

Evidence-based facial design of an interactive virtual advocate

W A Powell¹, T A Garner², D Tonks¹, T Lee¹

¹School of Creative Technologies, University of Portsmouth,
Winston Churchill Avenue, Portsmouth, UK

²Centre for Child Protection, University of Kent,
Giles Lane, Canterbury, UK

wendy.powell@port.ac.uk, t.garner@kent.ac.uk, daniel.tonks@port.ac.uk, thomas.lee@port.ac.uk

¹*www.port.ac.uk*, ²*www.kent.ac.uk*

ABSTRACT

RITA (Responsive InTeractive Advocate) is the vision for a computer software-based advocacy and companion service to support older adults and provide an alternative to institutional care. The RITA service will offer a preventative care approach, creating a digital champion who will learn an individual's needs and preferences over time, and be a friendly interface between users, family and professionals. This will involve the integration of a variety of technical components: (1) *The Face* - a realistic and emotionally expressive avatar, encouraging communication and interaction; (2) *The Mind* - a repository to store, organise and interpret personal and memory-related information representing the "essence" of a person, with user-defined access controls; (3) *The Heart* - an empathetic sensory interface which is able to understand and respond to the physical, emotional and psychological needs of the user. Each of these aspects presents a series of technical challenges, which will be addressed by combining existing state-of-the art techniques from a variety of disciplines, together with innovative processes and algorithms, to improve and extend functionality. RITA is being designed in consultation with user groups and service providers, and drawing extensively on existing research to inform the design and functionality of the system. In this short paper we introduce the design and development of the face of RITA.

1. INTRODUCTION

Around 3 million over 65's in England and Wales live alone (ONS 2011), with over half suffering from long-term health or mobility issues. Increasingly pressured lifestyles place burdens on extended family, restricting the opportunities for care and support (Merrill 1997; Silverstone and Hyman 2008), and increasing demands on the social care system mean that it is unable to address many of the issues associated with social isolation. Frequent changes between carers, alongside poor communication (across a fragmented healthcare system) presents a significant challenge in the provision of personalised support which is responsive to the needs, preferences, history and personality of an individual.

RITA (Responsive InTeractive Advocate) is the vision for a computer software-based advocacy and companion service that brings together three elements: a 3D virtual avatar and conversational agent, an 'essence' repository for storage and organisation of various forms of information pertaining to the user, and an empathetic communication system that is capable of understanding and responding to the psychological, social and emotional needs of the individual user. It is the first of these elements with which we are concerned here.

2. AVATAR CHARACTERISTICS: REQUIREMENTS AND EVIDENCE

Although it is recognised that visual rendering of an avatar influences the users perception of the characters personality (Dryer 1999), there is often a lack of systematic and informed design of virtual agents (Gulz, 2005; Gulz and Haale, 2006). Perception of character is subject to multiple and diverse influences, ranging from character type to clothing, facial expression to body language (Gulz 2005) and thus it can be challenging to decide on individual characteristics for any given avatar. However, a clearly defined purpose, together with elicitation of user requirement, can facilitate the process of meaningful and effective avatar design. The RITA animation team worked closely with a health care service design company in order to identify the key personality

characteristics required for the RITA avatar. The main characteristics identified were *trustworthiness* and *competence*.

The first consideration was whether the avatar should be human or non-human, and the level of realism necessary. Whilst animal or other non-human characters can be perceived as friendly and engaging, human characters are considered to be more attractive (Schneider et al, 2007). Research within this area is currently lacking in fully evidenced conclusions with regards to the more highly-specific questions, such as: *What are the advantages of using a human avatar within a healthcare and wellbeing context? Does a human avatar evoke a greater sense of professionalism and competence?* We can make a certain degree of inferences from the various existing healthcare-associated avatar products, the majority of which are human in design (see Osaine¹, Cisco²).

The Uncanny Valley theory (Mori, 1970; Mori et al, 2012; Tinwell, 2011) suggests a danger in that increasing the human realism of a character may lead to a sense of discomfort if the end result looks almost, but not quite, human. In light of this, the avatar is based on a human character but using techniques to maximise the realism and visual fidelity, in order to reduce the sense of discomfort whilst optimising the required characteristics.

There are a number of factors which have been identified as being key elements defining the character of virtual agents, including movement and hand gestures (Johnson, Rickel and Lester, 2000), voice and verbal communication (Nass et al, 1994; Cassell et al, 2000), facial and emotional expression (Lester et al, 2000), and facial characteristics (Gulz 2006). The notion that perception of character can be drawn from the face has remained a popular cultural belief over the centuries (Berry and McArthur, 1986) and it could be asserted that a personality profile can be projected onto a person based upon an observer's perception (physiognomic stereotyping, see Hummert et al, 2014). The face is a powerful communication tool in which the majority of human expressions are formed during conversation (see Argyle 2013). Consequently, the RITA avatar will initially be displayed from the neck upwards. This maintains the key features necessary for engaging interaction, whilst reducing the computational load required for high fidelity realistic animation of a full-body avatar. In light of this, the key features for consideration in avatar design will be facial characteristics, voice and facial expression.

2.1 Facial Characteristics

It is clear that people make enduring judgements of character based on less than one second of exposure to a face (Todorov 2008). The perception of trustworthiness and competence appears to be based on variations in a few key facial features, summarised in Table 1.

Table 1: Facial features associated with trustworthiness and competence.

Feature	Competence / Dominance	Trustworthiness / Honesty
Face shape	Not too round (Oosterhof 2008, Poutvaara 2009) Angular or square face (Zebrowitz 2008)	Round or baby face (Kleisner, 2013) Smooth smaller face (Todorov, 2013) Fatter face (van Vugt 2006) Pronounced cheekbones (Todorov, 2013)
Chin	Wider chin, not too round ((Zebrowitz 2005)	Wide chin (Todorov, 2013) Rounder broad chin (Kleisner, 2013) Shorter chin (Kleisner, 2013)
Nose	Bridge not pronounced (Zebrowitz 2008)	Slim nose with shallow nose sellion (Todorov, 2013)
Mouth	Thinner lips (Zebrowitz 2008)	Broad mouth (Kleisner et al, 2013) Smiling or upward corners (Kleisner et al, 2013)
Eyebrows	Lower brows (Zebrowitz 2008)	High inner eyebrows (Todorov et al, 2013) Eyebrows closer together (Kleisner et al, 2013)
Eyes	Not too round (Zebrowitz 2008)	Brown eyes (Kleisner 2013) Large, open eyes (Todorov et al, 2013)
Hair	Brunette (Kyle 1996)	No data

Based on combining the key attributes from the available evidence, the initial facial mesh was constructed with a fairly round face shape, but with a wider than average jaw, slightly pronounced cheekbones and with the more angular features slightly softened. The mouth is slightly broad, with an upward turn at the corners. The nose is slim and with a shallow sellion. The inner brows are slightly higher than average and not too far apart, with the brow ridge being less pronounced, resulting in less shadow on the eyes, creating an impression of brighter and more open eyes (Figure 1a).



Figure 1. (a) The base facial mesh designed for competence and trustworthiness (b) Actress engaging in motion capture process (c) emotions from actress animated on the base mesh (d) textured avatar.

2.2 Voice

The RITA avatar voice has two elements to consider, accent and tonal quality. The latter is dependent on context, emotional overlay and speech elements, but the characteristics of the accent persist through all dialogue and therefore warrant some consideration. The cultural variation that would likely dictate users from alternative ethnic backgrounds would possess different vocal preferences for their RITA (see Kooshabeh et al, 2014) suggests that, in an ideal scenario, the vocal characteristics of RITA would need to be fully customisable to best suit the individual user. Within the current stage of development however, we are required to focus upon a single accent. Research by Tamagawa and colleagues (2011) suggests that, within healthcare robotics, there is a significant user-preference towards regional-sounding accents. Within the UK, a 2009³ survey reported Received Pronunciation as the most appealing when speaking to a call centre, followed closely by a Scottish accent. Specifically, the Edinburgh accent has been associated with pleasantness and prestige⁴ BBC online⁵ published the results of an opinion poll that positioned Scottish accents as the ‘most reassuring’ during an emergency and accents from this region are largely connotative of competence, trustworthiness and friendliness. In light of this evidence, a Scottish actress was employed to provide source recordings as the basis for the voice of the avatar (Figure 1b).

The base mesh for the face was built in Autodesk Maya. A bone-based animation system drives the movements of the neck, jaw and eyelids, and blend shapes applied to simulate the macro and micro movements of the facial musculature. Once the mesh was fully rigged, the team used specialist facial capture software (Faceshift) to record facial movements and emotional expression from the actress. Animations were split into core speech and emotional state segments, before being imported into Maya and applied to the facial rig and refined by hand to optimise the fidelity (Figure 1c).

4. INTERACTION AND ONGOING DEVELOPMENT

The animated avatar is deployed in the Unity engine to allow real-time control of the avatar in response to signals from the control software. In order to build a modular and adaptable system, the “heart” and “mind” software send information to a decision-making system which will interpret the information relating to the users emotional and physiological state, personality and preferences, sending instructions to Unity to trigger the avatars response. Animation segments are coded and can be dynamically blended with emotional expression and audio files in order to generate appropriate speech, facial expression and synchronised speech.

The essence database is being developed and extended, and will be integrated with learning algorithms in order to populate itself over time as it interacts with the end-user. The modular RITA system can evolve over time, integrating developments in natural speech synthesis, data interpretation and emotional and physiological sensing. The project team engage regularly with service providers and service users in order to refine the capabilities and potential of the RITA advocate.

Acknowledgements: This research is supported by the Technology Strategy Board (Long Term Care Revolution). We thank our colleagues on the RITA project, Dr Jane Reeves (University of Kent), Dr Blair Dickson (Affective State Ltd) and Dr Valerie Carr (Snook) who provided insight and expertise to support the design of the avatar

5. REFERENCES

- Cassell, J, Sullivan, J, Prevost, S, and Churchill, EA, (2000), *Embodied conversational agents*, MIT, Cambridge, MA
- Dryer, DC, (1999), Getting personal with computers: How to design personalities for agents, *Applied Artificial Intelligence*, **13**, 2, pp.273-295.
- Gulz, A, (2005), Social enrichment by virtual characters—differential benefits, *Journal of Computer Assisted Learning*, **21**, 6, pp.405-418.
- Gulz, A, and Haake, M, (2006), Design of animated pedagogical agents - A look at their look. *International Journal of Human-Computer Studies*, **64**, 4, pp.322-339.
- Hummert, ML, (2014), 4 Age Changes in Facial Morphology, Emotional Communication, and Age Stereotyping. *The Oxford Handbook of Emotion, Social Cognition, and Problem Solving in Adulthood*, 47.
- Johnson, WL, Rickel, JW, and Lester, JC, (2000), Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. *International Journal of Artificial Intelligence in Education*, **11**, pp.47-78
- Kleisner, K, Priplatova, L, Frost, P, and Flegr, J, (2013), Trustworthy-looking face meets brown eyes. *PLoS one*, **8**, 1, e53285
- Khooshabeh, P, Dehghani, M, Nazarian, and A, Gratch, J, (2014), The cultural influence model: When accented natural language spoken by virtual characters matters, *Journal of Artificial Intelligence and Society*
- Kyle, DJ, and Mahler, HIM, (1996), The Effects Of Hair Color And Cosmetic Use On Perceptions Of A Female's Ability, *Psychology of Women Quarterly*, **20**, 3, pp.447-455.
- Merrill, D, (1997), *Caring for elderly parents: Juggling work, family, and care-giving in middle and working class families*. Westport, CT: Auburn House
- Mori, M, (1970), The uncanny valley, *Energy*, **7**, 4, pp.33-35.
- Mori, M, MacDorman, KF, and Kageki, N, (2012), The uncanny valley [from the field], *Robotics & Automation Magazine, IEEE*, **19**, 2, pp.98-100.
- Nass, C, Steuer, J, and Tauber, ER, (1994), Computers are social actors. *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*. ACM, Boston, MA, pp.72–78.
- Oosterhof, NN, and Todorov, A, (2008), *The functional basis of face evaluation*. Proc Natl Acad Sci U S A, **105**, 32, pp.11087-92.
- Poutvaara, P, Jordahl, H, and Berggren, N, (2009), Faces of politicians: Babyfacedness predicts inferred competence but not electoral success. *Journal of Experimental Social Psychology*, **45**, 5, pp.1132-1135
- Schneider, E, Wang, Y, and Yang, S, (2007), Exploring the uncanny valley with Japanese video game characters, *Proceedings of DiGRA 2007 Conference*
- Silverstone, B, and Hyman, HK, (2008), *You and Your Aging Parent: A Family Guide to Emotional, Social, Health, and Financial Problems*: Oxford University Press
- Tamagawa, R, Watson, CI, Kuo, IH, MacDonald, BA, and Broadbent, E, (2011), The effects of synthesized voice accents on user perceptions of robots. *International Journal of Social Robotics*, **3**, 3, pp.253-262.
- Tinwell, A, Grimshaw, M, Nabi, DA, and Williams, A, (2011), Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior*, **27**, 2, pp.741-749.
- Todorov, A, Dotsch, R, Porter, JM, Oosterhof, NN, and Falvello, VB, (2013), Validation of data-driven computational models of social perception of faces, *Emotion*, **13**, 4, p.724.
- Van Vugt, M, (2006), Evolutionary origins of leadership and followership. *Personality and Social Psychology Review*, **10**, pp.354-371.
- Zebrowitz, LA, and Montepare, JM, (2008), Social Psychological Face Perception: Why Appearance Matters, *Soc Personal Psychol Compass*. **2**, 3, p.1497.

¹Wall Street Journal Online

<http://online.wsj.com/news/articles/SB10001424052702304636404577291252691059524>

²The Medical Quack <http://ducknetweb.blogspot.co.uk/2009/10/cisco-and-healthcare-technology.html>

³Reuters <http://uk.reuters.com/article/2009/11/24/us-britain-branding-accents-idUKTRE5AN37C20091124>

⁴BBC Polls http://www.bbc.co.uk/voices/yourvoice/poll_results3.shtml

⁵BBC News <http://news.bbc.co.uk/1/hi/scotland/7754111.stm>