholm, Sweden.

Kessels, Ursula (2005). Fitting into the stereotype. how gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education* 20(3), 309-323.

Kihl, Maria (2003) *Välkommen I LTH? – en studie av D02:s och C02:s första år på Teknis*. [Welcome to LTH? A study of the first year at the college of engineering for computer technology students]. Report. Faculty of Engineering, Lund University, Sweden.

Lippa, Richard (2001) 'On deconstructing and reconstructing masculinity-femininity', *Journal of Research in Personality*, 35: 168-207.

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

Nass, Cliff, Isbister, Katherine & Lee, Eun-Ju (2000) 'Truth is beauty: researching embodied conversational agents', pp. 374-402 in Justine Cassell, Joseph Sullivan, Scott Prevost and Elizabeth Churchill (eds), *Embodied Conversational Agents*, Cambridge, MA: MIT Press.

Rommes Els, Overbeek Geertian Engels Rutger, de Kemp Raymond & Ron Scholte (2007) 'I'm not interested in computers': Gender-based Occupational Choices of Adolescents, *Information, Communication & Society*, 10(3): 299-319.

Rommes, Els (2007) 'Images and Identities; Sex, Sexuality and Soaps', available at: www.helsinki.fi/kristiina-instituutti/conference/pdf/rommes.pdf (last accessed May 2007).

Schneider, David (2003). *The Psychology of Stereotyping*, New York, NY: Guilford Press.

Turkle, Sherry. (1995). *Life on the Screen: Identity in the Age of the Internet*. New York, NY: Simon & Scuster <<ev ej

Voelker. David (1994). *The Effects of Image Size and Voice Volume on the Evaluation of Represented Faces*. PhD Thesis, Stanford University, Palo Alto, CA.

Reeves, Byron & Nass, Cliff (1996) The Media Equation: How People Treat Computers, Televisions and New Media like Real People and Places. New York, NY: Cambridge University Press.

LUNDS UNIVERSITET

E-husets tryckeri (2009), Lund, Sweden