

The Uncanny Valley: does it exist and is it related to presence ?

Harry Brenton (BespokeVR Ltd, harry@bvrhealth.com)

Marco Gillies (University College London, M.Gillies@ucl.ac.uk)

Daniel Ballin (BT, Daniel.Ballin@bt.com)

David Chatting (BT, David.Chatting@bt.com)

Abstract

The term 'Uncanny Valley' relates to the sense of unease and discomfort experienced when people look at realistic virtual humans. There is growing academic interest in the Uncanny Valley, however, our understanding of it is limited, and there has been a lack of rigorous investigation to determine whether the phenomenon actually exists. Researchers into immersive virtual reality have defined the term 'presence' as 'the feeling of being bodily in an externally-existing world'. The Uncanny Valley is relevant to a study of presence because it questions widely held assumptions about the correlation between realism and believability within a virtual world. This paper suggests that the theory of a 'break in presence' may contribute to our understanding of the Uncanny Valley. The uncanny effect may be analogous to a strong sense of presence clashing with cues indicating falsehood, within a VR world. The eyes of virtual characters are probably crucial elements in this interaction because of their key role in conveying intentional states to other organisms. As immersive environments grow increasingly realistic these may themselves generate a type of Uncanny Valley response thus far only reported when observing virtual humans.

Keywords

Uncanny, valley, realism, presence, social co-presence, virtual reality

1 Introduction

The term 'Uncanny Valley' originally referred to a graph plotting emotional response against the similarity of a robot to human appearance and movement (figure 1). As a machine acquires greater similarity to a human, it becomes more emotionally appealing to the observer. However, at a certain point, the increase in similarity brings about a very strong drop in believability and comfort. In the graphs below, the Uncanny Valley is represented by a sudden dip recording where observers have reacted negatively to something perceived as being disconcertingly close to human. The Uncanny Valley has been discussed by Reichardt (2004), Bryant (2001) and Dautenhahn (2004), who attribute the term to the Japanese roboticist Masahiro Mori (although the cited reference, Mori 1982. does not mention it).

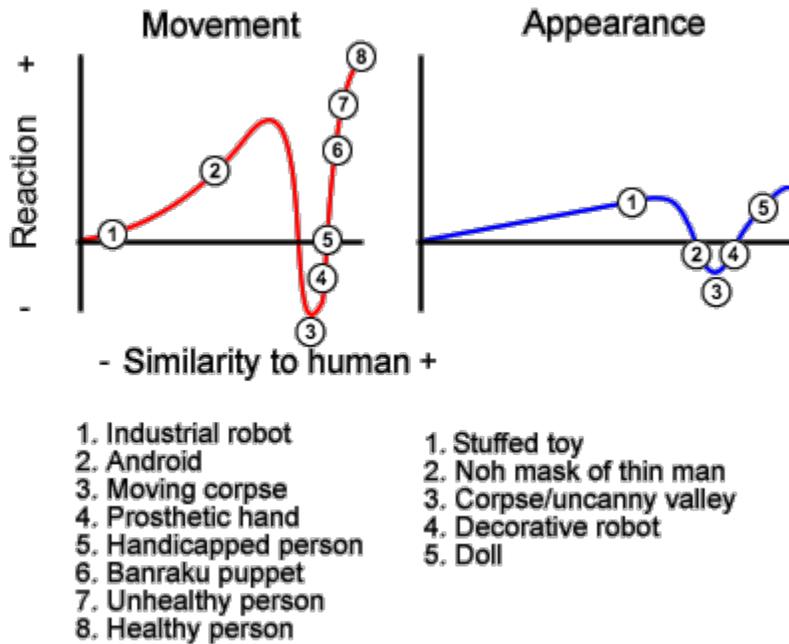


Figure 1. Graphs of emotional response against similarity to human appearance and movement (adapted from Reichardt and Bryant)

Although the Uncanny Valley has attracted great interest in the computer graphics and virtual reality community and is often cited at conferences, our understanding of it is limited and there has been little rigorous research into the phenomenon. This paper aims to provide a starting point for research and further debate.

2 Background

Computer generated characters in films, humanoids in horror movies and photorealistic sculptures are all useful examples of phenomena provoking uncanny responses to artificial humans. Tomlinson (2000) suggests that these responses often occur when an object can be categorised neither as living or dead.

2.1 Horror movies and Freud

In an essay on the nature of the uncanny, Freud (1955) describes his extreme discomfort at seeing someone wearing a prosthetic limb. He argues that uncanny reactions occur when something alien is presented in a familiar context or setting. Many horror movie directors have deliberately exploited this phenomenon to heighten an audience's sense of fear and dread. A classic example is Frankenstein's monster (Whale, 1931), an abomination who exists in a liminal realm between the living and the dead, simultaneously provoking sympathy and disgust. Other examples include the film *Invasion of The Body Snatchers* (Siegel, 1956), many films and novels by Stephen King and the film *Artificial Intelligence: AI* (Spielberg, 2001), in which a synthetic boy's ambiguous nature triggers disturbed and fearful reactions in his adoptive mother.

2.2 CGI characters in film

Realistic depiction of lifelike humans is more advanced in film computer-generated imagery (CGI) than within current real time virtual environments. Animators working with CGI characters have described the Uncanny Valley as a design limitation, which they have had to find creative ways to work around. The team behind *Shrek* (Adamson, 2001) made Princess Fiona less human because 'she was beginning to look too real, and the effect was getting distinctly unpleasant'. The animation director on *Final Fantasy: the spirits within* (Sakaguchi, 2001) describes an eerie sensation as he worked with increasingly photorealistic models: 'it begins to get grotesque. You start to feel like you're puppeteering a

corpse' (both quoted in Weschler, 2002). These comments make it clear why the Uncanny Valley is also referred to as the 'zombie effect' (Dautenhahn, 2004).

2.3 Photorealistic art

The sculptor Duane Hanson made highly realistic sculptures of people, often placed in locations where they could easily be mistaken for a real person. Viewers are confused because they are initially uncertain if they are looking at a human or a work of art. Hanson's sculpture *The Jogger* is particularly disturbing because it depicts a middle-aged man sitting on the floor in pain; a typical audience instinctively gathers around him to see if he needs help. Another artist working in this area is Ron Mueck (Heiner Bastian, 2004) who creates 'hyper-realistic' sculptures modelled at a range of different sizes and scales (Carter, 2003).

2.4 Living or dead?

Tomlinson (Tomlinson, 2000) argues that for survival reasons humans developed mental triggers that help us to quickly categorise entities as being alive or dead. By this measure Frankenstein's monster and Duane Hanson's *Jogger* are both abominations because they simultaneously trigger sympathetic responses indicating life and negative responses indicating death: 'A creature that seems almost human is a conundrum for our pigeon-holing mechanism. Often, these difficulties arise from the Uncanny Valley - where things are in the gray area between matching and not matching an abstraction. Abominations also arise from objects that fit multiple categories that were thought to be mutually exclusive'.

3 Does the Uncanny Valley exist?

The above examples provide good anecdotal evidence that an Uncanny Valley reaction can occur when viewing disconcertingly lifelike artificial humans. However that does not mean that a dip on a graph accurately describes the phenomenon. The two graphs in Figure 1 are widely circulated and redrawn from Reichardt (2004) but their validity should be questioned. For example, in the movement graph there is an asymptote going through points 1 & 2 and another asymptote going through points 3,4,5,6,7,8. The middle section is referred to as a valley but this is only because it has been drawn that way. It is effectively a step function, which would be better represented by two separate curves rather than a continuous line.

The graphs are cryptic and ambiguous: it is uncertain exactly what Mori meant by a 'doll' or a 'decorative robot'. Even more bewildering is the fact that neither the term Uncanny Valley nor the graph appears in the English translation of Mashiro's book *The Buddha in the Robot* (Mori, 2004) - although this could be because of a variant edition published only in Japan.

3.1 Why is the existence of the Uncanny Valley important ?

The roboticist David Hanson rejects the Uncanny Valley as 'pseudoscientific' (Ferber, 2003) and argues that robot designers should not be conceptually limited in their work by a theory that is scientifically unproven. Likewise, modellers and animators working with interactive computer characters should not necessarily treat the Uncanny Valley as an immutable law of design. This is becoming increasingly important with the release of games such as *Half Life 2* which are close to delivering real-time photo-realistic characters and environments on a standard PC. Thompson (2004) questions the assumption held by some game studios that high resolution photo-realistic characters will in themselves make a game better. Research that could offer substantiated guidelines for building realistic characters that do not fall foul of the Uncanny Valley would be useful for both game developers and designers of middleware tools for character creation.

3.2 Relevant studies with interactive animated characters

Avatars in collaborative virtual environments and autonomous social agents have been extensively studied experimentally. However, the number of studies that are relevant to the Uncanny Valley is limited. Few manipulate realism as an experimental condition, and generally the level of realism in the models is low. A further problem is that character designers appear to instinctively to shy away from models that exhibit signs of uncanniness.

Nowak & Biocca (2001) show that, up to a certain level, increasing realism has a positive effect on users' evaluation of a character. They also give anecdotal evidence that high realism can seem uncanny in certain circumstances. Tromp et al. (1998) give a particularly clear example, explicitly stating that a subject was disturbed by the fact that an avatar was graphically realistic while this realism was not matched with behavioural realism, making the avatar seem 'like a zombie'. This was reported only by one person, supporting the notion that the uncanny response may be subjective. Evidence that behavioural realism should be matched with graphical realism also comes from Garau et al., (2003), who found that while high behavioural realism was preferred for a graphically realistic avatar, the opposite was true with a graphically unrealistic one. Nass, Isbister, & Lee (2000) present a similar result: that personalities portrayed verbally have to match those portrayed non-verbally.

4. Is the Uncanny Valley related to presence ?

Slater (2004) presents a theory of presence that proposes that at any given moment in time the human brain formulates hypotheses about the world based on its perceptions. Various sensory percepts are organised into Gestalts, an overall hypothesis suggesting what we are perceiving and where we are. When humans experience a virtual reality system, the sensory percepts are ambiguous: a number of cues affirm the real location in a VR lab; another system of cues within the VR system indicates that we are in an alternative 'virtual' location. These sets of cues tend to form two alternative Gestalts. For example in a VR simulation of a meadow, one perceptual hypothesis suggests we are within a box with light projected onto four walls, another hypothesis indicates that we are within a field surrounded by flowers. At any time the brain will resolve to pick the hypothesis that corresponds to the location where we feel present. The hypothesis is likely to be chosen based on which of the two sets of cues is strongest and also based on internal mental factors - in fact we can be present in entirely mental spaces such as daydreams. However, slight changes in our percepts (or our mental state) can trigger a switch of hypothesis that Slater terms a Break in Presence (BIP).

Tomlinson (2000) suggests that the Uncanny Valley arises from an inability to categorise something as either dead or alive. This 'conundrum in our pigeon-holing mechanism' is analogous to Slater's Gestalt theory of presence because at any time the brain is choosing between two different hypotheses: 'this is a human being' versus 'this is not a human being'. BIPs and the uncanny effect are different domains because one concerns location and the other the nature of an object, however both theories are linked and underpinned by a consideration of the grey area between two possible hypotheses. If a virtual environment is too unconvincing, the user will remain entirely present in the real world and no BIPs will occur; if a character is insufficiently human-like, then the viewer will characterise it as non-human and will not experience the uncanny effect. Both BIPs and the uncanny effect occur when the two sets of cues are reasonably well balanced. However, Slater characterises a break in presence as a simple switch between two possible ways of looking at the world. People experiencing BIPs do not normally report a disturbing or uncanny effect.

There are two theories which qualify the relationship between the uncanny valley and BIPs:

- A switch between hypotheses is normally emotionally neutral but under certain conditions it becomes emotionally disturbing. This emotional disturbance is related to the level of realism involved: the stronger the cues that an object is human the stronger the 'this is human' hypothesis and the more emotionally jarring the switch. The depiction of a human is probably

more emotionally engaging than that of an environment because of biological conditioning in the way we relate to other human beings (Tomlinson 2000)

- The Uncanny Valley is not a switch but a superposition of two hypotheses. It may be the case that the 'this is human' hypothesis remains dominant: the non-human cues are strong enough to create doubts that we find disturbing, but not strong enough to cause a switch. This view has credibility because the uncanny effect appears to be an ongoing state rather than a momentary switch. The lack of a clear-cut switch would also explain what happens with low levels of realism. With an unrealistic character we can still maintain the belief that it is human to a degree (we can interact with it as if it were human) while mostly believing it is not human.

These considerations may provide valuable feedback into theories of presence within virtual environments. It may be worth investigating whether at certain times BIPs can be emotionally disturbing or whether there are interesting intermediary states between presence in one environment and another. An important question is whether there is an equivalent to the Uncanny Valley for environments as well as humans. As immersive environments grow increasingly realistic they may generate an Uncanny Valley response. This would suggest an amendment to the claim made by Waterworth (2003) that 'the level of experienced presence is an inverse function of the degree of abstraction of the media'. There may be a sudden strong drop-off in presence at a point when a high fidelity, low abstraction environment is almost but not quite distinguishable from reality. It is also possible that the Uncanny Valley is unique to our perception of human-like objects due to an innate attunement to human beings.

Even if the Uncanny Valley relates only to virtual characters and not to VR environments, there is a good practical reason for researchers into presence to examine the phenomenon. An uncanny character within an otherwise believable and realistic VR world may be the weak link that causes an unwanted break in presence.

5. Research and further work

Having a testable hypotheses is a prerequisite for a scientific study. However we are not proposing a complete and consistent theoretical model at this juncture because our understanding is still too limited. Instead we offer five working hypotheses intended to provoke further research and debate:

1. The Uncanny Valley response correlates to a mismatch or break in presence: cues indicating believability and realism clash with cues indicating falsehood
2. Increasing realism heightens sensitivity to cues indicating falsehood
3. Perceptual cues indicating falsehood are especially potent in the eyes of virtual characters
4. Research methodologies used to study breaks in presence and social co-presence can be used to study the Uncanny Valley
5. The Uncanny response is dynamic and subject to change over time

5.1 The Uncanny Valley response is related to presence and breaks in presence

Existing methods of studying and measuring presence (Slater and Steed 2000; Slater, Brogni and Steed 2003) might be adapted to study the Uncanny Valley hypothesis.

5.2 Increasing realism heightens sensitivity to cues indicating falsehood

Animated characters give off a number of perceptual cues that suggest they are people. For example, a simple character might move its eyes in a way that we recognise as similar to human behaviour, and so we accept it as being a person on some level. At all levels of realism it is possible to create characters that are 'not quite right' or that 'don't quite work', while others are appealing and easily accepted. Highly realistic characters operate within a different set of aesthetic constraints than those exhibiting lower realism. As a character becomes increasingly realistic it is constrained to the physical attributes of a real person and a viewer's tolerance for abstracted appearance and behavior is reduced.

Current technology allows increasingly graphically realistic characters but often their behaviour and movements do not match up to this realism. The high graphical realism gives strong cues that suggest the character is a person and thus raises high expectations for motion and behaviour. When the quality of these does not match up, it suggests the character is in fact not real. This creates a perceptual paradox which may generate the sense of the uncanny.

5.3 Perceptual cues indicating falsehood are especially potent in the eyes of virtual characters

The Uncanny Valley is probably related to our innate ability to extract social and emotional information from human faces. We are highly attuned to the cues given by faces, and therefore find it easier to spot small variations than when looking at inanimate objects (Donath, 2001). Sculptors and CGI artists aiming at high levels of realism often obscure the eyes of their creations, probably because the lack of an intelligent witnessing consciousness behind them would dispel the illusion of reality, as for example the agents in *The Matrix* who wear dark glasses, (Silberman, 2003). The artist Michael Najjar has created an eerie series of photographs depicting normal people with manipulated eyeballs The Nexus project (Najjar, 2004). This disturbing 'undead' effect has been used for so many zombies in horror movies that it has become a cliché parodied in films such as *Shaun of the Dead* (Wright, 2004). The eyes are especially powerful in our recognition of 'life' because they allow us to attribute a perceptual state to another organism (Weschler, 2002) and they also play a crucial role in understanding the intended behaviour of others (Baron-Cohen, 1995). When the eyes communicate a different (or ambiguous) intent to an ensuing unpredictable behaviour there is a mismatch which may promote uncanniness.

5.4 The uncanny response is culturally dynamic and subject to change over time

We have grown comfortable with representations of humans on film and it no longer seems uncanny to stare at gigantic flat representations of ourselves projected onto a screen. The Uncanny Valley describes an emotional reaction which may be subject to change over time. Currently 'uncanny' avatars may foster a climate of their own acceptance once we become used to looking at them. For example Duane Hanson's sculpture *The Jogger* is less uncanny the second time it is viewed because you are expecting it and have pre-classified it as a dead object. Thompson (2004) suggests that reactions to people with plastic surgery may indicate future responses to avatars in computer games: 'face-lifts used to look horrifically strange but now go unnoticed. Likewise, we've played with dead, fish-eyed characters for so long that they seem kinda normal'. 'We've' is an important word: in cultures where plastic surgery is unknown a facelift probably would provoke a far stronger response than in some western cultures. This raises an important issue for future research into the Uncanny Valley. The community of practice reading this paper is likely to have markedly different responses to realistic avatars because it has become inured to them by constant interaction.

6. Future work

Our main research aims are:

- To discover whether the uncanny response can be measured experimentally. As it is currently not well understood, this poses certain problems. It might be most useful to start by researching media in which realism is easier to achieve, such as film or animation, rather than interactive characters. Studies would proceed by manipulating characters and recording which are perceived as uncanny by subjects. Although initial evaluation is likely to be qualitative, we should aim to develop a reliable questionnaire which can be used by researchers in their experiments.
- To produce a set of heuristics that will help creators of animated characters to avoid the uncanny. Our hypotheses suggest that it is important to balance the realism of the various elements (graphics, animation, interaction) and to pay particular attention to certain aspects such as the appearance of the eyes.

Our hypotheses suggest a number of experimental tasks:

- Investigating the effect of mismatching the realism of various cues such as behaviour and appearance (an extension of the work done by Garau et al). This could be achieved by systematically manipulating the realism of cues and presenting them to users for comment. We aim to discover if it is possible to reliably reproduce the uncanny response in this way.
- Investigating the relationship between presence and the uncanny, by looking at correlations between measures of presence and measures of the uncanny response.
- Investigating face perception in greater depth, in particular the effect of manipulating different parts of the face, and determining if the eyes play a special role in causing the uncanny response.

References

- Half life 2 website (2004). <http://www.planethalf-life.com/half-life2/> [On-line].
- Adamson, A. & Jenson, V. (2001). Shrek.
Ref Type: Motion Picture
- Baron-Cohen, S. (1995). Mindblindness: an essay on autism and theory of mind. Cambridge MA: MIT Press.
- Bryant, D. (2001). Why are monster-movie zombies so horrifying and talking animals so fascinating? <http://www.arclight.net/~pdb/glimpses/valley.html> [On-line].
- Carter, J. (2003). Ron Mueck: Hyperrealist sculpture that keeps up with its neighbours. <http://www.bbc.co.uk/dna/collective/A1015435> [On-line].
- Dautenhahn, K. (2004). Design Spaces and Niche Spaces of Believable Social Robots. In (pp. 192-197). Berlin, Germany: IEEE Press.
- Donath, J. (2001). Mediated Faces. Cognitive Technology, 373-390.
- Ferber, D. (2003). The man who mistook his girlfriend for a robot. <http://www.csindy.com/csindy/2003-12-11/cover.htm> [On-line].
- Freud, S. (1955). The Uncanny" [Das Unheimliche trans. James Strachey 1919]. In (pp. 217-256). London: Hogarth Press.
- Garau, M., Slater, M., Vinayagamoorhty, V., Brogni, A., Steed, A., & Sasse, M. A. (2003). The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment. In Fort Lauderdale, FL, USA.
- Heiner Bastian (2004). Ron Mueck. Hatje Cantz.
- Mori, M. (1982). The Buddha in the Robot. Charles E Tuttle Co.
- Najjar, M. (2004). Michael Najjar Homepage. http://www.mkg-hamburg.de/english/ausstell/02_trien/artists/najjar.html [On-line].
- Nass, C., Isbister, K., & Lee, E. J. (2000). Truth is beauty: Researching conversational agents. In Embodied conversational agents (pp. 374-402). Cambridge, MA: MIT Press.
- Nowak, K. & Biocca, F. (2001). Understanding the influence of agency and anthropomorphism on co-presence, social presence and physical presence with virtual humans. In Philadelphia, USA.
- Reichardt, J. (2004). Robots: fact fiction and prediction.
- Sakaguchi, H. (2001). Final Fantasy: The Spirits Within.
Ref Type: Motion Picture
- Siegel, D. D. (1956). Invasion of the Body Snatchers.
Ref Type: Motion Picture
- Silberman, S (2003). MATRIX2 Bullet Time was just the beginning. Wired.
- M. Slater and A. Steed (2000) A Virtual Presence Counter (PDF), Presence: Teleoperators and Virtual Environments 9(5), 413-434.
- Slater, M (2002) Presence and the Sixth sense, in PRESENCE: Teleoperators and Virtual Environments, MIT Press 11(4) 435-439
- M. Slater, A. Brogni, A. Steed (2003) Physiological Responses to Breaks in Presence: A Pilot Study, Presence 2003: The 6th Annual International Workshop on Presence.
- Spielberg, S. d. (2001). Artificial Intelligence: AI.
Ref Type: Motion Picture

Thompson, C. (2004). The Undead Zone: Why realistic graphics make humans look creepy. <http://slate.msn.com/id/2102086> [On-line].

Tomlinson, B. (2000). Dead Technology. *Style*, 33, 316-335.

Tromp, J., Bullock, A., Steed, A., Sadagic, A., Slater, M., & Frécon, E. (1998). Small group behaviour experiments in the COVEN project. *IEEE Computer Graphics and Applications*, 53-63.

Waterworth, J. & Waterworth, E. (2003). The Meaning of Presence. *presence connect*.

Weschler, L (2002). Why Is This Man Smiling? Digital animators are closing in on the complex system that makes a face come alive. *Wired*.

Whale, J. d. (1931). *Frankenstein*.
Ref Type: Motion Picture

Wright, E. (2004). *Shaun of the Dead*.
Ref Type: Motion Picture