Connection Science

ISSN: 0954-0091 (Print) 1360-0494 (Online) Journal homepage: http://www.tandfonline.com/loi/ccos20

What makes virtual agents believable?

Anton Bogdanovych, Tomas Trescak & Simeon Simoff

To cite this article: Anton Bogdanovych, Tomas Trescak & Simeon Simoff (2016) What makes virtual agents believable?, Connection Science, 28:1, 83-108, DOI: 10.1080/09540091.2015.1130021

To link to this article: http://dx.doi.org/10.1080/09540091.2015.1130021

Published online: 03 Feb 2016.

CONNECTION SCIENCE, 2016
VOL. 28, NO. 1, 83–108
http://dx.doi.org/10.1080/09540091.2015.1130021

What makes virtual agents believable?

Anton Bogdanovych, Tomas Trescak and Simeon Simoff

MARCS Research Institute, School of Computing, Engineering and Mathematics, Western Sydney University, Penrith 2751 NSW, Australia

ABSTRACT

In this paper we investigate the concept of believability and make an attempt to isolate individual characteristics (features) that contribute to making virtual characters believable. As the result of this investigation we have produced a formalisation of believability and based on this formalisation built a computational framework focused on simulation of believable virtual agents that possess the identified features. In order to test whether the identified features are, in fact, responsible for agents being perceived as more believable, we have conducted a user study. In this study we tested user reactions towards the virtual characters that were created for a simulation of aboriginal inhabitants of a particular area of Sydney, Australia in 1770 A.D. The participants of our user study were exposed to short simulated scenes, in which virtual agents performed some behaviour in two different ways (while possessing a certain aspect of believability vs. not possessing it). The results of the study indicate that virtual agents that appear resource bounded, are aware of their environment, own interaction capabilities and their state in the world, agents that can adapt to changes in the environment and exist in correct social context are those that are being perceived as more believable. Further in the paper we discuss these and other believability features and provide a quantitative analysis of the level of contribution for each such feature to the overall perceived believability of a virtual agent.

ARTICLE HISTORY
Received 17 May 2015
Accepted 6 December 2015

KEYWORDS
Virtual agents; virtual reality; believability

1. Introduction

The term "believability" is frequently being used in various disciplines, but is loosely defined. Believability is an essential requirement of modern video games and distributed virtual worlds, hence, more researchers shift their focus to believable agents. As suggested by Livingstone (2006), "the need in modern computer games is not unbeatable AI, but believable AI". In terms of formalisation, believability resembles similarity with intelligence – it is hard to define and formalise. As the result we are witnessing conflicting definitions in existing works and lack of working formalisms for both concepts.

We argue that believability is a more practical concept than intelligence. Hence, we seek to better define the concept of believability by constructing a formal model of believability. In our work we predominantly focus on measuring and witnessing believability purely from

CONTACT Anton Bogdanovych a.bogdanovych@uws.edu.au

agent behaviour. Therefore, here we are mainly concerned with the concept of perceived believability.

This paper attempts to summarise existing believability research in an attempt to better understand the concept and formalise it. As a result, we present a definition of perceived believability, expand its key components and explain those in a formal way. We extend the fundamental work of Loyall (1997), (which we consider to be one of the most comprehensive attempts to analyse the concept of believability) by integrating recent research findings and formalising believability. Based on the resulting formalisation, we have developed a technological framework that integrates all the identified believability features through contemporary artificial intelligence (AI) techniques.

The remainder of the paper is structured as follows. In Section 2 we analyse existing works and definitions of believability to identify its key components and comprehensively define the concept. Section 3 presents the definition and formal model of believability. In Section 4 it is shown how the key components of this formalisation can be implemented. Section 5 presents a case study that uses the resulting believability formalisation for simulating everyday life of aboriginal Australians in 1770 A.D. The evaluation of the believability features, as well as our discussion on the level of contribution for each of the identified features towards perceived believability are presented in Section 6. Finally, Section 7 concludes the presentation and discusses future work.

2. The notion of believability

The notion of believability originates in the field of animation and theatre. A classical work of Walt Disney Studios on animated characters – the "illusion of life" (Thomas & Johnston, 1981) elaborates on the requirements for achieving believability. Believability and realism have been differentiated by Mateas (1999) and Doyle (2002). According to the authors, a believable character is not necessarily a real character, but *must be real in the context of its environment*. Believable agents and believable characters are differentiated in that believable agents are both computer based and interactive (Loyall, 1997).

Contemporary AI uses the term "believability" in relation to engaging life-like systems. Reactivity, interactivity and appropriate decision-making are the common characteristics of believability for autonomous agents (Riedl & Stern, 2006). These characteristics can also be extended with respect to the environment within which they operate.

2.1. Defining believability

In Mateas (1999) a believable character is defined as the one who seems life-like, whose actions make sense, who allows for suspension of disbelief. An extended definition of believable characters is given by Loyall (1997). Here a character is considered to be believable if it allows the audience to suspend their disbelief, but what is also important is a convincing portrayal of the personality of this character. Another definition that emphasises personality and focuses on agents rather than characters is presented in Lester and Stone (1997). Here believability is defined as the extent to which the users interacting with the agent come to believe that they are observing a sentient being with its own beliefs, desires and personality. A contemporary definition that is used in relation to video games states that believability of a virtual agent is associated with giving the illusion of being controlled by a human (Tencé, Buche, Loor, & Marc, 2010).

2.2. Exploring believability features

In our search for individual features that contribute to a virtual agent being perceived as believable we start with listing the key features identified by Loyall (1997), which is one of the most comprehensive works on believable agents and believability characteristics:

- *Personality*: Personality infuses everything a character does, from the way they talk and move to the way they think. What makes characters interesting are their unique ways of doing things. Personality is about the unique and not the general.
- *Emotion*: Characters exhibit their own emotions and respond to the emotions of others in personality-specific ways.
- *Self-motivation* – Characters do not just react to the activity of others. They have their own internal drives and desires, which they pursue regardless of whether or not others are interacting with them.
- *Change*: Characters change with time, in a way consistent with their personality.
- *Social relationships*: Characters engage in detailed interactions with others in a manner consistent with their existing relationships. In turn, these relationships change as a result of their interaction.
- *Consistency of expression*: Every character or agent has many avenues of expression depending on the medium in which it is expressed, for example an actor has facial expression and colour, body posture, movement, voice intonation, etc. To be believable at every moment all of those avenues of expression must work together to convey the unified message that is appropriate for the personality, feelings, situation, thinking, and other behaviours of the character. Breaking this consistency, even for a moment, causes the suspension of disbelief to be lost.
- *Illusion of life*: This is a collection of requirements such as pursuing multiple, simultaneous goals and actions, having broad capabilities (e.g. movement, perception, memory, language), and reacting quickly to stimuli in the environment.

The illusion of life is expanded by Loyall (1997) in terms of (i) appearance of goals; (ii) concurrent pursuit of goals and parallel action; (iii) reactiveness and responsiveness; (iv) situatedness; being resource bounded – body and mind; (v) exist in a social context; (vi) being broadly capable; and (vii) being well integrated (capabilities and behaviours).

Further investigation of the available literature on understanding believability has showed that some of the believability aspects need to be further clarified and the features proposed in Loyall (1997) need to be revised and updated. One of such features is what is labelled as "Emotion". Recent literature (Ortony, 2003) suggests that it is more appropriate to talk about emotional state rather than emotions in general.

2.2.1. Emotional state vs. emotions

Recent work on the use of emotions in achieving believability (Ortony, 2003) suggests that simply displaying emotions is not enough. Instead, the use of emotions has to be consistent with the personality of the virtual character, the situation that unfolds as well as the past history of situations resulting an emotional response. This emotional response is directly linked with the *emotional state* of the agent (Schachter & Singer, 1962). An important addition to the concept of emotions raised by Loyall (1997) is that to be believable a virtual character is not only expected to display emotions in certain situations that normally result an emotional reaction, but that the strength of this reaction depends on the past history of emotional encounters (e.g. agent accumulating anger through multiple unpleasant encounters and then showing a strong emotional reaction to a disproportionally small arousal, such as an accidental push from a stranger), the personality of this agent (e.g. in a similar situation a phlegmatic agent will react differently to a choleric) and the degree to which the situation causing an emotional response affects the emotional state of the agent (e.g. an accidental push by a stranger on a busy street would have a much smaller influence over your emotional state than being fired from your favourite job).

2.2.2. The role of environment

The majority of believability features identified by Loyall (1997) originate in the area of animation, where believable agents are seen as artificial characters that simply appear on the screen for the amusement of the viewer and do not engage in deep interaction with the human observer. Therefore, what is not particularly well investigated in this work is the interactivity aspect and, in particular, how the agent is capable of demonstrating awareness of its environment in interactions with the human. Loyall mentions situatedness and integrity, but these are analysed from the perspective of animated movie or cartoon characters that are not capable of actively interacting with the human user.

The importance of agent integration with the environment is highlighted in Ijaz, Bogdanovych, and Simoff (2011). In this work the emphasis is put on the awareness of the agents about their environment, own state in it, other participants and own interaction capabilities. The authors provide evidence that those features significantly improve the overall believability of the agents.

The concept of *awareness* introduced by Ijaz et al. (2011) is an essential aspect of believability for virtual characters capable of interacting with humans users and, in particular, interact via text conversations or via voice. In a conversations among humans we are normally aware of where we are (environment awareness), who we are (self-awareness) and generally how the interaction is progressing (interaction awareness). Therefore, awareness is an essential component of the believability of embodied conversational behaviour, which we label as "awareness believability". Furthermore, we describe each of the subcomponents of awareness believability.

2.2.3. Environment awareness

The importance of environment awareness for agent reasoning is best illustrated in Elpidorou (2010), where it is suggested that our consciousness does not arise from the brain alone but from the brain's exchange with its environment. Humans are embodied in space and use various cues related to space, like pointing and referring to areas of and things in it, in all they do (for more details see the Chapters 1,2 in O'Keefe and Nadel (1978)).

Existing literature presents a very limited picture on the use of environment awareness by animated agents. Agents like Cosmo (Lester, Voerman, Towns, & Callaway, 1999) and Steve (Johnson & Lester, 2000) are able to recognise and point to objects in the particular static environment but completely ignore their interactions with other participants, do not cater for dynamic environment and cannot orient themselves in a different environment. We suggest that awareness of environment objects alone would not be enough to achieve complete believability and virtual agents further require to be aware of other participants in the context of time.

Our key features of environment awareness include the positions of objects and avatars in the environment, how these evolve with time and the direction vectors associated with avatars (Gerhard, Moore, & Hobbs, 2004).

2.2.4. Self-awareness

Knowing own context and state within the environment, i.e. being self-aware, is essential for a virtual agent to interact believably (Doyle, 2002). To achieve this Doyle (Doyle, 2002) proposes to annotate the environment with meaningful labels and grant agents with access to this annotation. One of the most studied features of self-awareness for virtual agents and animated characters is social role awareness (Prendinger & Ishizuka, 2001). However, self-awareness is a much richer concept and many of its characteristics remain understudied, in particular existing works mostly ignore many vital characteristics that arise in dynamic environments.

Hallowell defines self-awareness (Hallowell, 1955) as the recognition of one's self as an object in the world of objects and highlights the importance of the perception as the key function of self-awareness.

2.2.5. Interaction awareness

Believability of interactions goes beyond traditional focus on modelling the visual co-presence (Gerhard, Moore, & Hobbs, 2005), Context awareness (perceiving other agents/objects in static environments) (Bickmore, Mauer, & Brown, 2007) and communication style (e.g. short vs. long utterances, usage of specific vocabulary) of the agents. Human behaviour in interactions is a result of the mix of being rational, informed, impulsive, and the ability to influence others and cope with the influences from others. All these nuances impact the richness of human interactions, hence, must be taken into account when considering the believability of interactions between virtual agents and humans.

Thus, interaction awareness is defined as the state of an agent who is "able to perceive important structural and/or dynamic aspects of an interaction that it observes or that it is itself engages in " (Dautenhahn, Ogden, & Quick, 2003).

2.2.6. Verbal behaviour

Apart from awareness, two significant components of believable behaviour that are missing in Loyall (1997) are verbal and non-verbal behaviours. These aspects can be seen as belonging to "being broadly capable" component of the illusion of life in Loyall's characteristic (Loyall, 1997), but are just casually mentioned due to agents being able to directly interact with human users via chat or voice not being a focus of Loyall's analysis. The majority of works on believable verbal behaviour, like Eliza (Weizenbaum, 1966) and ALICE (Wallace, 2004), are associated with chatter bots relying on scripted dialogues. Technically, chatter bots parse the user input and use keyword pointing, pattern matching and corpus based text retrieval to provide the most suitable answer from their "knowledge base"

(Gandhe and Traum, 2007), trying to keep a human engaged in a textual or auditory conversation. Interactive virtual characters capable of voice interaction with a human are not as common as chatter bots, but with recent advances in voice recognition became more frequent and believability of voice conversations with such virtual characters becomes and important focus of research (Kenny, Parsons, Gratch, & Rizzo, 2008).

2.2.7. Non-verbal behaviour

Humans complement verbal communication with non-verbal cues, such as facial expressions, body language, and gaze.

Facial expressions can be used to complement the word stream through expressing emotions. These emotional expressions have cross-cultural boundaries, but, generally, existing work deals with a list of emotional expressions: (happiness, sadness, fear, anger, disgust, agreement, disagreement, and surprise) as presented in Cunningham, Kleiner, Wallraven, and Bülthoff (2005).

Gestures allow humans to interact in a lively manner and are an important believability factor. Gesture selection and their correct execution may increase the expressivity of the conversation (Hartmann, Mancini, & Pelachaud, 2005). Believable gestures are related to gesture selection being correctly aligned with the flow of conversation and the generation of realistic movements of agent's upper limbs during the conversation (Hartmann et al., 2005).

Gaze helps to convey the cognitive state of a participant or synchronise a conversation as explained in Lee, Marsella, Traum, Gratch, and Lance (2007). Various gaze models such as avert, examining the current task, and gaze at visitors were simulated by Heylen, van Es, Nijholt, and van Dijk (2005). They measured the believability of an agent based on factors like satisfaction, engagement, natural eye- and head-movements and mental load among others; and this study showed the significant improvements in communication between humans and virtual agents as the result of improved gaze behaviour. Lance in Thiebaux, Lance, and Marsella (2009) contributed to the research on believable gaze by developing a hybrid approach combining head posture, torso posture, and movement velocity of these body parts with gaze shift.

2.2.8. Appearance

In addition to previously mentioned features from Loyall (1997), we add *unique and believable appearance* as an important characteristic of believable virtual agents. Kelley (1950) states that human behaviour towards others is shaped depending on differences in first impressions such that people who have favourable impressions of someone tend to interact more with that person than others having unfavourable impressions (Kelley, 1950). First impressions are, therefore, an important basis for whether humans will build relations with others and find their interactions believable (Bergmann, Eyssel, & Kopp, 2012). Another important line of research that connects appearance and believability investigates the phenomenon of uncanny valley (Mori, MacDorman, & Kageki, 2012), which states that there is a strong relationship connecting human-likeness and believability, but the correspondence between these is no linear and at some stage as the characters become more human-like their believability starts to drop rather than increase.

3. Formalising believability

Based on the analysis of existing works on believability, we try to isolate the key components of believability and define a believable virtual agent as follows.

Definition 1: *A* believable virtual agent *is an autonomous software agent situated in a* virtual environment *that is life-like in its appearance and behaviour, with a clearly defined personality and distinct emotional state, is driven by internal goals and beliefs, consistent in its behaviour, is capable of interacting with its environment and other participants, is aware of its surroundings and capable of changing its behaviour over time.*

Consequently, believability is formalised as follows:

$$\beta = \langle A^{\mathrm{T}}, P^{\mathrm{T}}, E^{\mathrm{T}}, L, SR, \Upsilon, \delta, Aw \rangle. \tag{1}$$

Here β is the believability of a virtual agent, A^{T} are the agent's appearance features, P^{T} is the agent's personality, E^{T} is to the emotional state of the agent, L corresponds to liveness, Aw represents agent's awareness, which we define later, SR stands for social relationships, Υ – represents the consistency constraints and δ – is the change function.

3.1. Appearance

To formalise the appearance, we assume the existence of parametric avatars associated with the corresponding believable virtual agents. These avatars are defined by their visual features, e.g. height, belly size, head size, etc. Trescak, Bogdanovych, Simoff, and Rodriguez (2012a). Each of these parameters has a value in the interval [0,1] where extremes are labelled by the specific state of the visual feature. For example a visual feature height, has a label for the minimum "short" and for the maximum "tall". The appearance (A^{T}) of an individual can then be represented by the following vector:

$$A^{\mathrm{T}} = [\alpha_1 \ldots \alpha_n], \forall i \in [1, n] : \alpha_i \in [0, 1]. \tag{2}$$

Here $[\alpha_1 \ldots \alpha_n]$ are parametric values specifying each of the appearance features.

3.2. Personality

While formalising the personality we consider the assumption of Egges, Kshirsagar, and Magnenat-Thalmann (2003) that a personality has n dimensions, where each dimension is represented by a value in the interval [0, 1]. A value of 0 corresponds to an absence of the dimension in the personality; a value of 1 corresponds to a maximum presence of the dimension in the personality. The personality P^{T} of an individual can then be represented by the following vector:

$$P^{\mathrm{T}} = [\beta_1 \ldots \beta_n], \forall i \in [1, n] : \beta_i \in [0, 1] \tag{3}$$

Here $[\beta_1 \ldots \beta_n]$ are numeric values responsible for each of the personality dimensions.

3.3. Emotional state

The emotional state (E^{T}) is defined following (Egges et al., 2003) as an m-dimensional vector, where all m dimensions of the emotional state are represented by a value in the interval [0,1]. A value of 0 corresponds to an absence of the emotion; a value of 1 corresponds to a maximum intensity of the emotion. Thus, the emotional state E^{T} can be formalised as

$$E^{\mathrm{T}} = \begin{cases} [\beta_1 \ldots \beta_m], \forall i \in [1, m] : \beta_i \in [0, 1] \text{ if } t > 0, \\ 0 \text{ if } t = 0. \end{cases} \tag{4}$$

Here $[\beta_1 \ldots \beta_m]$ represent numeric values that encode the level of intensity for each of the emotional dimensions.

3.4. Liveness

Liveness is agent's ability to express the illusion of life. It incorporates the illusion of life features from Loyall (1997), plus verbal and non-verbal behaviour, as follows:

$$L = < \mathrm{IL}, V_b, \mathrm{NV}_b > . \tag{5}$$

Here IL is a vector responsible for illusion of life, V_b represents verbal behaviour and NV_b represents non-verbal behaviour. We will further explain these elements below.

3.5. Illusion of life

We adapt Loyall's (1997) specification of "Illusion of life", uniting "situatedness" and "integration" into the concept of immersion in 3D virtual environments:

$$\begin{aligned} \mathrm{IL} = \langle &\mathrm{Goals, Concurrency, Immersion,} \\ &\mathrm{ResourceLimitation, SocialContext} \\ &\mathrm{BroadCapability, Reactivity, Proactiveness} \rangle \end{aligned} \tag{6}$$

3.6. Consistency

Consistency across the personality of an agent and other believability characteristics is ensured in our formalisation by the set of consistency constraints (Υ). We formalise those constraints as a penalty function that is 0, if emotional state of the agent and liveness features are inconsistent with the agent's personality and 1 otherwise.

$$\Upsilon : P^{\mathrm{T}} \times L \times E^{\mathrm{T}} \rightarrow \begin{cases} 1 - \text{if consistent,} \\ 0 - \text{if inconsistent.} \end{cases} \tag{7}$$

These constraints must ensure the consistency of the agent behaviour over the entire range of its believability features:

$$\forall p_j \in P^{\mathrm{T}}, \forall l_h \in L, \forall e_g \in E^{\mathrm{T}} : \Upsilon(p_j, l_h, e_g) = 1. \tag{8}$$

3.7. Change

Change (δ) is basically a learning function that allows an agent to update its believability features (β_i) in response to sensing a particular environment state (*EnvState*):

$$\delta : EnvState \times \beta_i \rightarrow \beta_i' \tag{9}$$

3.8. Social relationship

A social relationship (SR), formally speaking, can be represented by a function, which reflects on how the current role being assumed by an agent relates to the roles of other agents. This function results in a numeric value in a range $[0 \cdots 1]$. Here 0 represents no relationship between two roles and 1 – is the highest degree of relation.

$$\forall r_i, r_k \in \text{Roles} : SR = f(r_i, r_k) \in [0 \ldots 1]. \tag{10}$$

3.9. Awareness believability

Awareness is essential part of human conversational behaviour. In a conversation we are aware of where we are (environment awareness), who we are (self-awareness) and generally how the interaction is progressing (interaction awareness). Therefore, awareness is an essential component of the believability of embodied conversational behaviour, which we label as “awareness believability”. Furthermore, we develop each of the subcomponents of awareness believability.

So we can formalise awareness believability as follows:

$$Aw = \langle \text{EA,SA,IA} \rangle. \tag{11}$$

Here EA represents environment awareness, SA is the social-awareness and IA represents interaction awareness.

3.10. Environment awareness

As suggested by Ijaz et al. (2011), the key features of environment awareness include the positions of objects and avatars in the environment, how these evolve with time and the direction vectors associated with avatars. Thus, environment awareness is formalised as follows:

$$\text{EA} = \{\text{Objects,Avatars,Time}\}. \tag{12}$$

Here EA is the set of components of environment awareness and includes the objects in the environment, other avatars representing agents and human participants with respect to the current time.

3.11. Self-awareness

Knowing own context and state within the environment (being self-aware) is essential for a virtual agent to interact believably (Doyle, 2002). The formalisation of self-awareness proposed by Ijaz et al. (2011) is as follows:

$$\mathrm{SA} = \{G, P, B, Sc, St, ObjUsed, Role, Gest\}. \tag{13}$$

Here SA represents the set of components of self-awareness and includes the local goals of the agent (*G*), its current plans (*P*) and beliefs (*B*), current scene where the agent participates (*Sc*), its state within this scene (*St*), objects used by the agent (*ObjUsed*), the role it plays (*Role*) and the gestures being executed (*Gest*).

3.12. Interaction awareness

Human behaviour in interactions is a result of the mix of being rational, informed, impulsive, and the ability to influence others and cope with the influences from others. All these nuances impact the richness of human interactions, hence, must be taken into account when considering the believability of interactions between virtual agents and humans. Interaction awareness is defined as the state of an agent who is "able to perceive important structural and/or dynamic aspects of an interaction that it observes or that it is itself engaged in" Dautenhahn et al. (2003). The components of the interaction-awareness model as outlined in Ijaz et al. (2011) are presented below.

$$\mathrm{IA} = \{\mathrm{AV}_{vis}, \mathrm{AV}_{sc}, Act, Obj, State, Pos, Or\} \tag{14}$$

Here IA represents the set of components included in our interaction awareness model. AV_{vis} corresponds to the set of currently visible avatars. The AV_{sc} is a set of all avatars within the scene where the agent participates in a given moment. *Act* represents the set of actions each of the agents in the current scene is able to perform given its state. *Obj* refers to the list of objects the avatar can use. *State* is the state of the avatar in the world. *Pos* is the position of the agent in the virtual world and *Or* is agent's orientation vector in the virtual world space.

4. Implementation: the I^2B framework

Now that we have a formalisation of believability, next we present our attempt of developing a computational framework implementing this believability formalism. This framework supports the implementation of believable virtual agents for virtual worlds and game engines and is labelled I^2B (interactive, intelligent and believable). It is important to mention that here we do not attempt to develop a comprehensive general-purpose believability framework, but rather present a suggestion on how the aforementioned formalism can be practically implemented (with no claims for this implementation to be the most optimal, unique or comprehensive). The aim of this section is simply to show that the formalism from the previous section is practically useful and can act as a guide for building believable agents. Next we show how each of the components of the above believability formalism can be practically implemented using standard methods and best practices from the literature that were adjusted to fit the formalised models.

4.1. Appearance

Unity, the platform we have selected for testing our implementation, offers mechanisms to represent virtual agents as avatars and define them through a set of parametric features, e.g. (height, head size, arm length, skin colour, etc.). This way of modelling avatars is consistent with the aforementioned formalism.

4.2. Personality and emotional state

One of the most popular modern personality models used in computational psychology is OCEAN (or 'The Big Five') model proposed in John, Donahue, and Kentle (1990). We rely on this model in our framework. This model defines the following five personality traits: {Openness, Consciousness, Extroversion, Agreeableness and Neuroticism}. For modelling the emotional state we rely on the well-known OCC emotional model proposed in Ortony, Clore, and Collins (1988) and with computational implementation proposed in Bartneck (2002). In order for agents to be able to select an appropriate action reflective of its emotional reaction to the state of the environment that is most relevant for their personality, such action has to be annotated by the following *personality facets* (Howard & Howard, 1995): *temptation, gregariousness, assertiveness, excitement, familiarity, straightforwardness, altruism, compliance, modesty* and *correctness*. Using values of personality facets, the agent selects an action that provides the highest utility for its personality type (Bartneck, 2002; Howard & Howard, 1995). Thus, personality (P^T) and the emotional state (E^T) are implemented as an array of variables, where each variable represents a personality feature or an emotional state feature correspondingly.

4.3. Liveness

The implementation of various features of Liveness (L) is based on the Unity 3D technology[1] and the adaptation of a number of contemporary AI techniques. Next we focus on each of the liveness aspects.

4.3.1. Goal generation

A critical aspect in the illusion of life (IL) is to make an agent show that it has certain goals, which it can pursue in a concurrent fashion, as well as change them and prioritise in a reactive and proactive manner. We have developed and integrated all these features into unity (Trescak, Bogdanovych, Simoff, & Rodriguez, 2012b). Agent goal generation is based on agent *motivation*. In the current model, we support physiological motivation where agents proactively try to fulfil their physiological needs, such as hunger, thirst or fatigue. As a part of our future work, we want agents to consider other motivations, such as safety or belonging (social realisation). Furthermore, our model implements the BDI approach (Rao & Georgeff, 1995), allowing for all the standard features of agent-oriented programming offering C# classes for agents, events, plans, beliefs, goals and supporting the message communication, plan selection on receipt of an event, priority planning, etc. Programmers of virtual agents can express beliefs and desires of their agents, decide on the types of events they react to and design the plans to handle those events.

4.3.2. Planning

Every agent in our system relies on a number of plans to satisfy its goals. A plan is a set of instructions, triggered in response to some event. Those events arise as a result of a human- or agent-controlled avatar sending a text command or as a result of an environment state change. The I^2B framework supports static planning – when the entire plan is prescribed by a programmer and is executed by the agent without variation; and dynamic planning, when the agent can sense its current state in the environment and can react to environment changes re-evaluating its current plan. Rather than having a complete recipe for every situation the agent can encounter – the agent is simply given the list of possible actions and has to find a way of combining those to reach its goals. This search is done using a classical depth-first search algorithm (Tarjan, 1972), in which a path between the current state and the goal state is found by evaluating all available actions and analysing their pre-conditions and post-conditions.

4.3.3. Obstacle avoidance and locomotion

In order to believably immerse into its virtual environment and to support the illusion of life while interacting with its environment, the agent must be able to move around without being stuck at an obstacle. This required the implementation of obstacle avoidance techniques. Unity 3D (Pro) offers agent obstacle avoidance based on A^* algorithm (Russel & Norvig, 2003).

4.3.4. Object use

An important aspect of believability is the use objects in the environment (i.e. grabbing a spear, jumping on a boat). We have developed a designated library that provides a set of classes allowing agents to identify an object in the virtual world, attach it to the default attachment point, play a certain animation (i.e. rowing) associated with a given object, wear an object that is a piece of clothing, detach the piece of clothing, drop an object to the ground and detach the object and hide it in the avatar's inventory.

4.3.5. Non-verbal behaviour

Each agent is supplied with a list of possible gestures. Depending on the current emotional state an agent can select a certain gesture and play the corresponding animation. I^2B agents are also supplied with a programming solution dealing with idle gaze behaviour. When the agents are moving around, their gaze is not fixed. The gaze focus keeps changing by our attention-based model. The agent shifts its gaze between objects and avatars depending on the level of its interest in those. The increase and decay of the agent's interest in the surrounding objects will determine the shift in the gaze focus.

4.3.6. Verbal behaviour

The verbal behaviour of the I^2B agents is currently limited to exchanging text messages with other agents and text chats with humans. For chatting with humans I^2B agents employ the ALICE chat engine (Wallace, 2004) based on the AIML language. Each agent uses a number of AIML files that represent what can be seen as a common sense database. Additionally, every agent is supplied by personalised AIML files that define its personality and the data relevant for its role within the virtual society.

4.4. Social relationships

The virtual institutions (Bogdanovych, 2007) technology manages the social interactions and social relationships of the I^2B agents. The approach taken in virtual institutions is to “program” the environment first, in terms of the roles of the agents, their presence, possible scenes, the role flow of the agents between these scenes, interaction protocols of every scene, etc (see Bogdanovych, 2007 for more details on this process). With the help of the underlying virtual institution I^2B agents can also understand which social roles are being played by other agents or humans, and change their roles over time. Based on this information they can engage into believable social interactions and build social relationships. An agent’s personality and the emotional state are impacted by social interactions with others.

4.5. Consistency

Virtual institutions manage the set of rules (social norms) for all participants in the given virtual environment, subject to their roles, hence, they manage the consistency (Υ) of the agent behaviour. The institutional formalisation helps an I^2B agent to assign context to own actions and actions of other participants, thus being able to make the corresponding adjustments to its emotional state, personality and liveness.

4.6. Change

The I^2B technology supports the change (δ) through learning. The agents can be trained to respond to certain situations in a desired manner. They can learn at multiple levels of abstraction as described in Bogdanovych, Simoff, and Esteva (2008). The virtual institution structures the learning process and provides the context for learning. Through imitation the agents can learn new plans for various goals. Such plans are represented as recursive-arc graphs (similar to recursive decision trees) with probabilities being assigned to the arcs of the graph as the training continues. We have also created a method for training the I^2B agents to perform different verbal behaviour in various situations. Our method of modifying the AIML rules and assigning context to those is described in Ijaz et al. (2011).

4.7. Awareness

A virtual institution is essential in enabling the environment-, self- and interaction awareness of the I^2B agents. The institution helps the agent to understand which scene it is currently in, what is the current state of the scene, how other participants can change this state, etc. In combination with the ability to sense the surrounding objects and understand their types through annotations created by designers, the agents can include references to those objects in conversations with humans and into their decision-making. The details on integrating these features are presented in Ijaz et al. (2011).

4.8. Preparing to evaluate believability features

Having developed a general-purpose framework that supports the inclusion of the aforementioned believability features and easy mechanisms for switching these features on or off, as well as regulating the strength of each of the features has formed the ground for developing a case study where each such feature can be tested in isolation. With the help of the I^2B framework we are now able to place a number of virtual agents in a 3D simulated environment and simulate short scenarios that highlight one of the believability features. We are also able to produce two different versions of the same scenario where the surrounding environment will be identical and the only difference will be the behaviour of the agents, so that this behaviour will only differ in regards to the evaluated believability feature being enabled or disabled in the corresponding simulation. With this concept in mind we have developed a case study, presented in the next section, that was used for evaluating each of the identified believability features and measuring the degree to which each of them contributes to the overall perceived believability.

5. Case study: everyday life of the Darug people

For evaluating believability we have first built a simulation where individual agents were created via the I^2B framework. In this simulation we show everyday life of Aboriginal people from the Darug tribe, who used to live in the Parramatta basin (Australia, New South Wales) in year 1770 A.D. before the arrival of the first fleet and the establishment of the first European settlement in Australia. Each member of the tribe is represented by a virtual agent, whose model incorporates the believability features presented above. This simulation uses the aboriginal environment built for the Generations of Knowledge project (Trescak, Bogdanvych, & Simoff, 2015).

The simulation features an interactive 3D video game that takes the player on a quest to explore the life of an aboriginal clan in the Parramatta basin. A spiritual mentor and the guardian in the form of an aboriginal elder gradually introduces the participant to the daily life of aboriginal clans, the knowledge they possessed, rituals they performed, protocols they kept, etc. The elder familiarises the player with various clan members as they perform their every day activities such as tool making, painting, fishing or preparing food. During these interactions the player also learns about aboriginal medicine, astronomy, arts, as well as ceremonies, such as the smoking ceremony, and receives an introduction to their spiritual values.

Figure 1 depicts the (game) environment. The agents populating this environment show a slice of the aboriginal society. All agents are supplied with a number of internal goals and plans to reach those. For some simplistic activities, like eating it was more efficient to utilise static planning, while for others, like spear making and painting we utilised dynamic planning, so that the agents can better adjust to environment changes and interact with one another to resolve problems. The agent appearance was generated automatically using approaches from Trescak et al. (2012a). Figure 2 shows some selected agents performing their regular daily activities including collecting berries, having a feast, performing a smoke ceremony and chatting on the bank of the Parramatta river.

To give an example of the complexity of agent actions, consider the case of the spear maker. When there are no pending goals, the agent explores the forest. When it recognises the need for social interaction – it seeks for an interaction partner, approaches it and engages in a conversation. Meanwhile, the agent's hunger, thirst and fatigue levels are raising, possibly passing the threshold value, when the agent generates the goals:"hungry", "thirsty" or "tired". This tells the agent that it has to perform a specific action, such as feed, drink or rest. From its knowledge base, the agent can read all its possible actions and then use them to dynamically find a plan that leads to the goal. When a spear maker is requested

Figure 1. Simulation Environment.

Figure 2. Agents performing their daily activities.

to make a spear it dynamically constructs a chain of actions, based on the agent's current resources and the state of the environment. To make a spear, agent needs a lit fire, a spear wood and a bone axe. All these resources require further actions to obtain them, or resources to make them. Dynamic planning depends on agent capability to sense their environment and act upon these senses. For example, they can detect danger (e.g. fire). They can also perform various group activities. One of those activities is a feast. The feast is announced by the hunter, bringing a dead kangaroo.

The simulation of everyday life of the Darug people was used as a testbed and a case study to evaluate the believability features presented earlier. The aim of the evaluation was to first determine whether having each of the features present contributes to having the agents possessing the feature seen as more believable. We also wanted to investigate which of the aforementioned features would potentially increase believability more significantly than other features. To conduct such an evaluation it was important to create scenarios in which the features can be isolated and the contribution of other elements affecting believability is minimised. Having a simulation that we could control helped to make this possible as in our study we could simulate identical scenarios with identical background and identical agents and then modify only one parameter that is important for testing a particular feature.

6. Believability evaluation

There exists no formal definition of believability, nor there are clear methods to measure it. Thus, we have adapted and modified to our needs the approach in Gorman, Thurau, Bauckhage, and Humphrys (2006). The subjective nature of the approach has stimulated another aim of our work – the design of a rigorous evaluation of believable agents and calculating a believability index as a measure of their human-likeness and as a measure of comparison between different believability features.

6.1. The design of experiment

The aim of our study was to test the correctness of the following statement: Each of the identified believability features improves the overall perceived believability of virtual agents when compared with the agents who are lacking the corresponding feature. Based on this statement, we can formulate our study hypothesis as follows:

H_a: In the group that observed the virtual agent behaviour with the identified believability feature present, over 50% of the participants will consider the behaviour of the virtual agents being more believable than the behaviour of virtual agents without the corresponding feature.

In order to test this hypothesis we have conducted a study where each of the believability features was isolated and study subjects had to evaluate two scenarios (one where the agents presented on the screen possess this feature and one with agents lacking it). The study subjects had to evaluate both scenarios in terms of the believability of the virtual characters involved in those. Testing the hypothesis H_a involved verifying whether for each of the believability features the agents possessing it are perceived as more believable than those that do not have this feature.

To test the believability of our agents, we have designed an experiment and analysis technique adapting the methodology from Gorman et al. (2006). Similar to the work of Gorman et al. (2006) we have asked the users to focus on agents being perceived as human-like vs. artificial. Each of the stages of believability evaluation involved watching a video or participating in a short interactive scenario and then rating the agents present there.

6.2. Study participants

We have conducted a study with *65 participants* of diverse age and gender. The age distribution of participants was as follows: 18–24 (9.3%), 25–34 (39.5%), 35–50 (46.5%) and 50+

(4.7%). Out of all study participants 62.8% were males and 37.2% females. Our participants where recruited through online forums and social media. Some of the participants were undergraduate and postgraduate students from the Western Sydney University and were recruited on the university grounds.

6.3. Study methodology

The study has been conducted in two stages. The first stage (43 participants) analysed the characteristics identified by Loyall through an online survey .[2]

In this survey the participants were shown 15 question groups, where each group featured two videos (side-by-side) followed by questions asking which of the videos displayed character behaviour that is more artificial and which one is more believable. Next the participants were asked to evaluate the degree of this behaviour being human-like or artificial for each of the videos using the following Likert Scale: {*1:Definitely Human; 2: Probably Human; 3:Not Sure; 4:Probably Artificial; 5:Definitely Artificial*}.

The videos shown in the survey featured scenes from the Generations of Knowledge project, similar to those outlined in Figure 2. Each of the scenes tested a particular believability feature, so all agents in one of the videos possessed this feature, while all agents in the other video had this feature disabled. For example, when testing a personality aspect of believability, in one video we showed women with no distinct personalities collecting berries, while in the other video from non-verbal behaviour its clear that each of the women has one of the four classical extreme temperaments: choleric, melancholic, sanguine, and phlegmatic. We found that from the list of features presented in Loyall's work the self-motivation feature is too vague and it was difficult for us to produce a scenario for it using the people in the Darug tribe. So, for this question, instead we used the example provided by Loyall, which is a bull in one of the Walt Disney's cartoons that presumably illustrates self-motivation by blowing air through his nostrils and then running to attack. We have compared the corresponding part of this cartoon with another Disney cartoon showing a bull that does not blow air and runs to attack in response to seeing red colour. Apart from this question all other scenarios come from the simulation of Darug people. It was convenient to use our own simulation as in this case we had full control over the scenarios and could compare something that is very similar to the only difference being in the believability feature that is tested.

Based on the survey responses in regards to which video has been considered more believable we have computed the statistical significance of the obtained results. The ranking of the agents on the aforementioned Likert scale (Human vs. Artificial) helped us to compute believability indexes (described in the following section) and with their help measure how strongly each of the tested features influences believability.

The second stage of the study involved the remaining 22 participants and aimed at testing the awareness believability features. For evaluating the awareness believability we needed to have the possibility to dynamically interact with agents and converse with them, hence, presenting participants with video clips or other test mediums, as performed in Gorman et al. (2006), was not acceptable due to the issues of biased responses and guess work. To minimise both we ensured that (i) participants interact with our conversational agents in the actual virtual world; and (ii) the researcher conducting the study has no control over the routines of the agents with the flow of participants' conversations with them.

We also needed to present participants with highly immersive, engaging and interactive experience, which was essential in this study. From our past experience, we have learned that navigation in virtual worlds requires some practice. Hence, the experiments were supervised by a researcher in order to assist participants with interfacing the agents. Each participant was assisted to enter in the virtual world as an avatar. Through this avatar the participant was then requested to converse with the virtual agents, where our research assistant kept track of the conversation flow, ensuring that some of their questions were related to environment-, self- and interaction awareness of virtual agents. The assistant navigated the agent herself, while the participant was directing her where to go, whom to talk with and what to ask. This allowed the participant to focus on examining the believability of the agent's dialogue in the context of its environment, actions and the behaviour of other participants in the city. As a result the participant had to assess the believability of each dialogue on the Likert Scale: {*1:Definitely Human; 2: Probably Human; 3:Not Sure; 4:Probably Artificial; 5:Definitely Artificial*}. Depending on the topic, each of the dialogs was then associated with the corresponding label (e.g. interaction awareness, self-awareness, etc.) and then included in the statistics analysis.

In both studies the participants have evaluated the presented scenarios using the same scale. Their rating of the presented scenarios was later used for calculating the believability index as described further. The computed believability indexes were then used for comparing the contribution of each feature towards perceived believability. Apart from purely quantitative feedback participants also were able to provide qualitative feedback and explain why they ranked a particular scenario in a certain way.

6.4. Evaluating statistical significance

In order to estimate the statistical significance of the obtained results we tested the rejection of the null hypothesis (which assumed that participants could not tell the difference between the agents possessing the tested believability feature and those without it) in favour of the alternative hypothesis, which is H_a. The assumed significance level α was set at 0.05 (5% risk of incorrectly rejecting the null hypothesis).

In the first part of the study (evaluating Loyall's features from videos), the null hypothesis sample followed a normal distribution with a mean of 0.5. Here possible values were in the binary range of [0,1]. The value of 0 represented the fact that when deciding which of the videos was more believable the participant saw no difference or has incorrectly selected the agent that did not have the corresponding feature as more believable. The value of 1 in this sample corresponded to the fact that a participant was able to correctly identify the agent in possession of the tested feature as more believable. The statistical significance was then tested by comparing this sample with the resulting study sample for each of the features, where each participant's response was similarly recorded in the binary range [0, 1] following the same principle. For each of the tested believability features the resulting *P*-value has been computed and then used for determining whether this value is smaller than α and, therefore, is statistically significant.

For the second study (aware vs. un-aware agents), we had to evaluate statistical significance on the basis of fragments of interactions between test subjects and virtual agents that were ranked by participants on the 5-point Likert Scale (Human vs. Artificial). To obtain meaningful *P*-values, we have decided to treat responses with values ranging between

1 and 2 for aware agents as positive responses (1) and all others were treated as 0. For unaware agents the results in the range between 4 and 5 were considered positive responses. The study sample was compared with the normally distributed null hypothesis sample with the mean of 0.5.

The results of the statistical significance evaluation for both studies are presented in Section 6.6.

6.5. Measuring believability

The strategy we took for testing the hypothesis H_a was similar to that of Gorman et al. (2006), where believability indexes where used as a quantitative measure for a degree to which a certain simulated scenario is being perceived by test subjects as believable. With the use of believability indexes H_a can be modified as: if showing two scenarios where in the first one the agents shown possess the corresponding believability feature and in the second scenario they are lacking this feature, then we expect the resulting believability index in the first scenario B_1 to have a higher value than the believability index B_2 from the second scenario. In other words, testing H_a is reduced to testing whether $B_1 > B_2$ in all cases where B_1 corresponds to a scenario with the tested believability feature present and B_2 corresponds to the one without this feature.

For measuring believability we have modified the equations for believability index from Gorman et al. (2006) to reflect the interactive nature of some of our experiments, where the questions asked may differ across participants. Such index reflects participant's certainty with which s/he perceived a virtual agent as human-like or artificial. The equation for calculating the believability index for each study episode is shown below:

$$h_p(c_i) = \frac{|r_p(c_i) - A|}{A - B}, \tag{15}$$

where $h_p(c_i)$ is the perception of participant p of correspondence c_i as human-like and $r_p(c_i)$ is the rating of participant p for the same correspondence c_i. A and B represent the "Artificial" and "Human" value of the virtual agent response on the rating scale. Alternatively, $h_p(c_i)$ would be "0" if the respondent identified the virtual agent's response as "Artificial" or " 1" if s/he identified it as "Human", where all other values represent uncertain choices. The believability index for any participant is the average of perceptions:

$$b_n = \frac{\sum_{0<p\leq n} h_p(c_i)}{n}, \tag{16}$$

where n is the total number of responses per experiment. The overall believability index (B) in an encounter with a virtual agent, based on the rating given by m participants is

$$B = \frac{\sum b_n}{m}. \tag{17}$$

Such believability indexes were computed for every believability feature being tested. Testing the hypothesis H_a can then be reduced to testing whether the following equation holds for each of the evaluated features:

$$B_{\text{FeaturePresent}} > B_{\text{FeatureMissing}}. \tag{18}$$

Here $B_{\text{FeaturePresent}}$ represents the average score of perceived believability as rated by participants for all virtual agent encounters having the corresponding feature present, while $B_{\text{FeatureMissing}}$ is the average perceived believability for all encounters with the tested feature missing in the agent behaviour.

One benefit of using believability indexes is that they cannot only help with testing the hypothesis H_a, but can also be used as a quantifiable indicator as to how much does each feature contribute to the agents being perceived as believable. We can even use believability indexes as a measure of comparison across the tested believability features. However, in order to be able to compare the effect of each of the identified believability features it is not enough to simply compare the resulting believability indexes (B). Perceived believability of a virtual agent encounter (represented by B) depends on many factors like appearance, background, quality of animations, etc. Such elements are impossible to isolate and, thus, to eliminate their influence on the overall user assessment. Therefore, we have decided to compare believability features using the value of Δ from Equation (19).

$$\Delta = |B_{\text{FeaturePresent}} - B_{\text{FeatureMissing}}|. \tag{19}$$

In this way Δ would represent the absolute contribution of a particular feature towards perceived believability of an agent and acts as a more objective comparison measure as (given that the presence or absence of the tested feature is the main difference in the two agent encounters) it would ignore the contribution of other factors in the resulting score.

6.6. Data collection and analysis

Table 1 outlines the results of the first study (with 43 participants). Here, the first column contains the label that represents one of the believability features in the Loyall's classification. The following columns display the aggregated perceived believability for the feature being present and feature missing. The column, labelled as "*P*-value" displays the *P*-value that helps to decide on the statistical significance of the results. The column labelled as "Significance" displays the result of the statistical significance testing. Finally, the last column Δ presents the resulting comparison measure for a particular feature. The presented believability features are sorted by their value of Δ and, thus, the table displays features that make a higher contribution towards perceived believability at the top of the table and those with lower contribution at the bottom. The table is split into two sections, where the second section lists believability features associated with the illusion of life.

As seen in Table 1, some of the results were not statistically significant. In particular, the statistical significance test has failed for the following believability features: Consistency of expression, personality, social relationships, broad capability, appearance of goals, parallel action and situatedness. For these features the hypothesis H_a should be rejected. Furthermore, participant responses for testing one of the features (self-motivation) were statistically significant, but participants chose the video with the feature missing over the video with the feature present. Meaning that the H_a for this feature should be rejected as well.

Table 1. Believability comparison for Loyall's features.

Category	Feature present	Feature missing	*P*-value	Significant	Δ
Change	0.65	0.31	.016	Yes	0.34
Consistency of expression	0.57	0.40	.457	No	0.17
Emotion	0.52	0.39	.049	Yes	0.13
Personality	0.49	0.37	.457	No	0.12
Social relationships	0.52	0.41	.457	No	0.11
Self-motivation	0.27	0.35	.008	Yes, negative	0.08
Illusion of life					
Resource bounded	0.73	0.15	.002	Yes	0.58
Exist in social context	0.73	0.40	.016	Yes	0.33
Reactive and responsive	0.62	0.31	.029	Yes	0.31
Broadly capable (gestures)	0.64	0.34	.077	No	0.30
Appearance of goals	0.62	0.35	.165	No	0.27
Parallel action	0.72	0.47	.225	No	0.25
Broadly capable (voice)	0.65	0.43	.374	No	0.22
Situated	0.57	0.40	.374	No	0.17

Table 2. Believability comparison for aware and unaware agents.

Category	Unaware agents	Aware agents	*P*-value	Significant	Δ
Environment awareness	0.22	0.76	$< .001$	Yes	0.54
Self-awareness	0.26	0.75	$< .001$	Yes	0.49
Interaction awareness	0.30	0.77	.002	Yes	0.47
Overall awareness believability	0.27	0.76	$< .001$	Yes	0.49

Table 2 shows a summary of the results for testing environment-, self- and interaction awareness. Despite a smaller number of participants we were able to achieve high statistical significance for all tested features. This part of the study required a different approach as users needed to be able to dynamically interact with virtual agents rather than simply passively watch them performing in a video. Therefore, we have conducted a second study, where participants interacted with virtual agents in a gaming scenario (see Ijaz et al., 2011 for further details). This study was conducted based on text conversations with agents possessing environment-, self- and interaction awareness (further labelled as aware agents) and such interactions were contrasted with interactions with agents not having these awareness characteristics (further labelled as unaware agents). The study participants were divided into two different groups. The first group conversed with aware agents, the second group – with unaware agents. The experiments were conducted over a two-week period. After cleaning the incomplete responses, the data that were analysed included the responses of 22 participants – 11 per group.

The overall believability index for aware agents was 0.76 vs. 0.27 for unaware agents. The comparison along our awareness believability dimensions shows that for environment-aware queries in 76% of the cases participants perceived aware agent as human vs. only 22% of misclassification of unaware agents as humans. For queries about agent's own goals, plans, actions, etc. aware agents were ranked as human 75% of the times vs. 26% of misclassification of unaware agents as humans. In the case of interaction awareness, aware agents were believed to be humans for 77% cases vs. 30% of misclassification of unaware agents as humans. These results indicate that it was relatively difficult for participants to differentiate between aware agents and humans based on the restricted conversations that they had (targeting the corresponding features). The Δ column lists the resulting differences in perceived believability for each of the awareness features, allowing us to compare their overall contribution with that of the Loyall's features presented in Table 1.

6.7. Discussion of results

The results of the study suggest that from all believability features identified in this paper users found that showing that agents are resource bounded is of highest importance with $\Delta = 0.58$. In particular, virtual agents that have no visible limits to their physical strength (the aspect that was tested in our study) result in a high degree of disbelief that is higher than for any other feature that we have tested. This feature is also associated with a high statistical significance ($p = 0.002$), confirming the confidence of the test subjects in understanding its importance. If the agents are capable of interacting with human users, then the next important set of believability features (after being resource bounded) is agent awareness, meaning that agents are aware of their environment, interaction capabilities and own state in this environment. The weighted average of believability indexes for agent awareness is $\Delta = 0.49$. High average statistical significance ($p < .001$) indicates that test subjects were confident in their feedback.

The ability of agents to change their behaviour in adaptation to the changes in their environment is perceived to be the next key feature affecting believability with $\Delta = 0.34$ and $p = 0.016$ for significance. Closely behind is the ability to display that the agents exist in the correct social context of the simulation with the score of $\Delta = 0.33$ and similar level of statistical significance ($p = 0.016$). Being reactive and responsive is the next important feature with $\Delta = 0.31$ and $p = 0.0029$. The ability of virtual agents to display emotion was found to be less important than other features with $\Delta = 0.13$ and $p = 0.049$.

Unfortunately, for all other features we did not receive statistically significant results, so despite high Δ values we are unable to make reliable claims about their contribution to believability. The low statistical significance of these features could be associated with our study design, where videos that were shown to the study participants did not correctly convey the presence of the corresponding believability attributes. It might also be the case that with a higher number of study participants we would be able to increase statistical significance.

For one of the believability features described by Loyall (1997), namely self-motivation, we found evidence to suggest that this feature does not contribute to improving character believability. Loyall defines this feature in terms that can be interpreted as being pro-active rather than reactive. We found it problematic to simulate this feature within our aboriginal scenario, so in our study we used the examples of Disney animations that Loyall refers to in his work (Loyall, 1997). Users saw little difference between the two videos in regards to believability. The number of people who have correctly selected the video possessing the self-awareness feature as being more believable is only 26%, while 62% of participants responded that both presented videos are equally believable. Hence, we can say that hypothesis H_a can be rejected for this feature.

Overall, apart from self-motivation and features with low statistical significance, for all other features we have identified in this work the vast majority of test subject have correctly selected the video where characters displayed the feature in question as the one that is more believable, confirming the correctness of H_a.

6.8. Study limitations

While, to our knowledge, this is the first study of the kind that attempts to analyse individual believability features and quantify the contribution to the overall perceived believability, the results we have obtained should be treated with caution. In our case study we had a possibility to simulate identical situations with identical looking agents, placed within the identical surrounding environment and then could switch off the desired feature in one scenario and switch it on in another scenario, allowing for an accurate comparison. However, each of the features has been only tested with one scenario that we chose to be the most representative for the feature in question. To have more certainty in the correctness of our results each of the believability features requires to be exhaustively tested in a wide range of scenarios to be able to make more accurate claims in regards to its contribution towards perceived believability.

Additionally, when discovering the features that contribute to the overall believability of virtual agents we found that some of these features are either impossible to isolate or impossible to test in a way that the results are meaningful and useful. One of such features is being well integrated, identified by Loyall. Being well integrated is a very broad concept that includes pretty much every other feature being correctly interconnected with the story line, character personality, appearance and other features. We found that this feature is impossible to be tested in isolation and, therefore, did not evaluate it in our study. Additionally, we have mentioned that agent appearance significantly contributes towards believability perception of virtual agents, however, appearance is such a deep concept that testing it requires a complex dedicated study. Furthermore, appearance is known to be a subject to the uncanny valley phenomenon (Mori et al., 2012), so it was easy to obtain incorrect results in a study that tries to link appearance to believability simply because of this phenomenon and make incorrect assumptions about the reasons resulting a particular believability index.

Finally, we have tried to the best of our ability to illustrate the isolated features through scenarios that extensively demonstrate the presence or absence of the feature in question. However, we have to admit that not in all cases these scenarios were obvious for our test subjects. In some cases test subjects reported that their scoring was not due to the features that we have expected, but due to some other aspects accidentally introduced by us, like abnormally sharp turns in movement or animation being too slow, etc. This confusion has resulted low statistical significance of the results for some of the features, which would require further evaluation. On the other hand, we have to report that the analysis of open-ended responses for the cases where high statistical significance has been achieved shows that the vast majority of test subjects have correctly identified the features that were intended to be tested.

7. Conclusion and future work

We analysed existing literature in relation to believability of virtual agents. Based on this analysis we produced a revised definition of believability and a formal model. With the help of this formal model, we implemented a believability framework that can be used for simulating believable virtual agents. In building this framework, we have shown how each of the believability features identified in our study can be implemented using contemporary technologies. This framework was tested by developing a virtual reality simulation of the aboriginal Darug tribe from 1770 A.D in Western Sydney, Australia.

The identified believability features were tested through two user studies, where users evaluated the behaviour of virtual agents and compared such agents that posses the feature that is being tested with those that are missing the corresponding feature. The results of the studies suggest that many of the identified features do contribute to making virtual agents being perceived as more believable. Furthermore, by comparing believability indexes we were able to rate individual believability features by how much each of them contributes towards increasing the perception of the agents as more believable. Agents that are resource bounded, aware of their surroundings, interaction capabilities and own state in the environment, as well as agents that are capable of change, exist in the social context and are reactive and responsive were considered to be more believable than agents lacking these features.

The only feature for which we received evidence that it does not contribute to improving believability is self-motivation. Our test subjects saw no difference in the characters possessing this feature in comparison to those that do not have it. Some other features (consistency of expression, personality, social relationships, broad capability, appearance of goals, parallel action and situatedness) resulted positive responses suggesting that these might potentially contribute to increasing believability, but the statistical significance of the obtained results was low.

Future work will include running extensive studies with multiple scenarios that would help to exhaustively and more reliably test each of the believability features (especially those that resulted low statistical significance). In our next study, we plan to include more participants and will correct the glitches identified by the participants of the current study through the open-ended feedback. We will also continue investigating new aspects of believability in an attempt to identify new features and will work on addressing the limitations mentioned above.

Notes

1. http://unity3d.com
2. https://goo.gl/s1vP1L

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bartneck, C. (2002). *Integrating the OCC model of emotions in embodied characters*. In: Workshop on Virtual Conversational Characters, Citeseer.

Bergmann, K., Eyssel, F., & Kopp, S. (September 2012). *A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time*. In Yukiko Nakano, Michael Neff, Ana Paiva, & Marilyn Walker (Eds.), IVA'12: Proceedings of the 12th international conference on intelligent virtual agents (pp. 126–138). Berlin: Springer-Verlag.

Bickmore, T. W., Mauer, D., & Brown, T. (2007). *Context awareness in mobile relational agents*. In Catherine Pelachaud, Jean-Claude Martin, Elisabeth André, Gérard Chollet, Kostas Karpouzis, & Danielle

Pelé (Eds.), IVA '07: Proceedings of the 7th international conference on intelligent virtual agents (pp. 354–355). Berlin: Springer-Verlag.

Bogdanovych, A. (2007). *Virtual institutions* (PhD thesis), UTS, Sydney, Australia.

Bogdanovych, A., Simoff, S., & Esteva, M. (2008). *Virtual institutions: Normative environments facilitating imitation learning in virtual agents*. In Helmut Prendinger, James Lester, & Mitsuru Ishizuka (Eds.), Proceedings of the 8th international conference on intelligent virtual agents (IVA 2008). Lecture notes for computer science (pp. 456–464). Berlin: Springer.

Cunningham, D. W., Kleiner, M., Wallraven, C., & Bülthoff, H. H. (2005). Manipulating video sequences to determine the components of conversational facial expressions. *ACM Transaction on Applied Perception*, 2(3), 251–269.

Dautenhahn, K., Ogden, B., & Quick, T. (2003). *From embodied to socially embedded agents – implications for interaction-aware robots*.

Doyle, P. (2002). *Believability through context using "knowledge in the world" to create intelligent characters*. In Burg, B., Dale, J., Finin, T., Nakashima, H., Padgham, L., Sierra, C., & Willmott, S. (Eds.), AAMAS '02: Proceedings of the first international joint conference on autonomous agents and multiagent systems (pp. 342–349). New York, NY: ACM.

Egges, A., Kshirsagar, S., & Magnenat-Thalmann, N. (2003). *A model for personality and emotion simulation*. In: Knowledge-based intelligent information & engineering systems (KES 2003) (pp. 453–461). Berlin: Springer.

Elpidorou, A. (2010). Alva noë: Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness. *Minds Machines*, 20(1), 155–159.

Gandhe, S., & Traum, D. (2007). *Creating spoken dialogue characters from corpora without annotations*. In: Proceedings of Interspeech-07. Antwerp, Belgium, pp. 2201–2204.

Gerhard, M., Moore, D., & Hobbs, D. (2004). Embodiment and copresence in collaborative interfaces. *International Journal of Human–Computer Studies*, 61(4), 453–480.

Gerhard, M., Moore, D., & Hobbs, D. (2005). Close encounters of the virtual kind: Agents simulating copresence. *Applied Artificial Intelligence*, 19(3–4), 393–412.

Gorman, B., Thurau, C., Bauckhage, C., & Humphrys, M. (2006). *Believability Testing and Bayesian Imitation in Interactive Computer Games*. In Nolfi, S., Baldassare, G., Calabretta, R., Hallam, J., Marocco, D., Miglino, O., Meyer, J.-A., & Parisi, D. (Eds.), Proceedings of the 9th international conference on the simulation of adaptive behavior (SAB'06) (pp. 655–666), Volume LNAI 4095. Rome: Springer.

Hallowell, A. I. (1955). *Culture and experience*. Philadelphia: University of Pennsylvania Press.

Hartmann, B., Mancini, M., & Pelachaud, C. (2005). *Implementing expressive gesture synthesis for embodied conversational agents*. In: Gesture Workshop,188–199.

Heylen, D. K. J., van Es, I., Nijholt, A., & van Dijk, E. M. A. G. (2005). *Controlling the gaze of conversational agents*. In van Kuppevelt, J., Dybkjaer, L., & Bernsen, N.O. (Eds.), *Natural,intelligent and effective interaction in multimodal dialogue systems* (pp. 245–262). Berlin: Kluwer Academic Publishers.

Howard, P. J., & Howard, J. M. (1995). *The big five quickstart: An introduction to the five-factor model of personality for human resource professionals*. Retrieved from http://eric.ed.gov/?id = ED384754.

Ijaz, K., Bogdanovych, A., & Simoff, S. (2011). *Enhancing the believability of embodied conversational agents through environment-, self- and interaction awareness*. In Reynolds, M., (ed.), *Australasian computer science conference (ACSC 2011)* (pp. 107–116). Volume 113 of CRPIT, ACS: Perth.

John, O. P., Donahue, E., & Kentle, R. (1990). The 'big five'. Factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In Lawrence A. Pervin & Oliver P. John (Eds.)., *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.

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

Kelley, H. H. (1950). The Warm-Cold Variable in first impressions of Persons. *Journal of Personality*, 18(4), 431–439.

Kenny, P., Parsons, T. D., Gratch, J., & Rizzo, A. A. (2008). Evaluation of Justina: A virtual patient with ptsd. In Prendinger, Helmut, James Lester, & Ishizuka, Mitsuru (Eds.), *Intelligent virtual agents* (pp. 394–408). Berlin: Springer.

Lee, J., Marsella, S., Traum, D., Gratch, J., & Lance, B. (2007). *The rickel gaze model: A window on the mind of a virtual human*. In Pelachaud, C., Martin, J.C., Andr, E., Chollet, G., Karpouzis, K., & Pel, D. (Eds.), *Intelligent virtual agents*. Volume 4722 of lecture notes in computer science (pp. 296–303). Berlin: Springer.

Lester, J. C., & Stone, B. A. (1997). *Increasing believability in animated pedagogical agents*. In Joerg Mueller, Munindar P. Singh, & Anand S. Rao (Eds.), Proceedings of the first international conference on Autonomous agents (AGENTS '97) (pp. 16–21). New York, NY: ACM.

Lester, J. C., Voerman, J. L., Towns, S. G., & Callaway, C. B. (1999). Deictic believability: Coordinating gesture, locomotion, and speech in lifelike pedagogical agents. *Applied Artificial Intelligence*, 13, 383–414.

Livingstone, D. (2006). Turing's test and believable AI in games. *Computer Entertainment*, 4(1), 1–13.

Loyall, A. B. (1997). *Believable agents: building interactive personalities*. PhD thesis, Computer Science Department, Pittsburgh, PA, USA.

Mateas, M. (1999). *Artificial intelligence today*. Berlin: Springer-Verlag, pp. 297–328.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *Robotics & Automation Magazine, IEEE*, 19(2), 98–100.

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a cognitive map*. Oxford: Oxford University Press.

Ortony, A. (2003). *On making believable emotional agents believable*. Cambridge, MA: MIT Press.

Ortony, A., Clore, G., & Collins, A. (1988). *Cognitive structure of emotions*. Cambridge, MA: Cambridge University Press.

Prendinger, H., & Ishizuka, M. (2001). Social role awareness in animated agents. In: *AGENTS '01: Proceedings of the fifth international conference on autonomous agents* (pp. 270–277), ACM: New York, NY.

Rao, A. S., & Georgeff, M. P. (1995). *BDI agents: From theory to practice*. In: ICMAS. Volume 95, 312–319.

Riedl, M. O., & Stern, A. (2006). *Failing believably: Toward drama management with autonomous actors in interactive narratives*. In: TIDSE. 195–206.

Russel, S., & Norvig, P. (2003). *Artificial intelligence a modern approach*. Upper Saddle River, NJ: Prentice-Hall. Pearson Education International.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379–00.

Tarjan, R. (1972). Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2), 146–160.

Tencé, F., Buche, C., Loor, P. D., & Marc, O. (2010). *The challenge of believability in video games: Definitions, agents models and imitation learning*. CoRR abs/1009.0451.

Thiebaux, M., Lance, B., & Marsella, S. (2009). *Real-time expressive gaze animation for virtual humans*. In: AAMAS (1). 321–328.

Thomas, F., & Johnston, O. (1981). *Disney animation : The illusion of life* (1st ed.). New York, NY : Abbeville Press.

Trescak, T., Bogdanovych, A., Simoff, S., & Rodriguez, I. (2012a). *Generating diverse ethnic groups with genetic algorithms*. In: VRST '12: Proceedings of the 18th ACM symposium on Virtual reality software and technology, ACM Request Permissions.

Trescak, T., Bogdanovych, A., Simoff, S., & Rodriguez, I. (2012b). *Generating diverse ethnic groups with genetic algorithms*. In: Proceedings of the 18th ACM symposium on virtual reality software and technology (VRST '12), New York, NY, USA, pp. 1–8.

Trescak, T., Bogdanvych, A., & Simoff, S. (2015). *The aboriginal dreaming meets virtual reality*. In: Proceedings of the annual conference of the Alliance of Digital Humanities Organizations.

Wallace, R. (2004). *The Anatomy of A.L.I.C.E.* A.L.I.C.E. Artificial Intelligence Foundation.

Weizenbaum, J. (1966). Eliza – a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45.