



Toward Believable Acting for Autonomous Animated Characters

Cassidy Curtis
cassidycurtis@google.com
Google Research
USA

Peter F McDermott
petermcd@google.com
Google Research
USA

Alonso Martinez
martinezalonso@google.com
Google Research
USA

Tianyu Liu
tianyuliuchn@gmail.com
Google Research
USA

Sigurdur Orn Adalgeirsson
siggiorn@google.com
Google Research
USA

JD Velásquez
jdvelasquez@google.com
Google Research
USA

Dei Gaztelumendi
dei@deimorph.com
Google Research
USA

Horia Stefan Ciurdar
horia@google.com
Google Research
USA

W. Bradley Knox
bradknox@google.com
Google Research
USA

Norberto Adrian Goussies
norbertogoussies@hotmail.com
Google Research
USA

Palash Nandy
palash@google.com
Google Research
USA

ABSTRACT

This paper describes design principles and a system, based on reinforcement learning and procedural animation, to create an autonomous character capable of believable acting—exhibiting a responsive and expressive illusion of interactive life, grounded in its subjective experience of its world. The design principles incorporate knowledge from animation, human-computer interaction, and psychology, articulating guidelines that, when followed, support a viewer’s suspension of disbelief. The system’s reinforcement learning brain generates action, emotion, and attention signals based on motivational drives, and its procedural animation system translates those signals into expressive biophysical movement in real time. We demonstrate the system on a stylized quadruped character in a virtual habitat. In a user study, participants rated the character favorably on animacy and ability to experience emotions, which is consistent with finding the character believable.

CCS CONCEPTS

- Computing methodologies → Procedural animation; Reinforcement learning; Intelligent agents.

KEYWORDS

believability, acting, autonomous characters, synthetic characters, animation, reinforcement learning



This work is licensed under a Creative Commons Attribution International 4.0 License.

MIG '22, November 3–5, 2022, Guanajuato, Mexico
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9888-6/22/11.
<https://doi.org/10.1145/3561975.3562941>

ACM Reference Format:

Cassidy Curtis, Sigurdur Orn Adalgeirsson, Horia Stefan Ciurdar, Peter F McDermott, JD Velásquez, W. Bradley Knox, Alonso Martinez, Dei Gaztelumendi, Norberto Adrian Goussies, Tianyu Liu, and Palash Nandy. 2022. Toward Believable Acting for Autonomous Animated Characters. In *ACM SIGGRAPH Conference on Motion, Interaction and Games (MIG '22)*, November 3–5, 2022, Guanajuato, Mexico. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3561975.3562941>

1 INTRODUCTION

*Suspension of disbelief*¹ is a remarkable human gift: it allows a viewer to believe in a fictional character despite knowing that it is not real. It unlocks the viewer’s ability to be fully immersed in the story, relate to the character’s point of view, and feel empathy for its plight. A key requirement is that the fiction be internally consistent—particularly, the character’s behavior must be believable within the rules of its world.

To create a performance that sustains the viewer’s suspension of disbelief requires simultaneous command of the physical and psychological aspects of the character’s behavior, in a way that reveals its inner life, grounded in its relationship to its past, future, fellow characters, and surroundings. The craft of creating such a performance is known as *acting* [Strasberg and Chaillet 2021].

A believable autonomous character—one that you can relate to and empathize with as easily as any other fictional character—has been a dream for decades. But in interactive media like games, such characters remain rare. Standard game animation techniques require developers to specify a stimulus-response mapping in advance, and it is impractical to prepare a bespoke response for every combination of contingencies a character could possibly encounter. Thus game characters tend to alternate between two modes: full interactive autonomy with limited acting ability, and non-interactive clips (or cutscenes) featuring more nuanced acting captured from

¹Sometimes also characterized as *secondary belief* [Tolkien and Tolkien 1984].

live actors or hand-authored by animators. But seldom are they autonomous and believable at the same time.

Our goal in this work is to enable an autonomous character to act believably throughout an entire interactive experience. Specifically, this paper offers the following novel contributions:

- (1) A set of design principles that contribute to the believability of an autonomous character’s behavior, incorporating knowledge from the fields of psychology, animation, and human-computer interaction.
- (2) A system that illustrates how one can generate behavior that demonstrates most of these principles, autonomously and in real time, by using internally motivated reinforcement learning to produce action, emotion and attention signals that are coherent and grounded in the character’s subjective experience. When expressed through a layered procedural animation system², these signals provide the observable evidence needed to sustain the viewer’s suspension of disbelief.

We demonstrate our system using a simple non-speaking animal character that allows us to illustrate the principles with maximum clarity. Finally, we validate our system through a user study.

2 RELATED WORK

Acting has been a topic of interest in computer animation for decades, notably in the work of Ken Perlin and colleagues, who developed a variety of procedural animation methods for expressing emotion and personality through movement [Perlin 2003; Perlin and Goldberg 1996; Perlin and Seidman 2008]. Although these “digital actors” required a human user to direct their actions and emotions using authoring tools, these papers envisioned a future in which such characters could act autonomously. Our system is a step toward realizing this shared vision: it is truly autonomous, creating its own emergent performance spontaneously in real time.

Our system sits in the context of a long history of experimental autonomous beings. Braintenberg [1986] showed that even simple vehicles can engender the illusion of emotions and desires through their mechanically deterministic behavior. Bates et al. [1994] proposed that a believable animated character could be achieved by combining knowledge from artificial intelligence and animation. More recently, BabyX [Sagar et al. 2014] combined neural models of perception and emotion with a photorealistic human face rig to plausibly simulate certain interactions with an infant. Our character could be considered a kind of *animat* [Wilson 1991], in that its behavior is emergent from basic motivations.

The term *believability* is used variously in different fields of study. In interactive narrative, it generally refers to script writing, describing a character’s decisions only at an abstract or cognitive level [Shirvani and Ware 2019]. In game research, it most often refers to gameplay: the player’s sense of immersion in a game experience [Warpefelt 2016], or a player’s perception of the humanness of NPCs [Umarov and Mozgovoy 2012]. In contrast, our work is concerned with the moment-to-moment believability of a character’s acting.

Our principles of believability (see Section 3.1) extend and refine Gomes et al.’s (2013) *dimensions of believability*, a list of descriptive attributes meant to be used as a metric for evaluating interactive

characters. To our knowledge that metric has never been validated as an experimental instrument or tested in its entirety, although parts of it have been used to evaluate simple robots [Marmpena 2021] and minimal virtual game agents [Rosenkind 2015]. We apply our list not as an evaluation rubric, but as design principles to guide character development.

There are some notable examples of interactive characters in games that exhibit many of the principles of believability: Trico from *Last Guardian* [Cooper 2018] and the horses in *Red Dead Redemption II* [Kleanthous 2021] are both highly convincing depictions of living animals, which show some degree of autonomy even while the player is riding them. The character Lucy from the virtual reality animated short *Wolves in the Walls* [Billington and Shamash 2018], while not truly autonomous, has moments of contingent interaction when she uses eye contact and physical interaction to establish rapport with the viewer.

Machine learning has been used with great success to solve challenging procedural animation problems like physically-driven locomotion [Liu and Hodgins 2017; Tao et al. 2022], object perception and muscle control [Nakada et al. 2018], and gaze kinematics [Klein et al. 2019]. In contrast, we apply reinforcement learning to the character’s acting (specifically its choice of actions, attention targets, and emotions), and address its locomotion using relatively simple heuristic methods, inspired by games like *Spore* [Hecker et al. 2008], *Overgrowth* [Rosen 2014] and *Rain World* [Jakobsson and Therrien 2016], which share a similarly minimalist layered approach.

Driving character behavior through models of motivation has been used in AI to study the cognitive architectures of the mind [Bach 2012; McCall et al. 2020], to enable more efficient exploration of an RL environment [Bellemare et al. 2016; Pathak et al. 2017], to generate discrete action policies for NPCs [Forgette and Katchabaw 2014; Merrick and Maher 2007], and to model player behavior in games for testing purposes [Roohi et al. 2018]. Our system uses motivation both for driving behavior, as well as for grounded emotional expression, with more emotional specificity and nuance than previous work.

To determine our character’s emotions, its RL-based brain creates authentic signals (i.e. grounded in things that matter to it) that indicate emotion upon their appraisal (see Section 5). Past work exists on emotion in ML [Strömfelt et al. 2017] and specifically in RL agents, as surveyed by Moerland et al. [2018]. As the survey outlines, research on emotion in RL differs based on whether the emotion affects the learning process or not; whether emotions can be affected by external events (e.g. eating or falling) and by signals internal to the agent (e.g. state novelty); and how RL-related quantities such as the state values, rewards, and transition probabilities are mapped to emotion concepts, if at all. Our approach to deriving emotion perhaps most closely resembles that of Jacobs et al. [2014] which maps the agent’s estimated state value to a hope and fear emotional dimension and maps other RL-based statistics to a joy and distress dimension. Notable deviations of our system from [Jacobs et al. 2014] are detailed in Section 5.

²Not itself a novel contribution, but described in Section 6 for completeness.

3 BELIEVABILITY

Believability depends on both character and observer. Specifically, we define a character as *believable* if its observed behavior meets or exceeds the viewer's expectations. Those expectations depend on the character's visual design, the viewer's prior beliefs, and the viewer's understanding of the rules of the character's world. For example, a photorealistic human character might reasonably be expected to speak, wear clothing, and, if injured, suffer lasting harm. A cartoon coyote may quite believably do none of these things.

The difficulty of sustaining believability depends greatly on these expectations. One can easily imagine a being so simple—such as a bacterium or a blade of grass—that portraying it believably would be trivial. But in this work we are specifically interested in the type of character that is *relatable*, i.e. one that sets expectations of having an inner life with intentions, desires, and emotions. Our principles strive to describe what's needed for a relatable character to sustain the viewer's suspension of disbelief.

Believability achieved can be lost at any time, and therefore must be actively maintained moment by moment. An action that is believable once might not be believable if repeated, or in the context of a longer sequence of events. Thus, it is important to consider a character's believability across all time scales, from a fraction of a second to the duration of an entire experience.

3.1 Principles of believability

An acting performance operates on many levels simultaneously, from its physical qualities to its psychological ones. Thus we have found it helpful to break down the monolithic concept of believability into a set of more specific design principles, each of which is more tractable to analysis and simulation. This list is tailored to the domain of animated acting, combining insights from the fields of psychology, animation, and human-computer interaction, and in particular draws inspiration from Gomes et al. [2013], who also decomposed believability into a set of individual attributes. (For a detailed comparison, see Appendix A.)

For a character's acting performance to be believable, it should demonstrate the following principles:

- (1) **Coherence of identity:** *maintains consistent characteristics over time so as to appear to have a single identity.* A common violation of this principle is to allow an entirely different algorithm to take over behavior in certain situations, such as when an NPC abruptly delivers pre-scripted dialog, or a toy robot's persona is replaced by a virtual assistant.
- (2) **Physical movement:** *obeys physical laws consistent with the rules of its world.* Note that the rules of the character's world may differ from the rules of our world. This is acceptable so long as those rules are internally consistent.
- (3) **Biological movement:** *moves as if supported by a biological structure such as rigid bones with flexible joints and muscles.* The human visual perception system is primed to recognize biological movement [Johansson 1973] which makes this quality especially important for animated characters. Animation principles such as *squash and stretch, arcs, slow in and out, follow through and overlapping action* [Thomas and Johnston 1981] all exist primarily to support the perception of biological movement.

- (4) **Self-propulsion:** *appears to move on its own initiative, rather than only being pushed or pulled by external forces.* Self-propelled entities are perceived by infants as agents, and their actions are perceived as goal-directed [Luo and Bailargeon 2005]. A common violation of this principle is a marionette whose movements are obviously controlled by a puppeteer pulling strings. Skillful puppetry involves creating an illusion of self-propulsion so compelling that the viewer forgets the puppeteer and strings exist.
- (5) **Contingent interaction:** *reacts in a timely way to events in the world.* Humans, even infants, attribute mental states, particularly perception and goal-directedness, to objects if they react contingently to their environment [Johnson and Ma 2005]. A common violation of this principle is when an autonomous character plays back a fixed animation clip for some time period, during which it cannot respond to interruptions of any kind.
- (6) **Self-motivation:** *appears to act on its own internal motivations, and pursue goals to satisfy its needs.* Goal-oriented behavior is recognized in psychology as a precursor for perception of agency and animacy, even in infants [Gergely and Csibra 2003]. More complex characters should also be capable of concurrent pursuit of multiple goals [Loyall 1997].
- (7) **Attention:** *appears to pay attention to relevant entities in the world around it.* This includes *attention appropriateness*, i.e. that the object of the character's attention is something relevant to it, and *attention legibility*, that this fact is displayed in a way that the viewer can perceive. Attention provides a powerful tool to communicate a character's inner state to the viewer—what it can perceive, is thinking about, or considers important. Timely changes in attention also provide an opportunity to demonstrate *contingent interaction*.
- (8) **Emotion:** *clearly conveys emotions appropriate to the viewer's understanding of its circumstances, motivations, and goals.* This too requires both *appropriateness* and *legibility*. Note that multiple aspects of a character's behavior can convey information about its emotional state: its choice of actions (or inaction), modulations of those actions (e.g. speed of movement), or explicit display of emotional affect (e.g. posture or facial expression).
- (9) **Explainability:** *behaves in a way that allows the viewer to form an explanation consistent with the observed evidence.* This principle applies to a performance taken as a whole: it is not necessary that every action be completely explainable the moment it occurs—in fact, the mystery can deepen the viewer's engagement in the experience—as long as the viewer can retroactively explain it later [Gaver et al. 2003]. For this principle the timing and sequencing of events is important: for example, an action will be more explainable if preceded by a relevant shift in attention. A common violation of believability occurs when a character reacts to an event that it didn't perceive.
- (10) **Thought:** *displays evidence of thinking in a biologically plausible way.* By *thinking* we mean activity that is purely mental, such as remembering, considering, understanding, planning or deciding. Overtly portraying such thought processes helps convey a sense of inner life, and is especially important for

characters from which the viewer expects any degree of intelligence. Thinking can also affect the character's attention and emotions: a character that thinks may attend to its thoughts, and show changes in emotion contingent on those thoughts. The timing of such changes should reflect the viewer's expectation of how long they would naturally take. A software-based character can violate this principle in either direction, e.g. by reacting instantly to stimuli that it should need some time to understand, or by pausing for an implausibly long duration to plan movement.

- (11) **Sociality:** *senses and acts on inferences about the beliefs, desires, and intentions of other beings.* Humans are profoundly social creatures and reason about other beings with an intentional model [Lieberman 2013]. We therefore expect a relatable character to reason and act similarly towards us or other agents in its world.
- (12) **Personality:** *has unique behavioral and expressive traits that distinguish it from its peers as an individual.* The dramatic arts, particularly writing and acting, stress the importance of unique behavior to prevent the character from being seen as ‘stock’ or unconvincing—and designers of believable characters would be well served to think about opportunities to express uniqueness even in the details of movement [Loyall 1997].
- (13) **Change with experience:** *behaves in a way that shows the effect of its past experiences.* By change with experience we do not mean merely acquiring information about the world, but rather change in behavior patterns due to experiences relevant to the character’s motivations. Such behavior change suggests that the experience was meaningful to the character and creates room to read emotions into the experience. Even if the viewer did not witness the event that caused the change, the behavior itself is often sufficient. Artificial agents that change from experience are hypothesized as more life-like and shown to engender greater empathy in humans [Darling et al. 2015].

This list of principles is not a recipe, but rather a set of tools for guiding one’s choices when designing and animating an autonomous character. The principles do not all apply equally to every character: each one is only necessary to the extent that the viewer’s expectations require it. For example, a mechanical robot might set low expectations for *biological movement*, and a photo-realistic squirrel may be perfectly believable without evidence of *thought*.

4 ILLUSTRATIVE IMPLEMENTATION

To illustrate the importance of these principles, and to show by example how they can be put to use, we have built a system to demonstrate them using reinforcement learning (RL) and procedural animation.

While we believe the principles to be applicable to characters at all levels of complexity, most of them can be demonstrated using a simple animal character endowed with basic, universally relatable needs like safety, nourishment, and rest.

An intentional outcome of our approach is that the character’s attention, emotions, and action policy all emerge from this same

set of underlying motivations, grounding every aspect of the character’s behavior in its sensorimotor experience of its world. This is important for believability and particularly *explainability*, because the viewer will intuitively seek the simplest logical explanation for the character’s observed behavior. An effective way to ensure that the behavior has a simple explanation is to derive every aspect of it from a single source of truth.

Of the thirteen principles, our implementation focuses on the first nine. Although our system can appear to demonstrate the remaining principles in brief evaluations, to produce such behavior robustly over long time periods is an area for future work (see Section 10).

4.1 System overview

Our system is comprised of two main parts: a *brain* trained with RL to choose a desirable course of action (Section 5) and a procedurally animated *body* which produces a continuous stream of expressive, biophysical movement (Section 6). The brain consists of a drive-based intrinsic reward system, an action policy and set of value functions produced by RL, and emotion and attention modules that rely on the policy and value functions. The body provides the brain with observations about its environment, and the brain sends command signals to the body (see Figure 1). This cycle of observation and command occurs at a frequency of 5Hz, analogous to the traditional animation timing scheme of a handful of “beats” per second [Thomas and Johnston 1981].

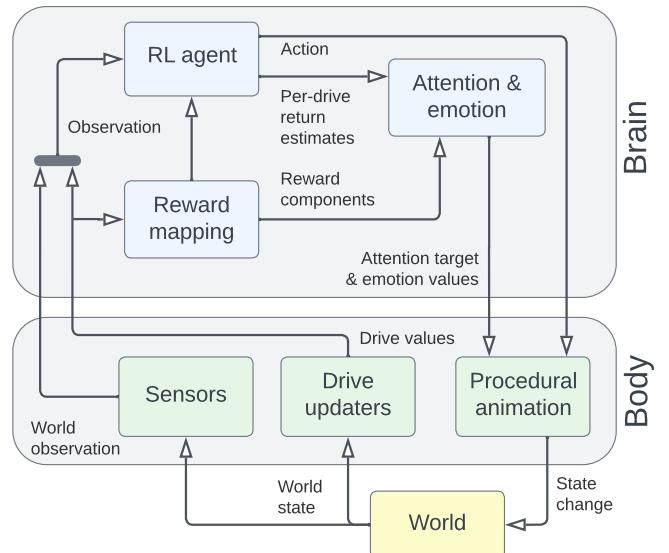


Figure 1: Diagram showing the major components and signal pathways in our system.

4.2 Environment

Our character lives in a simple 3D landscape (Figure 2) with gentle hills, treacherous cliffs, and rocks and other obstacles that can obstruct its view and movement. Fruit grows on trees and bushes, falls from them, and rolls down the hillsides.

Our character's world also contains *critters*, which are simple pest-like creatures programmed to compete for food and attack or flee when threatened, and *the user*, who can help or harm the character by petting, feeding, or poking it.



Figure 2: Our character, Axo (a), seen here in its native habitat, competing with a few hostile critters (b) for a fruit (c).

4.3 Character design

A character's visual design sets the viewer's expectations for its capabilities. If the character resembles a familiar species, it activates prior beliefs about similar animals. Also, more detailed designs set higher expectations of realistic movement. Humans excel at filling in missing details: viewers can perceive a complete walking figure from just a handful of moving dots [Johansson 1973]. However, viewers are far less tolerant of *conflicting* information [Chaminade et al. 2007; Saygin et al. 2012]. Thus, the more complex the character design, the greater the risk of breaking the viewer's suspension of disbelief.

For these reasons, we opted for a simplified, stylized visual design. Our character, Axo, is a quadruped made of primitive shapes, with rudimentary features designed to convey only the most basic information through posture, movement and expression (see Figures 2 and 3). This ensures that all information shown to the viewer is entirely under our system's control, and sets the viewer's expectations appropriately, increasing the likelihood that the character's behavior will meet or exceed them.

5 REINFORCEMENT LEARNING BRAIN

In support of the believability principles, we chose to implement our character's brain using RL (see Figure 1). The objective of RL is to maximize the accumulation of reward, and many of its algorithms pursue this objective by creating value functions that estimate how much reward can be accumulated from a given state on average [Sutton and Barto 2018]. This approach naturally aligns with several of our believability principles in the following ways: We construct the reward signal such that it only depends on the character's motivational drives (see Section 5.1). This ensures that a well-trained behavior policy will optimize actions to achieve desired

motivational states, supporting the principle of *self-motivation*. The reward is also aligned with the character's immediate emotional state, as it describes the current state's desirability with respect to each motivational drive. The value estimates offer predictions of how well the character's motivational needs will be met in the future, which also provides rich signals for anticipatory emotion and attention, supporting the corresponding believability principles. Since the behavior, emotion, and attention systems are all driven by the same motivational drives, they are constrained to be coherent, to be authentically grounded in the character's experience, and to require minimal ad-hoc machinery [Wilson 1991].

We also want characters to exhibit complex and interesting behavior that emerges from the interplay and conflict between simple components (i.e. the motivational drives), as well as between the constraints of the character's embodiment and its interaction with its environment. RL inherently generates this emergent behavior to in order to maximize the cumulative rewards. Further, by using a general learning method to generate behavior, the brain is somewhat agnostic to the specific environment or character design, allowing adaptation to changes in either by retraining instead of reprogramming. Lastly, RL has the ability to naturally support the *change with experience* principle, though its use for that purpose is future work (see Section 10).

5.1 Reward design

An important step in the design of a believable character is choosing what should matter to it, and therefore what should drive its behavior. This corresponds to the principle of *self-motivation*. Inspired by Maslow's Hierarchy of Needs [Maslow 1943], we defined a fundamental set of *motivational drives*, closely related to the Hullian drives described by Konidaris and Barto [2006]. These drives organize the character's appraisal of events and states of its world. Specifically, the reward we designed for RL to maximize is solely dependent on the values of these motivational drives, making them the foundation of the character's behavior.

Our design is inspired by a refinement to the classic agent/environment RL paradigm that was introduced by Barto [2013, Figure 2], in which an "organism" contains an RL agent and an internal environment, and this internal environment interacts with an external environment. In this paradigm of *intrinsically motivated RL*, reward is produced as a combination of internal and external signals and events. In our implementation, the internal environment includes both the body layer and the brain layer (excluding for the RL agent itself) in Figure 1, and the external environment is the world.

Our current system includes three basic drives: nourishment, safety, and rest. We chose these drives because of how relatable and legible they are to the user in a simple environment, as well as for how the interplay and conflict between them can result in complex, interesting emergent behavior and emotional expression.

We define D to be the set of all drives. The mechanics of each drive $d \in D$, described as a tuple (p_d, α_d, U_d) , determine how to calculate a corresponding drive value $d_t \in [0, 1]$ at time t . In the absence of relevant stimuli, a drive value slowly regresses to its set-point p_d by increasing or decreasing by α_d at each time step. Drive values are also affected by events in the environment (e.g.

nourishment increases upon eating) according to each functional rule u in the rule set U_d for that drive (see detailed instantiations of rules in Appendix C.1). The update of a given drive value d_t at time t , as a function of the state s and action taken a , is defined as follows:

$$d_t = d_{t-1} + \alpha_d * \text{sign}(p_d - d_{t-1}) + \sum_{u \in U_d} u(s, a)$$

Each drive also defines a function w_d that maps the drive value to a reward component value. We then define the reward signal at time t , r_t , to simply be the sum of these reward component values, one per motivational drive:

$$r_t = \sum_{d \in D} w_d(d_t)$$

In our current system, each w_d is a strictly increasing piecewise linear function, which allows us to define the change in reward to be less sensitive to change in drive values in some ranges, and more so in others (for example, to model that the character should be more severely punished for very low nourishment values when it is very hungry), see detailed instantiations of reward mappings in Appendix C.2.

In summary, a character can increase or decrease its reward by acting to influence the probability of various drive-relevant events occurring, which in turn increases or decreases its drive values. These changes in drive values change the character's reward.

5.2 Emotions and attention

To demonstrate the principle of *emotion appropriateness*, the character's displayed emotions must be grounded in its current subjective experience. And to demonstrate *attention appropriateness*, it must observably attend to whatever is most relevant to its current emotional state and behavior. When juxtaposed against the character's choice of actions, these two observable details combine to support the principle of *explainability* (see Figure 3). Thus, in addition to sending action commands to the body, the brain must also produce emotion and attention signals.

The emotion system generates three types of emotions. *Immediate* emotions reflect the current state of each motivation drive and its reward component. *Anticipatory* emotions predict future reward for each motivational drive (e.g. *safety anticipation* will be low when danger is near). Lastly, *surprise* emotions are generated from the derivatives of the anticipatory emotions, enabling the character to react to sudden changes in its prospects. For a detailed description of how these signals are generated, see Appendix E.

Prediction of future reward, or value, has previously been used to provide anticipatory emotion signals [Jacobs et al. 2014], but structuring our reward as a sum of motivational reward components enables improvements upon such past work. Using the same data that is generated to train a behavior policy, the brain concurrently trains a drive-decomposed value function to predict cumulative future reward for each individual motivational component. This drive-wise prediction provides more specific signals for anticipatory emotional expression than would a single value function.

The attention system identifies the most emotionally salient object for each drive by first iterating through all visible objects and

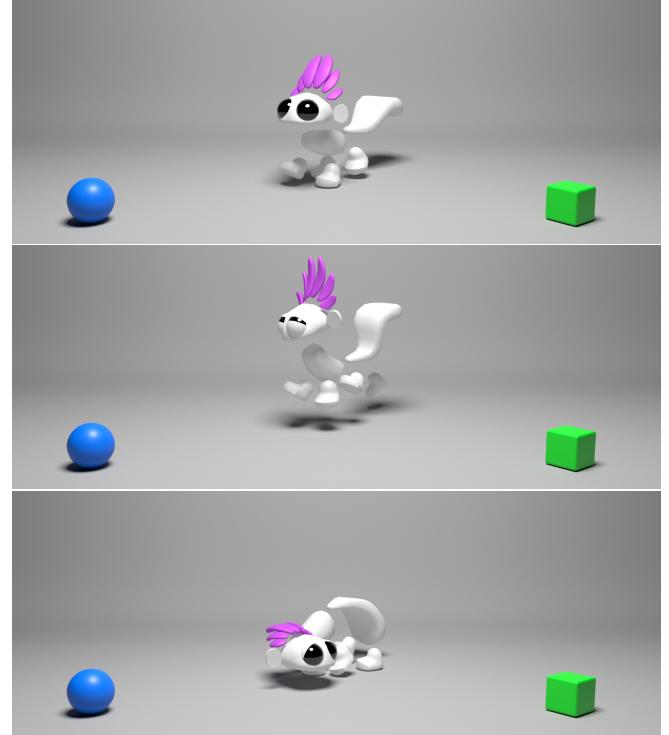


Figure 3: Faced with a choice between two objects, Axo decides to move to the left. Without any additional information, the viewer cannot explain why Axo made this choice (top). With the visual cues of attention and emotion, the viewer can clearly distinguish between the case where Axo is excited about the blue sphere (middle), versus fearful of the green cube (bottom). These cues enable the viewer to explain the behavior, which also endows the objects with meaning.

evaluating the drive-decomposed value function with a counterfactually generated observation where the corresponding object is removed. The object that, when removed, results in the largest change becomes the referent for the corresponding anticipatory emotion. The referent with the largest change is then chosen as the character's attention target.

5.3 User modeling

Following the principles of *sociality* and *change with experience*, we designed some opportunities for basic interaction between our character and the user.

We detect the user's presence and head position using the device's built-in camera and mirror them with an avatar at the edge of the environment. This places the user in the virtual space, as if the screen was their first-person view.

The user is thus observable by and can interact with the character, either positively by petting and feeding it, or negatively by poking it, via touchscreen gestures. Besides their immediate impact on the character's drives (feeding increases nourishment, poking decreases safety, etc.), the system tracks (and includes in the observation) a scalar statistic of these interactions, called the user score. This is a

heuristically estimated quantity, based on occurrences of negative and positive interactions. This allows training the RL policy with simulated users, each assigned and behaving according to a fixed score from a wide range of potential scores. This prepares the character to adapt its behavior to the changing score of the real user at runtime.

Despite its simplicity, this approach was sufficient to engender basic reasoning about the user and for primitive social behaviors to emerge (see Section 7).

5.4 Observations and actions

The RL observation vector consists of both information about the character’s internal state such as the values of its motivational drives, and sensed elements of the environment such as relative positions of visible objects as well as navigability of the terrain. For more detail see Appendix B.

The RL policy outputs an action vector of six real-valued numbers. The first two elements are the chosen *forward* and *turn* accelerations. The remaining four values represent discrete action commands (*grab*, *drop*, *eat*, and *roar*). The discrete action corresponding to the highest value is selected, unless none exceed a given threshold, in which case no discrete action is selected. The procedural animation system receives the action command and attempts to execute it in the world; whether it is successful or not depends on the state of the world and preconditions of the action (e.g. objects that are out of reach can not be picked up).

5.5 Training

The policy is trained using the Proximal Policy Optimization RL algorithm [Schulman et al. 2017] using parameters and minor modifications specified in Appendix D. Our training infrastructure [Espeholt et al. 2019] coordinates several thousand parallel agents, each running episodes that last four minutes in simulation time but significantly less wall-clock time. A full training run takes one to four wall-clock hours and represents approximately 55,000 simulation hours of experience. Given our choice of algorithm and architecture, this training duration is the empirical threshold at which the performance plateaus and we consider the policy ready for evaluation (see Section 5.6).

Training episodes are instantiated from a curriculum of five to ten episode templates. Each template describes the initial configuration for a stochastically instantiated scene—including landscape features, food sources, and motivational drive values—designed to expose the RL agent to salient scenarios that accelerate its acquisition of useful behaviors and increase its robustness to environment variations.

Thus our workflow for developing an autonomous character replaces direct behavior scripting with the more high-level task of authoring training scenarios. Rather than narrowing the situational space and explicitly describing every expected behavior within that space, we elicit the emergence of complex behavior autonomously, by describing a set of training scenarios and allowing the RL agent to learn behavior within this curriculum.

5.6 Testing

Once trained, we validate that the brain’s action choices and emotional reactions function as expected by running the RL policy

through a series of tests. Some tests pass or fail based on whether certain conditions are met, such as *eats when hungry*, *escapes canyon*, *gets surprised by critters*, and *avoids falling from cliff*. Other tests, which we call *emotional calisthenics*,³ are more subjective in nature: we place Axo in a situation designed to trigger an interesting sequence of behaviors, and we visually evaluate the believability of the resulting behavioral performance as a whole. (See examples in the supplemental video.)

6 PROCEDURAL ANIMATION

Several of the believability principles depend directly on the quality and responsiveness of a character’s animation. Thus in this section we describe our procedural animation system, not because it is novel, but to illustrate the minimum requirements for believability, and to aid the reader who may wish to reproduce our results.

The animation system must translate the brain’s nuanced action, attention, and emotion signals into continuous expressive movement. And to exhibit the principle of *contingent interaction*, that movement must respond swiftly and gracefully to those signals, which can change at any time.

Our animation system consists of several components that run concurrently, each responsible for one aspect of the character’s performance: locomotion, gait, attention, discrete actions, emotional expression, and physical simulation. By layering the effects of these simple components, our system produces a wide variety of complex emergent behaviors from minimal hand-crafted input. (All of Axo’s animation is generated from 27 short clips totaling less than 30 seconds of material.)

6.1 Locomotion and gait system

The brain’s continuous movement signals can specify a range of movement from fast and intense to slow and subtle. The locomotion system must capture the nuance these signals provide, translating them into plausible locomotion that demonstrates *self-propulsion* as well as *physical* and *biological movement*.

Our locomotion system works similarly to that found in many game engines: it is rooted in a base node that traverses the virtual world’s navigation mesh based on the brain’s continuous movement signals, which smoothly control its forward acceleration and Y rotation velocity. A gait system procedurally animates Axo’s feet, rib cage, and pelvis relative to this base node, with timing driven by a phase value based on distance traveled. As the base node’s velocity changes, the system blends between gaits, resulting in smooth transitions, e.g. from walk to trot to gallop. Each gait is hand-crafted by an animator to caricature the weight shifts and physical relationships between body parts, creating the illusion that the character’s own muscles are driving it, supporting *self-propulsion* and *biological movement*. The feet automatically conform to the underlying terrain to ensure solid ground contact, supporting *physical movement*.

6.2 Attention and discrete actions

The attention system supports the principle of *attention legibility* by translating the brain’s attention signals into continuous movement

³So named by analogy to the physical calisthenics tests commonly used to test deformations in character rigging.

designed to clearly convey to the viewer the character’s focus of interest at all times: when the brain’s attention shifts to a new referent object, it triggers a blink animation on the character’s eyes, and the head turns crisply to follow the new referent.⁴ The timing of this action draws the viewer’s eye to the moment of change.

The brain’s discrete action commands *grab*, *drop*, and *eat* trigger procedural head animations that perform the respective actions, if conditions allow. (For example, a *grab* command will only result in an action if an object is within reach, and an *eat* command will only take effect if Axo is holding food.) The *roar* command triggers animation on the character’s entire body and face, and emits a signal to scare away critters.

6.3 Emotion verbs and adverbs

The emotion expression system supports the principles of *emotion legibility* and *explainability* by translating the brain’s emotion signals into human-readable acting via salient changes in posture, gait, and facial expression.

The brain produces nine real-valued emotion signals, consisting of an immediate, anticipated, and surprise value for each of its three motivational drives (safety, nourishment and rest). The body’s emotion expression system filters these raw signals and produces triggering events when they cross certain thresholds. These filtered signals and triggers are used to control a finite state machine consisting of different *emotion adverb* states (described in more detail in Appendix F).

Some emotions are only legible when they have the stage entirely to themselves. For these emotions, we create *emotion verbs*, which are discrete, time-bounded actions that override other actions and expressions throughout their duration. For example, the *surprise* emotion verb, triggered by the brain’s surprise signals, causes the character to stop whatever it’s doing, blink and stand up alertly for one second before continuing about its business.

6.4 Physical simulation

The locomotion, gait, attention, and emotion systems run concurrently, resulting in a single, unified performance that continuously expresses the brain’s motivational state. Driven by this performance are simple physics-based spring joint chains for the character’s headdress and tail, which provide automatic overlapping action and follow-through, reinforcing the principle of *physical movement*. These flexible appendages also support *self-propulsion* by making it clear that the character’s body is what causes them to move.

7 RESULTS

We have observed our system demonstrating a variety of interesting emergent behaviors. For example, because Axo’s initial impression of the user is somewhat negative, when the user is present Axo will visibly hesitate, approaching and retreating, exhibiting a precarious balance between its fear of the user and its desire for the plentiful fruit near the camera. After the user has built up trust by feeding Axo (Section 5.3), Axo exhibits a different behavior that has the appearance of playing “peek-a-boo”—showing disappointment and

⁴Note that this animation is purely for the viewer’s benefit: the character’s head orientation has no actual effect on its perception, because its sensors are located at its base node.

confusion when the user hides from Axo’s view, and delighted surprise when the user returns—a charming and completely unplanned result of its particular combination of motivational drives.

By observing intermediate RL policies saved partway through a single satisfactory training run, we can see how the character’s skills and emotions develop throughout the training process. After the first 3% of training steps, Axo has already learned to avoid critters—but has not yet learned to seek out and eat food for nourishment. We first observe food-seeking behavior after 25% of training. At every stage of development that we have observed, the character’s displayed attention and emotions appear consistent with its action policy. For instance, the “younger” versions of Axo that do not seek food also display no positive emotions about food. Also, at 25% of training, Axo is quite timid when confronted by critters, whereas the fully developed Axo is more brazen—suggesting how the character’s personality can seem to mature as its skills develop. For examples of these behaviors, see our supplemental video.

7.1 User study

There is as yet no established measure of believability for synthetic character animation. Nor is there broad consensus on benchmarks for evaluating interactive synthetic characters of any kind [Fitrianie et al. 2020]. Thus we lean on existing tools to serve as a proxy. Specifically, in our user studies we applied two instruments from the fields of human-robot interaction and psychology: the Godspeed questionnaires [Bartneck et al. 2009], which benchmark experience with and perception of synthetic characters (particularly used with social robots in HRI), and Dimensions of Mind Perception [Gray et al. 2007], which measures the degree to which subjects perceive an entity (human or otherwise) as having a mind.

We conducted a user study with 54 users, 80% female, with diverse backgrounds, recruited through a screening process designed to be inclusive and equitable. Data from 50 completed surveys was collected under informed consent. Respondents ran an interactive application featuring Axo in its native habitat (Figure 2). Users were instructed to run the application at full screen on laptop computers in their homes for at least 60 minutes. We did not specify how actively users should engage with the app, allowing them to control where their experience lay on the spectrum from passive to interactive. Users then responded to a survey with 32 items, including seven-point differential scale prompts on multiple constructs (anthropomorphism, animacy, likeability, and perceived intelligence) as part of the Godspeed questionnaires, as well as prompts for seven-point ratings (“strongly agree” to “strongly disagree”) on mind perception dimensions, including mental and emotional capacities. (See Appendix G for the full survey.)

Results from the Godspeed questionnaires are plotted in Figures 4 and 10. The character generally received its highest scores on the animacy and likeability Godspeed questionnaires. On every semantic differential of these two questionnaires, at least 75% of users gave ratings at or above the midpoint. The other two Godspeed questionnaires—on anthropomorphism and perceived intelligence—nonetheless also tended to have more favorable than unfavorable responses: on every semantic differential, at least 50% of users gave ratings at or above the midpoint.

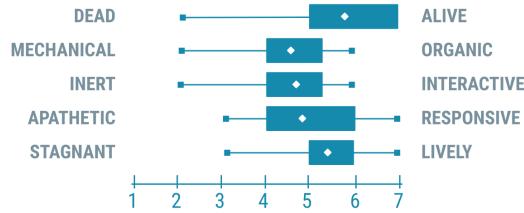


Figure 4: Results from the Godspeed questionnaire for animacy in our user study. These box-and-whisker plots show minimum, first quartile, mean, third quartile, and maximum responses.

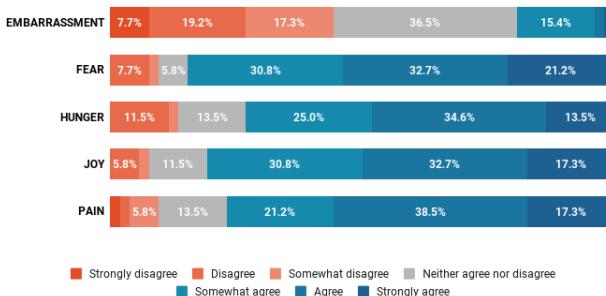


Figure 5: User study results from the evaluation of the dimensions of mind perception.

Figure 5 shows some results from users' evaluations of the dimensions of mind perception. A majority of users agreed or strongly agreed that the character experiences fear, hunger, joy, and pain. We also asked about embarrassment, an emotion we did not explicitly design the character to exhibit and would not expect to emerge without a drive or other mechanism related to social inclusion; users correspondingly tended to disagree or express neutrality regarding the character feeling embarrassment.

Participants were also asked to describe Axo in their own words. Of the responses that included pronouns, 41% used the pronoun *it*, whereas 59% used personifying pronouns such as *he*, *she* or the singular *they*, a behavior consistent with the perception of an animal as someone, not something [Wales and Siemund 2009].

8 CONCLUSION

Our goal in this work is to enable the creation of autonomous characters capable of believable acting throughout an interactive experience. In support of this goal we have laid out what we consider important principles of believability, which we hope will be useful to practitioners engaged in developing such characters. We have built one example system that exhibits these principles using reinforcement learning and layered procedural animation. Designing our character's brain holistically around these principles and basing its actions, emotions, and attention on internal motivations results in emergent behavior that is coherent, responsive, and grounded in its subjective experience, and thus feels authentic. Observing the character's growth throughout training shows all aspects of its

acting performance to be tightly coupled, and suggests that even its perceived personality is to some extent emergent from the learning process. While there is no established benchmark for measuring believability, we have found and used suitable proxy measures, and are encouraged by the results.

9 ETHICAL CONSIDERATIONS

We believe one should always be forthright about a character's synthetic nature. Those who deploy such a character must ensure its nature is understood by users, either inherently via the character's design, as with Axo, or via an explicit statement for a photorealistic character. Users can then suspend disbelief to benefit from an emotional experience with the character, while knowing that it is not real.

Increased believability may also come with risks and downsides, such as the potential for empathetic distress, social isolation, or allowing synthetic characters to influence important life decisions. Research into believability must therefore be complemented by research into corresponding risks and benefits.

10 FUTURE WORK

We consider the list of principles in Section 3.1 to be prerequisites for believability in autonomous animated characters. However, further empirical work is needed to support the relationship between these principles and believability. In this vein, fruitful directions for future work are to measure believability more directly, correlate the individual principles with established instruments, show the impact of each principle using ablation, and establish baselines for comparison.

With the exception of the heuristic user score described in Section 5.3, no additional learning takes place at runtime in our current system. We leave for future work the design of RL solutions that learn at runtime, in ways that further support *change with experience*, including learning at appropriate rates, learning reliably, and demonstrating curiosity. We implemented minimal support for *sociality* through interactions with the user that directly impact the existing motivational drives. More compelling social behavior might be achieved through the additions of a dedicated social drive and more sophisticated interaction modalities.

While our current system displays behaviors that create a short-lived illusion of demonstrating the principle of *thought*, we do not believe this illusion would hold up under longer time frames or more complex scenarios. In future work, we would like to demonstrate thinking more overtly using a construct genuinely grounded in the character's experience, e.g. a mental model of the world that could trigger specific animation cues upon incorporating new information.

As is typical for RL solutions, the effective planning horizon exhibited by our policy is dependent on algorithm choice, discount rate, and decision frequency. The current result is a character that only accounts for events less than a minute into the future. Follow-up work could focus on designing more sophisticated solutions that result in the pursuit of longer term interests.

ACKNOWLEDGMENTS

We would like to thank Paul Lee, Neth Nom, Aaron Stiles, and Jeffrey Miao for their creative contributions to this project; and David Salesin, Brian Curless, and Johnny Soraker for their valuable feedback.

REFERENCES

- Joscha Bach. 2012. A framework for emergent emotions, based on motivation and cognitive modulators. *International Journal of Synthetic Emotions* 3, 1 (2012), 43–63.
- Christoph Bartneck, Dana Kulic, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1, 1 (2009), 71–81.
- Andrew G Barto. 2013. Intrinsic motivation and reinforcement learning. In *Intrinsically motivated learning in natural and artificial systems*. Springer, 17–47.
- Joseph Bates et al. 1994. The role of emotion in believable agents. *Commun. ACM* 37, 7 (1994), 122–125.
- Marc Bellamente, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, and Remi Munos. 2016. Unifying count-based exploration and intrinsic motivation. *Advances in neural information processing systems* 29 (2016), 1471–1479.
- Pete Billington and Jessica Shamash. 2018. Wolves in the walls: chapter 1. In *ACM SIGGRAPH 2018 Virtual, Augmented, and Mixed Reality*, 1–1.
- Valentino Braitenberg. 1986. *Vehicles: Experiments in synthetic psychology*. MIT press.
- Thierry Chaminade, Jessica Hodgins, and Mitsuo Kawato. 2007. Anthropomorphism influences perception of computer-animated characters' actions. *Social cognitive and affective neuroscience* 2, 3 (2007), 206–216.
- Jonathan Cooper. 2018. The Last Guardian: Procedural Animation. Game Anim: Video Game Animaiton Explained. <https://www.gameanim.com/2018/01/24/last-guardian-procedural-animation/>
- Kate Darling, Palash Nandy, and Cynthia Breazeal. 2015. Empathic concern and the effect of stories in human-robot interaction. In *24th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 770–775.
- Lasse Espeholt, Raphaël Marinier, Piotr Stanczyk, Ke Wang, and Marcin Michalski. 2019. Seed RL: Scalable and efficient deep-RL with accelerated central inference. *arXiv preprint arXiv:1910.06591* (2019).
- Siska Fitrianie, Merijn Bruijnes, Deborah Richards, Andrea Bönsch, and Willem-Paul Brinkman. 2020. The 19 unifying questionnaire constructs of artificial social agents: An IVA community analysis. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, 1–8.
- Jacquelyne Forgette and Michael Katchabaw. 2014. Enabling motivated believable agents with reinforcement learning. In *Games Media Entertainment*. IEEE, 1–8.
- William W Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 233–240.
- György Gergely and Gergely Csibra. 2003. Teleological reasoning in infancy: The naïve theory of rational action. *Trends in cognitive sciences* 7, 7 (2003), 287–292.
- Paulo Gomes, Ana Paiva, Carlos Martinho, and Arnav Jhala. 2013. Metrics for character believability in interactive narrative. In *International conference on interactive digital storytelling*. Springer, 223–228.
- Heather M Gray, Kurt Gray, and Daniel M Wegner. 2007. Dimensions of mind perception. *science* 315, 5812 (2007), 619–619.
- Chris Hecker, Bernd Raabe, Ryan W Enslow, John DeWeese, Jordan Maynard, and Kees van Prooijen. 2008. Real-time motion retargeting to highly varied user-created morphologies. *ACM Transactions on Graphics (TOG)* 27, 3 (2008), 1–11.
- Elmer Jacobs, Joost Broekens, and Catholijn Jonker. 2014. Emergent dynamics of joy, distress, hope and fear in reinforcement learning agents. In *Adaptive learning agents workshop at AAMAS2014*.
- Joar Jakobsson and James Therrien. 2016. The Rain World Animation Process. Game Developers Conference 2016. <https://youtu.be/sVntwsrjNe4>
- Gunnar Johansson. 1973. Visual perception of biological motion and a model for its analysis. *Perception & psychophysics* 14, 2 (1973), 201–211.
- Susan C Johnson and Erica Ma. 2005. The role of agent behavior in mentalistic attributions by observers. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005*. IEEE, 723–728.
- Tobias Kleanthous. 2021. Making the Believable Horses of 'Red Dead Redemption II'. Game Developers Conference 2021. <https://youtu.be/8vtCqfFAjKQ>
- Alex Klein, Zerrin Yumak, Arjen Beij, and A. Frank van der Stappen. 2019. Data-Driven Gaze Animation Using Recurrent Neural Networks. In *Motion, Interaction and Games (MIG '19)*. ACM, New York, Article 4, 11 pages.
- George Konidaris and Andrew Barto. 2006. An adaptive robot motivational system. In *International Conference on Simulation of Adaptive Behavior*. Springer, 346–356.
- Matthew D Lieberman. 2013. *Social: Why our brains are wired to connect*. OUP Oxford.
- Libin Liu and Jessica Hodgins. 2017. Learning to schedule control fragments for physics-based characters using deep q-learning. *ACM Transactions on Graphics (TOG)* 36, 3 (2017), 1–14.
- Aaron B Loyall. 1997. *Believable Agents: Building Interactive Personalities*. Technical Report. Carnegie-Mellon University Department of Computer Science.
- Yuyan Luo and Renée Baillargeon. 2005. Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science* 16, 8 (2005), 601–608.
- Asimina Marmpena. 2021. *Emotional body language synthesis for humanoid robots*. Ph. D. Dissertation. University of Plymouth.
- Abraham Harold Maslow. 1943. A theory of human motivation. *Psychological review* 50, 4 (1943), 370.
- Ryan J McCall, Stan Franklin, Usef Faghhihi, Javier Snaider, and Sean Kugele. 2020. Artificial Motivation for Cognitive Software Agents. *Journal of Artificial General Intelligence* 11, 1 (2020), 38–69.
- Kathryn Elizabeth Merrick and Mary Lou Maher. 2007. Motivated reinforcement learning for adaptive characters in open-ended simulation games. In *Proceedings of the int'l conference on advances in computer entertainment technology*, 127–134.
- Thomas M Moerland, Joost Broekens, and Catholijn M Jonker. 2018. Emotion in reinforcement learning agents and robots: a survey. *Machine Learning* 107, 2 (2018), 443–480.
- Masaki Nakada, Tao Zhou, Honglin Chen, Tomer Weiss, and Demetri Terzopoulos. 2018. Deep learning of biomimetic sensorimotor control for biomechanical human animation. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–15.
- Deepak Pathak, Pulkit Agrawal, Alexei A Efros, and Trevor Darrell. 2017. Curiosity-driven exploration by self-supervised prediction. In *International conference on machine learning*. PMLR, 2778–2787.
- Ken Perlin. 2003. Building Virtual Actors Who Can Really Act. In *Virtual Storytelling: Using Virtual Reality Technologies for Storytelling*. Olivier Balet, Gérard Subsol, and Patrice Torguet (Eds.). Springer, Berlin, Heidelberg, 127–134.
- Ken Perlin and Athomas Goldberg. 1996. Improv: A system for scripting interactive actors in virtual worlds. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 205–216.
- Ken Perlin and Gerry Seidman. 2008. Autonomous digital actors. In *International Workshop on Motion in Games*. Springer, 246–255.
- Shaghayegh Roohi, Jari Takatalo, Christian Guckelsberger, and Perttu Hämäläinen. 2018. Review of intrinsic motivation in simulation-based game testing. In *Proceedings of the 2018 chi conference on human factors in computing systems*, 1–13.
- David Rosen. 2014. An Indie Approach to Procedural Animation. Game Developers Conference 2014. <https://youtu.be/LNidsMexsSE>
- Micah Marlon Rosenkind. 2015. *Creating Believable, Emergent Behaviour in Virtual Agents, Using a Synthetic Psychology Approach*. Ph. D. Dissertation. U. of Brighton.
- Mark Sagar, David Bullivant, Paul Robertson, Oleg Efimov, Khurram Jawed, Rathesh Kalarat, and Tim Wu. 2014. A neurobehavioural framework for autonomous animation of virtual human faces. In *SIGGRAPH Asia 2014 Autonomous Virtual Humans and Social Robot for Telepresence*, 1–10.
- Ayse Pinar Saygin, Thierry Chaminade, Hiroshi Ishiguro, Jon Driver, and Chris Frith. 2012. The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social cognitive and affective neuroscience* 7, 4 (2012), 413–422.
- John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
- Alireza Shirvani and Stephen G Ware. 2019. A plan-based personality model for story characters. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 15. 188–194.
- Lee Strasberg and Ned Chaillet. 2021. Acting. Encyclopedia Britannica. <https://www.britannica.com/art/acting>
- Harald Strömfelt, Yue Zhang, and Björn W Schuller. 2017. Emotion-augmented machine learning: overview of an emerging domain. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 305–312.
- Richard S Sutton and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.
- Tianxin Tao, Matthew Wilson, Ruiyu Gou, and Michiel van de Panne. 2022. Learning to Get Up. In *ACM SIGGRAPH 2022 Conference Proceedings* (Vancouver, BC, Canada). ACM, Article 47, 10 pages.
- Frank Thomas and Ollie Johnston. 1981. *The Illusion of Life: Disney Animation*. Hyperion New York.
- John Ronald Reuel Tolkien and Christopher Tolkien. 1984. On Fairy Stories. In *The Monsters and the Critics, and Other Essays*. Boston: Houghton Mifflin.
- Iskander Umarov and Maxim Mozgovoy. 2012. Believable and effective AI agents in virtual worlds: Current state and future perspectives. *International Journal of Gaming and Computer-Mediated Simulations (IJGEMS)* 4, 2 (2012), 37–59.
- Katie Wales and Pieter Siemund. 2009. Pronominal gender in English: A study of English varieties from a cross-linguistic perspective. *English Language and Linguistics* 13, 3 (2009), 508.
- Henrik Warpefelt. 2016. *The Non-Player Character: Exploring the believability of NPC presentation and behavior*. Ph. D. Dissertation. Department of Computer and Systems Sciences, Stockholm University.
- Stewart W Wilson. 1991. The animat path to AI. In *From Animals to Animats: Proceedings of the First International Conference on the Simulation of Adaptive Behavior*, J. A. Meyer and S. W. Wilson (Eds.). MIT Press/Bradford Books.

A PRINCIPLES OF BELIEVABILITY: COMPARISONS TO RELATED WORK

Gomes et al. [2013] proposed a list of attributes thought to contribute to a character’s believability. We made the following refinements to that list: We added *physical movement*, *biological movement*, *self-propulsion*, *contingent interaction*, *self-motivation*, and *thought*. We replaced *awareness* with *attention appropriateness* and *legibility*, to clarify that it’s the character’s observable behavior that matters, not whether it is literally aware of its surroundings. We replaced *emotional expressiveness* with *emotion appropriateness* and *legibility*, which more precisely specify what’s important for believability (for example, a stoic character like Mr. Spock can be quite believable without being very emotionally expressive.) We replaced *understandability* with *explainability*, to emphasize the viewer’s active role in constructing a mental model of the character’s intentions and desires. We removed *visual impact*, which we consider an aspect of character design (see discussion in Section 4.3). We also removed *predictability*, which we found too vague to be useful: for every case near either extreme of the predictability spectrum that a viewer might find unbelievable, the other principles are more informative at explaining why. For example, a “too unpredictable” character would likely also fail to demonstrate *explainability*, *self-motivation*, or *coherence of identity*, and a “too predictable” character would fail to show *change with experience* or *contingent interaction*.

B OBSERVATION STRUCTURE

The RL observation vector consists of information about both the character’s internal state and sensed elements of the environment. Since we are not concerned with solving low-level problems like visual perception, the design of the environment sensors is intentionally simple and abstract, but biologically plausible in that they only report information that an animal could reasonably be expected to observe using some combination of its sensory organs.

The observed internal state includes the three drive values (nourishment, safety, and rest), the change of each drive value since the last time step, the character’s egocentric velocity, whether the character currently holds an object, and its current evaluation of the user score (see Section 5.3).

To represent observations of the character’s environment, a 3D ray-tracing sensor reports the type and position of each unobstructed object within the character’s field of view, defined as a 70° cone extending forward from the character’s base node (see Figure 6). We also report the last-seen relative position of each object that is no longer visible and the time since it was last seen, for up to a minute. Up to 21 objects are reported.

Lastly, the navigability of the terrain is observed via a 2D ray-tracing sensor that reports distances to the edge of the navigable terrain at seven fixed angular intervals to the front of the character and three shorter range intervals to the back (Figure 7). This navigability sensor also reports the height of the jump at that edge if applicable. The combined data reported by these sensors results in the observation being a real-valued vector of size 130.

C REWARD DESIGN DETAILS

Below we provide complete details for the specification of our motivational drives and reward function, as introduced in Section 5.1.

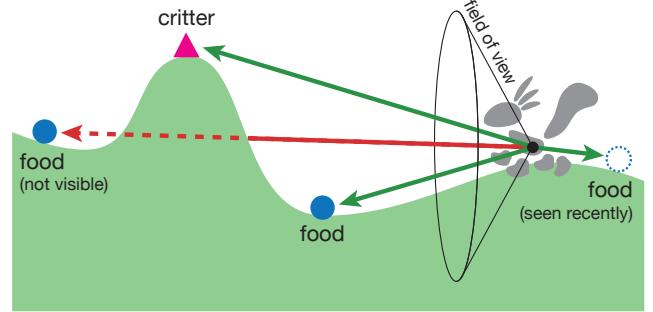


Figure 6: Diagram of the ray-tracing object sensor.

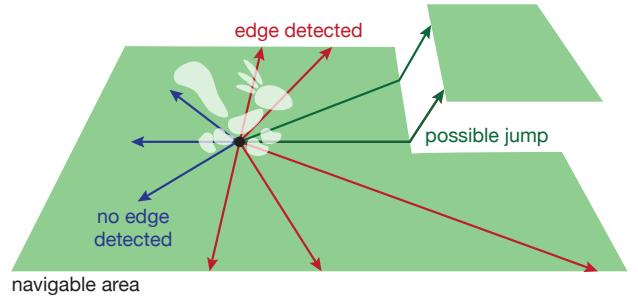


Figure 7: Diagram of the navigability sensor.

C.1 Drive parameters and update rules

In our character we implement three drives: nourishment, safety, and rest. The value of each drive is in the range [0, 1]. Drive state values are updated every time step, which is approximately every 200 ms.

C.1.1 Nourishment.

- The initial value is 1.
- The set point p_d is 0.
- The rate of return to the set point, α_d , is 0.0005.
- Update rule u_0 : For each unit of nutrition of food objects that are eaten, increase this drive value by 0.3. The fruits in the environment have a nutrition value of 1.

C.1.2 Safety.

- The initial value is 0.5.
- The set point p_d is 0.5.
- The rate of return to the set point, α_d , is 0.001.
- Update rule u_0 : When a critter touches the character, decrease this drive value by a sampled value from the uniform distribution over [0, 0.8].
- Update rule u_1 : When the user pokes the character, decrease this drive value by:
 - 0.264 if poked on its feet and
 - 0.2 if poked on its head, chest, or hips.
- Update rule u_2 : When the user pets the character, increase this drive value by the following, per second of petting:
 - 0.16 if petted on its feet;
 - 0.32 if petted on its head or chest; and

- 0.6 if petted on its hips.
- Update rule u_3 : When the character falls down from a height greater than 0.3 m, decrease this drive value by an amount linearly related to the height in the range [0, 0.8], applying the maximum decrease of 0.8 for all heights ≥ 0.5 m.

C.1.3 Rest.

- The initial value is 1.
- The set point p_d is 1.
- The rate of return to the set point, α_d , is 0.00125.
- Update rule u_0 : For each meter per second of speed of movement in the horizontal plane, decrease this drive value by 0.0015.
- Update rule u_1 : For each meter per second of vertical speed, decrease this drive value by 0.00225.
- Update rule u_2 : For each radian per second of rotational speed, decrease this drive value by 1.5×10^{-6} .
- Update rule u_3 : For each roar action, decrease this drive value by 0.24.

C.2 Reward calculation

Reward for each time step is the sum of the three reward components, one from each of the drives listed above. The function w_d for each drive, which maps the drive state value to a reward component value, is shown in Figure 8.

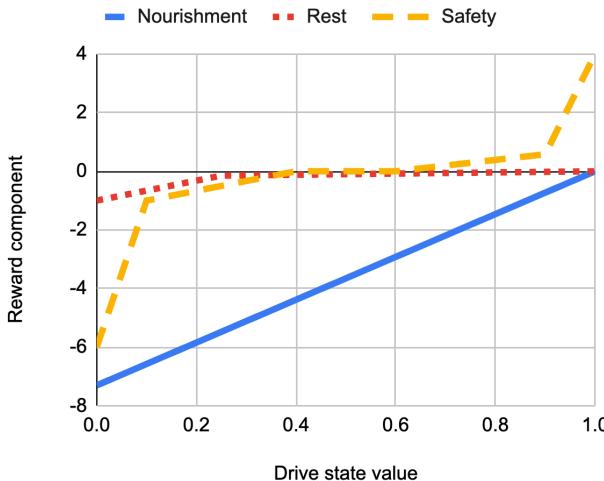


Figure 8: Reward mappings w_d for each drive.

To avoid constant eating, the nourishment reward is further multiplied by an appetite value, which is in the range [0, 1]. At a high-level: when the character’s nourishment is high its appetite decreases, and when it is low then its appetite increases. The non-linearity of triggering the increase/decrease by thresholds on the nourishment value results in dynamics where eating food provides disproportionately more value when the appetite is high vs. low, resulting in behavior where the character eats to satisfaction and then stops until appetite is high again, vs. continuously eating food as the nourishment value oscillates in a saw-tooth pattern. On each step, the appetite value is modified as such:

```

1: if nourishment  $\geq 0.8$  then
2:   increment  $\leftarrow -0.005$ 
3: else if nourishment  $\leq 0.2$  then
4:   increment  $\leftarrow 0.005$ 
5: end if
6: appetite  $\leftarrow \max(0, \min(1, appetite + increment))$ 

```

D TRAINING DETAILS

We train using the Proximal Policy Optimization (PPO) algorithm [Schulman et al. 2017] with the following parameters:

- Adam step size (α): 2.5e-4
- Entropy bonus coefficient: 1.7e-6
- Minibatch size: 512
- Unroll length: 8
- Discount (γ): 0.95
- GAE parameter (λ): 0.9
- Number of policy network layers: 4
- Number of value (and anticipation) network layers: 6
- Number of nodes per layer: 512

In addition to the behavior policy and value function networks, our system adds a drive-decomposed return-prediction network to aid in emotion generation. This new network is identical to the value function network, except that instead of outputting a single value for the overall discounted return, it outputs a vector of values, where each value estimates what the discounted return would be if only its corresponding drive provided reward.

We modified PPO’s loss function (eqn. 9 in Schulman et al. [2017]) by replacing the $L_t^{VF}(\theta)$ term with $(L_t^{VF}(\theta) + L_t^{DF}(\theta))$ where L_t^{DF} is the mean loss of the drive-decomposed network, calculated for each head in the same way as for the value function loss.

E EMOTION SIGNAL GENERATION

We generate anticipated emotion signals by subtracting each drive-decomposed return estimate from its moving average and scale the result by a calibrated constant to generate values in the range expected by the emotion expression system.

We trigger a per-drive surprise signal if two conditions hold: the drive has no active surprise from the previous time step, and the difference between its return estimate and its moving average goes beyond a calibrated positive or negative threshold. Once triggered, these positive or negative surprise signals decay to zero after one second.

Both sets of anticipated emotion and surprise signals are attached to the most salient referent object using counterfactually generated observations as described in Section 5.2.

F EMOTION EXPRESSION SYSTEM

The procedural emotion expression system uses a finite state machine consisting of emotion adverb states, which modify the character’s posture, facial expression, breathing, and gait; and emotion verbs, which trigger specific time-bounded actions.

Each adverb state may contain a single pose, or may allow blending between two or more poses based on the values of specific filtered emotion signals from the RL brain. For example, the *fear* adverb state blends between the “fear” and “neutral” poses (see Figure 9) as the value of the filtered *safety anticipation* signal varies

between 0.25 and 0.6. This allows the system to display subtle, nuanced changes in response to fluctuations in the brain's emotion signals, and keeps the character feeling alive and dynamic.

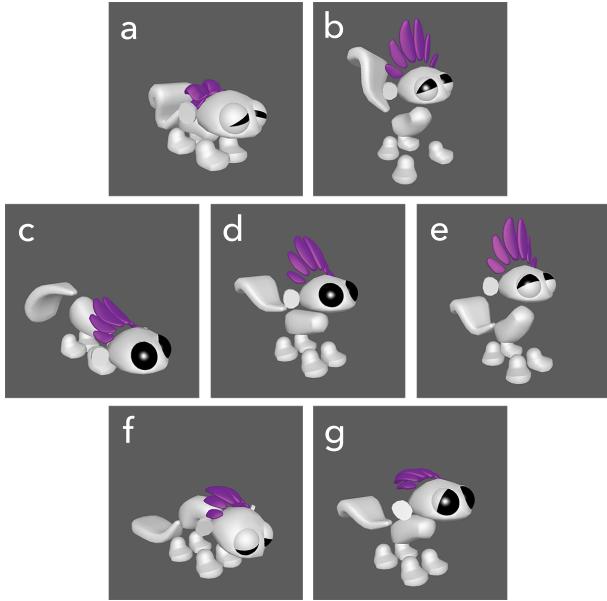


Figure 9: Postures and facial expressions used in emotion adverb states: (a) pain, (b) pleasure, (c) fear, (d) neutral, (e) hope, (f) exhaustion, (g) hunger.

If the system is not busy with an emotion verb (such as *surprise*) or other physical action (such as *roar*), it triggers a transition to a new adverb state when one of the following thresholds is crossed:

- *nourishment anticipation* > 0.55 → *hope*
- *safety anticipation* > 0.8 → *hope*
- *safety anticipation* < 0.4 → *fear*
- *safety* < 0.2 → *pain*
- *safety* > 0.8 → *pleasure*
- *nourishment* > 0.8 → *satisfaction*
- *nourishment* < 0.2 → *hunger*
- *rest* < 0.2 → *exhaustion*

These state transitions are sharp and fast (1/4 second in duration), which makes them highly salient, drawing the viewer's attention to the fact that an important emotional change has occurred.

G USER STUDY SURVEY QUESTIONS

- (1) Please enter the number you see for "Total runtime"
- (2) Please enter the number you see for "User present time"
- (3) Please enter the number you see for "User mouse clicks"
- (4) For how long did you feel engaged while participating in the Axo experience?
- (5) What was your overall impression of the Axo experience?
- (6) What was your impression of Axo?
- (7) What was your impression of the environment where Axo lives?
- (8) Did you interact with objects in the environment (e.g., pick, move and drop food, critters)?

- (9) Did you pet Axo?
- (10) Did you poke Axo?
- (11) Did you feed Axo?
- (12) Did you notice the day and night cycles?
- (13) Did you leave your camera on throughout the entire time you ran the Axo app?
- (14) To what extent do you agree or disagree with the following statements?
 - I believe my interactions with Axo were engaging
 - I believe Axo demonstrated interest in me
 - I believe Axo is capable of conveying thoughts or feelings to others
 - I believe Axo is capable of having experiences and being aware of things
 - I believe Axo is capable of longing or hoping for things
 - I believe Axo is capable of experiencing embarrassment
 - I believe Axo is capable of feeling afraid or fearful
 - I believe Axo is capable of feeling hungry
 - I believe Axo is capable of experiencing joy
 - I believe Axo is capable of experiencing pain
 - I believe Axo is capable of understanding how others are feeling
 - I believe Axo is capable of remembering things
 - I believe Axo is capable of working toward a goal
 - I believe Axo is capable of having personality traits that make it unique from others
- (15) Please rate your impression of Axo on these scales:
 - Fake:Natural
 - Machinelike:Humanlike
 - Unconscious:Conscious
 - Artificial:Lifelike
 - Moving rigidly:Moving elegantly
- (16) Please rate your impression of Axo on these scales:
 - Dead:Alive
 - Mechanical:Organic
 - Inert:Interactive
 - Apathetic:Responsive
 - Stagnant:Lively
- (17) Please rate your impression of Axo on these scales:
 - Dislike:Like
 - Unfriendly:Friendly
 - Unkind:Kind
 - Unpleasant:Pleasant
 - Awful:Nice
- (18) Please rate your impression of Axo on these scales:
 - Incompetent:Competent
 - Ignorant:Knowledgeable
 - Irresponsible:Responsible
 - Unintelligent:Intelligent
 - Foolish:Sensible
- (19) What, if anything, made Axo seem lifelike or not lifelike?
- (20) Imagine there was an event that would destroy all characters (i.e., Axo and the pyramidal critters) and their entire environment, but you could choose to save one of the characters. Which character would you choose to save?

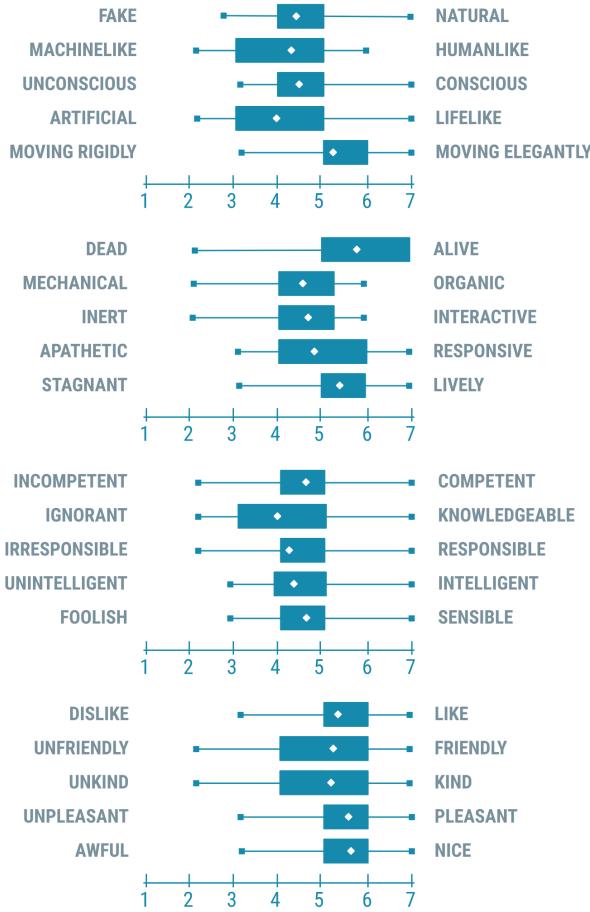


Figure 10: Results from the Godspeed questionnaires for anthropomorphism, animacy, perceived intelligence, and likeability. These box-and-whisker plots show minimum, first quartile, mean, third quartile, and maximum responses.

- (21) Considering your answer to the previous question, what actions would you take, if any, to save the character from being hurt?
- (22) Imagine that you could make one of the characters happy. Which character would you choose to make happy?
- (23) Considering your answer to the previous question, what actions would you take, if any, to make the character happy?
- (24) Do you feel this experience was more of a lean-in, engaged game where you actively participate or more of an ambient, less actively engaged experience?
- (25) Now that you have interacted with Axo, in which contexts or situations would you want to have this type of experience?
- (26) Which of the following would be reasons why you would not want to interact with Axo?
- (27) On which of these devices would you find the Axo Experience most relevant to you?
- (28) What kinds of experiences have you had with a digital character?
- (29) How does Axo compare to those experiences?
- (30) What else could make the Axo experience better for you?
- (31) Would you like to be notified of future opportunities to experience and test our latest efforts in this area?
- (32) What is your gender? Choose all that apply.

G.1 Survey Results

Results from the Dimensions of Mind questionnaires are shown in Table 1. The full results from the Godspeed questionnaires are shown in Table 2 and plotted in Figure 10.

	Strongly Disagree	Disagree	Somewhat Disagree	Neither	Somewhat Agree	Agree	Strongly Agree
Embarrassment	7.7%	19.2%	17.3%	36.5%	15.4%	1.9%	1.9%
Fear	0.0%	7.7%	1.9%	5.8%	30.8%	32.7%	21.2%
Hunger	0.0%	11.5%	1.9%	13.5%	25.0%	34.6%	13.5%
Joy	0.0%	5.8%	1.9%	11.5%	30.8%	32.7%	17.3%
Pain	1.9%	1.9%	5.8%	13.5%	21.2%	38.5%	17.3%

Table 1: User study results from the evaluation of the dimensions of mind perception, for questions about Axo’s ability to experience emotions. The table displays percentage of responses on a 7-point scale from Strongly Disagree to Strongly Agree.

ANTHROPOMORPHISM	MIN	MAX	MEAN	STDEV	Q1	Q3
Fake ←→ Natural	2.75	7.00	4.33	1.52	3.75	5.00
Machinelike ←→ Humanlike	2.00	6.00	4.21	1.46	3.00	5.00
Unconscious ←→ Conscious	3.00	7.00	4.44	1.34	4.00	5.00
Artificial ←→ Lifelike	2.00	7.00	3.90	1.52	3.00	5.00
Moving Rigidly ←→ Moving Elegantly	3.00	7.00	5.19	1.33	5.00	6.00
ANIMACY	MIN	MAX	MEAN	STDEV	Q1	Q3
Dead ←→ Alive	2.00	7.00	5.83	0.99	5.00	7.00
Mechanical ←→ Organic	2.00	6.00	4.62	1.35	4.00	5.25
Inert ←→ Interactive	2.00	6.00	4.73	1.33	4.00	5.25
Apathetic ←→ Responsive	3.00	7.00	4.81	1.40	4.00	6.00
Stagnant ←→ Lively	3.00	7.00	5.46	1.18	5.00	6.00
PERCEIVED INTELLIGENCE	MIN	MAX	MEAN	STDEV	Q1	Q3
Incompetent ←→ Competent	2.00	7.00	4.52	1.32	4.00	5.00
Ignorant ←→ Knowledgeable	2.00	7.00	3.90	1.33	3.00	5.00
Irresponsible ←→ Responsible	2.00	7.00	4.19	0.94	4.00	4.00
Unintelligent ←→ Intelligent	3.00	7.00	4.25	1.19	3.75	5.00
Foolish ←→ Sensible	3.00	7.00	4.60	1.18	4.00	5.00
LIKEABILITY	MIN	MAX	MEAN	STDEV	Q1	Q3
Like ←→ Dislike	3.00	7.00	5.27	1.23	5.00	6.00
Unfriendly ←→ Friendly	2.00	7.00	5.23	1.37	4.00	6.00
Unkind ←→ Kind	2.00	7.00	5.15	1.13	4.00	6.00
Unpleasant ←→ Pleasant	3.00	7.00	5.52	1.15	5.00	6.00
Awful ←→ Nice	3.00	7.00	5.58	0.99	5.00	6.00

Table 2: Results from the Godspeed questionnaire in our user study. Results display minimum, maximum, mean, standard deviation and interquartile ranges of responses to five 7-point semantic differential scales for the concepts of anthropomorphism, animacy, perceived intelligence, and likeability.