1 2 9 0

UNIVERSIDADE Ð COIMBRA

Nuno Vicente Borges Correia Malaquias Barreto

Studying Believability
In Videogame Creature Design

Tese no âmbito do Programa Doutoral em Ciências e Tecnologias de Informação orientada pelos Professores Doutores Licínio Gomes Roque e Fernando Penousal Machado e apresentada ao Departamento de Engenheira Informática da Faculdade de Ciências e Tecnologias da Universidade de Coimbra.

Agosto de 2019

UNIVERSIDADE DE COIMBRA

DOCTORAL THESIS

Studying Believability in Videogame Creature Design

Author:
Nuno BARRETO

Supervisor:
Dr. Licínio ROQUE
Dr. Penousal MACHADO

Submitted in fulfillment of the requirements
for the degree of Doctoral Program in Information Sciences and Technologies

Centro de Informática e Sistemas da Universidade de Coimbra
Departamento de Engenharia Informática

August 26, 2019

"It's just an illusion that I ... really exist; I'm just representative of a type."

In Do Androids Dream of Electric Sheep, Philip K. Dick

UNIVERSIDADE DE COIMBRA

Abstract

Faculdade de Ciências e Tecnologias da Universidade de Coimbra

Departamento de Engenharia Informática

Doctoral Program in Information Sciences and Technologies

Studying Believability in Videogame Creature Design

by Nuno BARRETO

Videogames are shifting towards creating, and simulating, meticulously detailed environments, and worlds. It becomes evident that, in order to support this world-building process, they must be populated by beings. These beings can come in the form of creatures. By assuming them to be zoomorphic entities that do not qualify as fundamentally human-like, whether or not they have anthropomorphic features, this thesis attempts to study how can we help designers create 'believable' creatures.

As content creation is a lengthy and costly process, increasing with the project's complexity, our first solution revolved around automating this process. However, this initial effort did not prove to be successful, due to the lack of a way to assess our creations. Consequently, we reflected upon our initial goals and pivot the research accordingly. To this extent, an additional literature review was made, framing creatures, within the context of a videogame artifact. This lead us to 'believability' and our proposal was materialized in the form of a construct and a Likert Scale, based on principles drawn from biology, animation, illustration and artificial intelligence.

The idea behind the scale was to ultimately help designers improve their artifacts under a new perspective, that of 'believability'. As such, it underwent four revision processes: after developing the scale's 46 original items, it was administrated as a questionnaire. The resulting data was analyzed through a Principal Component Analysis and it suggested retaining 26 items, out of the original 46, spread across 4 dimensions. The scale was administrated, a second time, in a similar setup, so as to perform a goodness-of-fit test. This was done through Confirmatory Factor Analysis. Nevertheless, results recommended further validation. As such, a third revision was made and an additional administering step. This time, through an Intraclass Correlation Coefficient, they indicated the scale was capable of providing consistent measurements across 6 experts. However, some of its items were not being interpreted as intended. With this in mind, a group interview was conducted in order to refine the scale's statements to further increase their readability.

After iterating upon the scale, we then used it to guide the design of a videogame prototype, named Orphibs Evolvapalooza. This prototype was developed as a proof-of-concept, meant to showcase the possibility of designing a creature around the scale's items, translating them into computational aspects and incorporating these aspects into the creature's agent architecture. It was first created in paper form and, after an usability test, made into software. The latter was followed by a second usability test, before being deployed online.

Data was collected from the people who played the game. This allowed us to verify the scale once more, taking into account its internal consistency and the measure agreement among players. The information gathered was used in our initial goal of automating creature content creation.

This case study demonstrated it is possible to create a *quasi-automated* system that helps explore a search-space defined by the scale itself. The scale helps filter ideas from a large set, narrowing down the creature's final presentation. Procedural Content Generation approaches, in particular those under a mixed-initiative paradigm, can make use of this insight.

By offering a non-anthropomorphic perspective on how 'believability' can be assessed on videogame creatures, this work expands the Believable Agent research field. This new insight, the construct underlying the scale, and the scale itself are the main contributions of this work. As far as we can determine, the scale can be a useful part of the creative process behind creatures.

Keywords: Creature Design, Believability, Video Games

UNIVERSIDADE DE COIMBRA

Resumo

Faculdade de Ciências e Tecnologias da Universidade de Coimbra
Departamento de Engenharia Informática

Programa Doutoral em Ciências e Tecnologias da Informação

Studying Believability in Videogame Creature Design

por Nuno BARRETO

Os videojogos estão-se a focar, cada vez mais, em criar, e simular, mundos meticulosamente detalhados. Torna-se evidente que estes mundos necessitam de ser preenchidos por entidades. Algumas destas podem vir na forma de criaturas. Ao assumi-las como zoomorfas, desprovidas de quaisquer características antropomórficas que as tornem percetivelmente humanóides, esta investigação tem como objetivo estudar maneiras de ajudar designers a criar criaturas 'credíveis'.

Como a criação de conteúdos é um processo caro e moroso, que aumenta em proporção à complexidade do projeto, a nossa solução inicial envolvia a criação de um método automatizado. No entanto, o nosso primeiro esforço não foi bem-sucedido, devido à falta de avaliação. Seguiu-se, portanto, uma introspeção que resultou numa mudança de direção, da investigação, e, após uma revisão adicional da literatura que enquadrava criaturas no contexto de um artefacto vídeo-lúdico, chegou-se ao conceito de 'credibilidade'. Isto permitiu propor um construto e uma Escala de Likert, cujas bases advém das áreas de biologia, animação, ilustração e inteligência artificial.

A escala teve como objetivo ajudar designers de criaturas a melhorar os seus artefactos. Como tal, a escala sofreu quatro processos de revisão: após a criação dos seus 46 itens originais, esta foi administrada no formato de um questionário. Os resultados foram analisados, através de uma Análise de Componentes Principais que sugeriu reter 26, dos 46 itens originais, e distribuí-los por 4 dimensões. Posteriormente, a escala foi novamente administrada onde, através de uma Análise Factorial Confirmatoria, foram feitos testes de aderência. Aqui, os resultados sugeriram validações subsequentes. Deste modo, foi efetuado uma terceira revisão, sucedida de outra aplicação. Desta vez, calculando Coeficientes de Correlação IntraClasses, foi indicado que havia, em grande parte, concordância entre 6 avaliadores especialistas. No entanto, concluiu-se também, que alguns dos itens estavam a ser interpretados incorretamente. Como tal, foi efetuado uma entrevista de grupo para melhorar a legibilidade destes.

Após a escala ter sido iterada, esta foi usada para guiar o design de um protótipo, de um videojogo, chamado Orphibs Evolvapalooza. Este protótipo foi construído para ser uma prova de conceito que demonstra a possibilidade de desenhar criaturas, tendo em conta os itens da escala, de traduzí-los em computações e incorporá-los numa arquitetura de agentes. O protótipo foi inicialmente implementado na forma de papel e, após um teste de usabilidade, foi transformado em software. Este último sofreu um segundo teste de usabilidade, antes de ser disponibilizado online.

Foram colhidos dados das pessoas que o jogaram. Isto permitiu verificar a escala de acordo com a sua consistência interna e medir o grau de concordância entre avaliadores. Todo este processo proporcionou uma oportunidade para refletir no nosso objetivo inicial, o de automatizar a criação de criaturas.

Este caso de estudo demonstrou que era possível desenvolver um sistema 'quase' automático que ajuda a explorar um espaço de procura definido pela própria escala que, por sua vez, ajuda a filtrar ideias de um conjunto maior, contribuindo para afunilar a apresentação final da criatura. Abordagens de Geração Procedimental de Conteúdos, em particular aquelas que usem um paradigma de Iniciativa Mista, podem usufruir deste conhecimento.

O trabalho aqui feito oferece uma perspectiva não-antropomórfica sobre como 'credibilidade' pode ser vista em criaturas de videojogos. Como tal, esta investigação expande o campo de Agentes Credíveis. Esta perspectiva, o construto por detrás da escala e a própria escala são os contributos principais deste trabalho e, pelo que pudémos determinar, a escala pode ser uma parte útil no processo creativo por detrás da concepção de criaturas.

Palavras-Chave: Design de Criaturas, Credibilidade, Videojogos

Acknowledgements

To my girlfriend, Fátima Martins and to my family, particularly my parents and grandmother, for all their love and support,

To my advising professors Licínio Roque, and Penousal Machado, for guiding me throughout this perilous journey and helping me focus on what was important,

To David 'David-Sama' Fial, for all his help, support, editing, and his famous 'Secluded Philosopher / Greek Oracle' wisdom,

To my friend, and colleague, Rui Craveirinha for all his insights and his help in 'making science', as we call it,

To my friends Henrique Batista, Daniel Lobo, Eduardo Bento, Tiago Prelhaz and Pedro Caetano, for lending me their shoulders (to cry on), for their insight and for beta-testing most (if not all) of the prototypes from this thesis (and enduring the resulting headaches),

To my friends Pedro Oliveira and Sérgio Alves for realizing the artistic vision of Orphibs Evolvapalooza,

And last, but not least, to everyone else who helped me along the way and conquer this intellectual 'Moby Dick'.

This thesis was partly supported by the CISUC ICIS project CENTRO-07-0224-FEDER-002003 and the FCT Ph.D. Grant SFRH/BD/100080/2014.

Contents

List of Figures

List of Tables

List of Abbreviations

CBS	**C**reature **B**elivability **S**cale
CFA	**C**onfirmatory **F**actor **A**nalysis
DSR	**D**esign **S**cience **R**esearch
EFA	**E**xploratory **F**actor **A**nalysis
FPS	**F**irst **P**erson **S**hooter
GA	**G**enetic **A**lgorithm
ICC	**I**ntraclass **C**orrelation **C**oefficient
IQR	**I**nter**Q**uartile **R**ange
PCA	**P**rincipal **C**omponents **A**nalysis
PCG	**P**rocedural **C**ontent **G**eneration
RPG	**R**ole **P**laying **G**ame
RTS	**R**eal-**T**ime **S**trategy

Chapter 1

Introduction

As technology advances, so does the potential for creating vast game worlds. Indeed, in recent years, these worlds have been increasingly larger. For instance, The Witcher 3: Wild Hunt (CD Projekt RED, 2015) was advertised as having a world size superior to its previous installment (Winegarner, 2013). Another more recent example is Red Dead Redemption 2 (Rockstar Studios, 2018); its world (estimated to have $75km^2$) is considered larger that the company's previous game, Grand Theft Auto V (Sohrabi-Shiraz, 2018), a game already known for its vast dimension (Dane, 2014). Regardless of how imaginative these worlds may be, they provide the means for players to lose any sense of technological mediation (Lombard and Ditton, 1997) and immerse themselves in such worlds. However, this immersion, or 'suspension-of-disbelief', may break if they feel empty, or lack some sort of fauna to support them (Rosenkind, Winstanley, and Blake, 2008).

Fauna, or rather creatures, comprise several types of audio-visual content, and creating them comes at a cost (in the form of time and man-power (Hendrikx et al., 2013; Togelius et al., 2013; Merrick et al., 2013)). To mitigate this, there are automated solutions, named Procedural Content Generation (PCG), aimed to reduce production costs for developers (Hendrikx et al., 2013; Togelius et al., 2013; Nareyek, 2007). Nonetheless, this approach still faces many challenges. For instance, PCG content may be perceived as bland or non-creative (Togelius et al., 2013). Additionally, techniques for generating asset compositions, such as creatures, are still ongoing research (Hendrikx et al., 2013). Even with the proposal of a system capable of creating, and assembling creatures, the risk of them being perceived as bland still remains.

By studying the intricacies behind the production of creative artifacts (Csikszentmihalyi, 1997), it becomes apparent this whole process alternates between a creation phase and an evaluation one. While the literature has several instances tackling each of the creatures' components individually (morphology (Guo et al., 2014), animation (Hecker et al., 2008) and behavior (Lionhead Studios, 2001)), with some even trying to assemble them altogether (Sims, 1994), these methods lack any kind of non-functional evaluation to assess how a creature fares when interpreted in a videogame context. We argue that, before being able better direct the production of these PCG systems, one such evaluation method must first exist.

Our research focuses on studying how to support the 'suspension-of-disbelief' mediated by the videogames' fictional worlds. This concept is known as 'believability' and it plays an integral part in improving player experience through engagement (Warpefelt, Johansson, and Verhagen, 2013). As a research field, there is some work in the creation of believable beings (Eladhari et al., 2011; Mateas, 2003) albeit, most, if not all, concentrates on human characters. As a final point, it is worth noting that, colloquially, although believability may be regarded as a synonym for realism, the literature contradicts this by citing cartoons as test cases (Thomas and Johnston, 1997).

To be believable, an agent must appear in such a way that players are able to accept it "is alive and thinking" (Rosenkind, Winstanley, and Blake, 2012). In humans, this means a set of criteria must be met, including showing personality, emotion, self-motivation, change and social relationships (Thomas and Johnston, 1997). Creatures, on the other hand, interact with humans differently (for instance, they lack a human-understandable language) so the aforementioned criteria might not hold true.

We conjecture potential criteria should be reflected in the audio-visual perception of several elements. As such, we propose an instrument (Wobbrock and Kientz, 2016) capable of measuring believability so that creature design, automated, or otherwise, can be made towards maximizing its perception. Furthermore, such a tool can help inform and guide the iterative creation process through a set of derived heuristics or rules (Csikszentmihalyi, 1997). In the end, the knowledge gathered could allow us to answer "How can we empower creature designers to devise (automated) creatures through an evaluation instrument, capable of measuring creature assets in a contextualizing environment?"

1.1 Research Summary

Our research effort started with the definition of the research problem, following a literature review in Procedural Content Generation of creatures. This was succeeded by the creation of two proof-of-concept prototypes known, respectively, as Orphibs II and the Savanna Simulation. These were meant to study the feasibility of creating a system capable of outputting complete creatures, into a videogame. After the latter prototype underwent a survey analysis, it was concluded these systems lacked some form of evaluation method to assess creatures, framed in a videogame context.

This led to a second revision of the literature, this time, focusing on potential measurable characteristics that could be rated, during a creature artifact's evaluation phase. From this second revision, we found 'believability', a construct defined by a set of elements that, when perceived, observers create a mental model of an artifact that is interchangeable from its 'real-life' counterpart, even if in a fictional context. However, the state-of-the-art in believability was shown to be anthropocentric. As a result, we proposed a creature believability model, from which we derived the Creature Believability Scale. The latter was created as an evaluation method to assess creatures, in a videogame context.

This psychometric scale underwent an iterative revision process. Specifically, there were 4 revisions, before its current form. Its first revision resulted from the evaluation of its initial construction process. This was made with an Exploratory Factor Analysis, using the data collected from its administering (this consisted in observing several gameplay movie clips with associated questionnaires). The scale was administered a second time, using the previous methodology, albeit with the objective of validating its underlying theoretical model, through a Confirmatory Factor Analysis. From this, results suggested the scale's underlying model only reasonably fit the data obtained. As a consequence, it suffered an additional revision. By this time, the experimental methodology was revised to provide a setup where participants could interact directly with videogame artifacts, instead of passively watching movie clips. With this in mind, the third revision process was made to assess the scale's agreement among observing videogame experts. After conducting this experiment, results suggested these videogame experts were able to agree on most evaluations given to 20 different commercial videogames. However, they also suggested some of the items were not being interpreted correctly. So, we carried out follow-up interviews to understand the scale's potential pitfalls. For this reason, we revised the scale an additional time, with the goal of improving its readability. Finally, this revision was refined in a group interview, resulting in its current form.

With the scale's fourth revision, we took the opportunity to revise the previously made initial prototypes, now under the guidelines resulting from the scale's statements. This allowed us to create a videogame, called Orphibs Evolvapalooza, so as to illustrate how the scale's items can guide the creation of believable creatures. At the same time, we incorporated, in said videogame, a generative system for instantiating creature archetypes; this, in turn, allowed us to collect evaluation data on several different creatures. At its core, Orphibs Evolvapalooza is designed around the idea of using the Creature Believability Scale to evaluate creatures, in order to evolve better ones. This prototype allowed us to confirm the scale's initial experiment, this time in an interactive play context.

The studies conducted during this research effort show the believability model we proposed can be a valuable tool in creature design. Particularly, they show how an evaluation instrument can be derived in order to facilitate both the assessment and the creation of virtual creatures. Additionally, this process can be further supplemented by automated systems to explore search-spaces centered on this model as illustrated in the Orphibs Evolvapalooza case study. In the end, the research documented in this thesis gives a new, contextualized, perspective on how to think about creature design and provides the means to empower designers.

1.2 Document Structure

This thesis is structured in the following manner: Chapter 2 presents the State-of-the-art survey, or literature review, conducted during the time of this research. It starts with the motivations behind PCG and how it can help automate content creation, in general. It then instantiates on creature PCG and bridges this to believability. With some information on the background behind this concept, it focuses on how the state-of-the-art attempts to create, and measure, believability in videogame agents and links this to back to creatures. This is accompanied by reflections on the potential gaps found during the literature's review, framed by our research goals.

The latter are detailed in Chapter 3. It lists the objectives responsible for structuring this research (in the form of research questions), its methodology and planning. In the planning section, the research's timeline and the produced artifacts are outlined.

Following the definition of the problem, present in the previous 2 chapters, the empirical research effort starts at Chapter 4. This details the initial steps made and how these helped shape the overall research. Mainly, how the construction of 2 prototypes, built towards an automated creature design system, informed the creation of a tool meant to measure believability. This tool is then instantiated, in Chapter 5, in the form of a psychometric scale. This chapter depicts the iterations the scale underwent in order to be used in a practical context.

This context, as explained in Chapter 6, consists in an exercise where a software videogame is produced, guided by the scale outputted in the previous chapter. This videogame integrates a generative system capable of producing instances of creatures and players are prompted to evaluate them. This experiment is then used to verify some of the initial results obtained during the scale's early administering. Additionally, the videogame was built as a proof-of-concept to illustrate how the scale can guide the design of 'believable creatures'.

From all of these research efforts, we established the initial steps towards a systematic study of believability in videogame creatures. Furthermore, this offered us a new perspective on how creatures can be designed. These conclusions, extended in Chapter 7, are built from the thesis' chapters' several discussions, where we reflect upon the implications each step had towards the knowledge base. Moreover, we linked these to the research questions asked in Chapter 3. Finally, we end this thesis with a future work introspection. For example, how can our effort help stimulate discussion on creature creation PCG systems and how can creature believability be extrapolated to other media.

Chapter 2

Literature Review

Before deconstructing creatures as a designable artifact, and to help bind the research's scope, we must first understand the semantics behind the word. As said in the former chapter, our goal lies in designing living beings that are more proximate to 'lesser animals' (Darwin, 1889) than actual humans.

Cambridge Dictionary defines 'Creatures' as either "any large or small living thing that can move independently" or "(...) a life form that is unusual or imaginary" (*Cambridge Dictionary* 2018). However, these definitions are too vague. So much so that any human, or even humanoid, can be considered a creature.

With this in mind, we introduce our own definition. We define 'Creatures' as zoomorphic entities (with an animal, or animal-like, form), inspired by (contemporary or otherwise) living or fictional beings (Borges and Guerrero, 2002), that lack the ability to speak verbally, high cognitive functions and resemblance to a human being, even if they may possess other vaguely anthropomorphic features (i.e. having human-like proportions and being emotionally expressive). While broad, it is specific enough to limit the amount of possible living beings to a subset, excluding any thing resembling a human. The reason for this separation, revolves around the fact that humans, or humanoids, in videogames, take on the role of characters, and have their own predefined set of expectations. Creatures, on the other hand, populate the videogames' fictional worlds, sometimes without actually contributing to the videogame's narrative. For this to work, creatures must be 'believable' and this concept plays an integral part in improving player experience through engagement (Warpefelt, Johansson, and Verhagen, 2013).

In videogame development, particularly that of 3D videogames, we can consider creatures as the interpreted results of an audiovisual composition of several elements including a mesh (or 3D model), animations, textures, behaviors and sound. Usually, these elements are all simultaneously present when a creature is in a videogame, albeit with some exceptions (i.e. using only audio cues, as part of a soundscape, to imply a creature's existence/presence).

Traditionally, these elements are either produced by individuals or a multidisciplinary team, both with the aid of specialized editing tools (Martinho, Santos, and Prada, 2014). Nonetheless, as technology advances, players begin to demand increasingly complex content (Hendrikx et al., 2013). This comes at a cost, as assets require time and man-power to produce. Consequently, this causes videogame development budgets to rise so publishers, and investors, pressure developers for more conservative design choices in order to produce profitable products (Hendrikx et al., 2013; Togelius et al., 2011a). Seeking to improve the content creation process, there have been attempts to partly (Hendrikx et al., 2013; Yannakakis, Liapis, and Alexopoulos, 2014; Khaled, Nelson, and Barr, 2013), or even completely (Hendrikx et al., 2013; Togelius et al., 2011a; Yannakakis, 2012), automatize it. This is named Procedural Content Generation and it can be applied to a wide variety of content, including creatures.

2.1 Procedural Content Generation

Procedural Content Generation (PCG) has its roots in 'process intensity', a concept introduced by Crawford (1987). In his essay, Crawford (1987) defines this concept as the "degree to which a program emphasizes processes instead of data" (Crawford, 1987). He adds this can be used to determine the utility of any piece of software, particularly videogames, running on a computer: the more 'process intense' a piece of software is, the more valuable it is perceived. This is because such software takes advantage of the computers' ability to operate upon data. At its core, PCG can be viewed as software used to process data in order to generate videogame content. This is even latent in its definition.

PCG is defined as "the algorithmic creation of videogame content with limited or indirect user input" by Togelius et al. (2011a) and Togelius et al. (2011b) and Shaker, Togelius, and Nelson (2016). They further describe content as any videogame component including sounds, 3D models, textures, levels, etc.

However the authors explicitly exclude entity behavior. Yet, there is no apparent consensus when considering the latter as content (while Rosenkind, Winstanley, and Blake (2012) argue in favor, Togelius et al. (2011a) and Yannakakis (2012) argue against). Following the former line-of-thought, we consider content to be any artifact experienced by the player and thus, if behavior has a visual and audio feedback, it is content, and if it was created with "limited or indirect user input" (Togelius et al., 2011a), it is also procedurally generated. Moreover, discarding entity behavior, from the procedural generation equation, may impact other types of content as it is argued it plays a key role in supporting animations (Cerezo, Pina, and Serón, 1999; Multon et al., 1999).

Commercially, there have been three main instances where PCG is used. During its infancy, it was considered a necessity to overcome hardware limitations. Games including Elite (D. Braben and I. Bell, 1984) and Starflight (Binary Systems, 1986), both space exploration videogames with an ambitious design, applied PCG to generate entire universes that could otherwise be impossible to fit on 1980s computers' memory. While nowadays, due to hardware advances, this is not as popular as it were, in 2004, and 2006, .kkrieger (Farbrausch, 2004) and RoboBlitz (Entertainment, 2006) respectively, used PCG to reduce their files' size.

Another commercial application, and contemporary to the former one, is to aid in improving videogames' replay value. While videogames as old as Rogue (A.I. Design, 1980) (who gave birth to the genre 'rogue-like') capitalize upon this (by generating different levels during each playthrough) it was only when Diablo (Blizzard North, 1996) was released that this technique was popularized. Recently, several videogames, with 'procedurally-powered infinite playthroughs', have been appearing, with Minecraft (Mojang, 2011) and No Man's Sky (Hello Games, 2016) being some of the most notorious examples.

The third application revolves around specialized design tools. These tools usually unload some repetitive aspects of content creation, allowing human designers to focus on more creative tasks. Some examples include Euphoria (NaturalMotion, 2008), an animation tool based on the work of Sims (1994), SpeedTree (Interactive Data Visualization, Inc., 2013), a L-Grammar based tree generator and Houdini (Side Effects Software, 2015) a 3D animation software supporting a node-based visual language, capable of automating visual effects.

In research, PCG follows a slightly wider variety of directions going from algorithms to new types of generatable content and field applications. As such, several taxonomies have been proposed, throughout the years, as a means to group them. Each taxonomy offers a different perspective to sort these directions.

As an example, Togelius et al. (2011a) present a taxonomy that categorizes techniques according to their morphology and application: content created during the videogames (online) or before its development (offline); content required to complete the videogames (necessary) or to supplement it (optional); generators receiving, as input, random seeds or a parameter vector; stochastic or deterministic generation; and, algorithms that include an evaluation function (generate-and-test), or not (constructive).

Hendrikx et al. (2013) offer a different perspective in analyzing PCG. Instead of grouping techniques, they developed a hierarchical taxonomy for the generated videogame content: its base starts with Game bits (the videogame's most elementary units including textures, sounds, etc.). This is followed by Game Space (the world where the player navigates), Game Systems (the simulation of complex systems), Game Scenarios (the logical sequence of the videogames' events), Game design (the rules and goals) and Derived content (side-products of the videogame's world including leaderboards and news).

Another approach is the one made by Khaled, Nelson, and Barr (2013). They developed a set of design metaphors to help understand the relationships between designers and PCG. These metaphors include Tools (mechanisms meant to achieve design goals and empower the designer), Materials (artifacts procedurally generated that can be edited by designers), Designers (algorithms that aim to solve game-design goals) and Experts (monitors and analysts of gameplay data).

Recently, Craveirinha, Barreto, and Roque (2016) surveyed these different proposals. In their review, they deconstructed PCG as a two-moment activity (generation and evaluation) following Csikszentmihalyi's reflection on the production of human creative artifacts (Csikszentmihalyi, 1997). Furthermore, they added PCG partakes in three main roles: they can be either automated agents, 'PCG designers', replacing humans during the creation process; computer-aided design tools, or 'PCG design tools', meant to empower human creators; and, videogame artifacts, known as 'PCG games', created with, or without, the former two and encompassing in-game content generation. This distinction helped clarify the existence of three actors participating during either moments: the human designer, the computer and the player. With this in mind, Craveirinha, Barreto, and Roque (2016) framed the surveyed taxonomies according to the actors and their roles during the whole process. In the end, they point the need for a supplementary research on how humans can work, in synergy, with PCG systems. Particularly, some of the roles they identify are still underdeveloped, namely Designers as explicit evaluators and human-computer co-design.

Conversely, Togelius et al. (2013) offer a different perspective on future research directions. One more focused on the content itself. Through a reflection on future goals and current challenges in the PCG research field, the authors pinpoint the need to experiment on generating more complex content (a concept coined multi-level multi-content). This is used to identify an agglomerate of more granular content. An example used by the authors is videogames worlds: vast structures encompassing fauna, flora, and other videogame-related aspects such as quests, dialogue, etc. They cite this example as something PCG should aspire to create. However, while they state Dwarf Fortress (Tarn Adams, 2006) as a videogame already capable of generating such worlds, they are perceived as generic, or bland. This is something Togelius et al. (2013) identify as another challenge PCG faces today. Indeed, they stress the need to develop means to allow PCG to output creative content.

2.1.1 Creativity in Procedural Content Generation

Csikszentmihalyi defines creativity as a process resulting in a novel and valuable artifact, which contributes to the evolution of culture (Csikszentmihalyi, 1997). He also introduces the notion of a Creative System, a model composed of the synergy between the artifact's author, its domain (or set of symbols and rules) and its field (or group of 'gatekeepers'). Additionally, he states that potentially creative authors must have an extensive knowledge base of the domain, practice in the domain's manipulation, motivation and, most importantly, the ability to filter ideas from a wide selection. Undergoing a creative process then becomes an iterative one where an output goes through several iterations (each including a creation and an evaluation phase) as the author manipulates and evaluates his/her intermediate results.

Offering a computational approach, Boden (2009) augments this notion by diving it into three types: combinatorial creativity (the association of ideas that are not directly linked), exploratory creativity (the traversal of a well-defined search space) and transformational creativity (a meta-exploratory approach that involves exploring variations in rules and constraints that define the search space itself). These concepts were later formalized by Wiggins (2006) through the creation of a mathematical model (a tuple definition the conceptual universe in addition to its traversal rules) for a framework of search algorithms, generalizing upon traditional AI searching methods. These ideas, that computers can output creative artifacts, form the fundamentals of the Computational Creativity research field.

While Computational Creativity studies several approaches to create artistic artifacts (music (Cope, 2006), paintings (Machado, 2006), poems (Colton, Goodwin, and Veale, 2018), solving problems creatively (Hélie and Sun, 2010) to name a few examples) and metaphysics (Colton, Mántaras, and Stock, 2013; McCormack and d'Inverno, 2012) that go beyond the scope of this thesis, there are, nonetheless, attempts to bridge the gap between this field and PCG. As stated previously, one of the current challenges PCG faces is, still, its current tendency to generate relatively bland content as 'PCG algorithms are rarely classified as creative' (Yannakakis, Liapis, and Alexopoulos, 2014). Yet, creative artifacts are assumed to be preferred to non-creative ones (Dormans and Leijnen, 2013).

This is true, even though searching algorithms have been present throughout the PCG literature (Togelius et al., 2011a; Liapis, Yannakakis, and Togelius, 2013a; Liapis, Yannakakis, and Togelius, 2012; Dahlskog and Togelius, 2014; Togelius, Justinussen, and Hartzen, 2012). Nonetheless, there are several works where creativity models were used to enhance PCG (Ribeiro et al., 2003; Cook and Colton, 2012; Cook, Colton, and Pease, 2012; Dormans and Leijnen, 2013). Furthermore, Merrick et al. (2013) offer some insights in how this could be better achieved as they state "automated videogames design requires not only the ability to generate content, but also the ability to judge and ensure the novelty, quality and cultural value of generated content." (Merrick et al., 2013).

Contrary to the previously cited works, where authorship is centered around computers, there are approaches more in line with the future directions presented by Craveirinha, Barreto, and Roque (2016). These attempt to foster creativity, through a cooperation between humans and computers, where the latter plays the role of a colleague (Lubart, 2005). Implemented as design tools, some examples include Tanagra (Smith, Whitehead, and Mateas, 2010; Smith, Whitehead, and Mateas, 2011) and Sentient Sketchbook (Liapis, Yannakakis, and Togelius, 2013b) for level design and the Authorial Game Evolution tool for game design (Craveirinha, 2018). The paradigm shift towards a more participating human designer gives form to human-computer interactions where humans and computers can iterate each others' intermediate artifacts in order to produce a final output.

All in all, introducing creativity as an integral part of PCG systems, and algorithms, may be a potential solution to overcome the perceived blandness that PCG is notoriously known for. Indeed, there are some works studying different implementations, and their respective implications. These however, seem to cover only a few subsets of content, leaving out creature design.

2.1.2 Procedural Content Generation in Creature Design

As mentioned, creatures, as 3D videogame assets, are compositions of 3D models (we will refer to these as morphologies, from now on), animations, textures, behaviors and sound. This is because they are usually synthetic actors, present in the videogames' worlds, taking on the roles of either 'Adversaries' (Treanor et al., 2015) or 'Spectacles', meant to 'bring life' to their environments and provide a richer context for the player to get immersed in (Treanor et al., 2015). So, as far as synthetic actors go, Cerezo, Pina, and Serón (1999) argue that behaviors and animations are key elements to convey life to them. Although a composition of these elements is necessary in creature design, most literature in PCG tackles each element individually. Indeed, Hendrikx et al. (2013) propose creatures as complex content (multi-content (Togelius et al., 2013), in fact), stating their generation, as whole, still requires further research. So, the following subsections illustrate state-of-the-art approaches to generate morphologies, animations and behaviors.

Procedurally Generated Morphologies

Most of the work present in the literature follows a grammatical approach: assuming a creature can be well defined, morphologies are constructed through means of a set of rules, using individual pieces. This is similar to the methods used in other types of content such as vegetation and buildings (Smelik et al., 2009; Shlyakhter et al., 2001; Tobler, Maierhofer, and Wilkie, 2002; Martin, 2005; Mathias et al., 2011), resulting in a shape composed of smaller ones, linked together.

One of the most fundamental works in this area is that of Sims (1994): Through means of an evolutionary process, morphologies are created from the transcription of a grammar graph. While results are mere geometric shapes, this work inspired others, depicted in this chapter.

Simmons, Wilhelms, and Van Gelder (2002) present a mechanism to generate animal models by altering a preexisting one, named the canonical model. They instantiate this by feeding the system a horse model where variations, including a pony, were originated. However, the authors point out this system cannot create creatures autonomously as canonical models need to be input *a priori*.

Possibly one of the few examples, in creature design, using explicit creativity models, is the one proposed by Ribeiro et al. (2003). By using Divago (Pereira and Cardoso, 2000), a conceptual blending (Fauconnier and Turner, 1998) tool, they were able to generate creatures from two conceptual maps, a descriptive script and a preexisting database of decomposable morphologies. The resulting output could then be used as is or fed to the tool's database.

Directly inspired by computer generated architecture, Ilčík et al. (2010) developed extended shape grammars for describing mophologies. This extension was used to develop organic creatures, among various objects. Though, they concluded the grammars still neglected some elements including preservation of mass, extreme positions and collisions' resolution.

Attempting to introduce an interactive design tool (Takagid, 2001), Hudson (2013) devised an evolutionary process to generate rigged creatures. Using Houdini (Side Effects Software, 2015), Hudson evolved structures to create creature sets, from which users could choose to evolve further and even crossover with other evolved creatures to output new ones.

Guo et al. (2014) proposes a process where creatures (named monsters by the authors) are originated from the deformation of preexisting ones (similar to Ribeiro et al. (2003) and Simmons, Wilhelms, and Van Gelder (2002)). This works by decomposing the inputs into a group of shapes and then applying transformations to them. Experiments were conducted to validate the system but, while results showed the monsters inspired human designers' works, the software occasionally combined incompatible parts.

Procedurally Generated Animations

The surveys conducted by Multon et al. (1999) and Cerezo, Pina, and Serón (1999) identify two main approaches for procedural animation: *kinematics* and *dynamics*. While both these methods can be expressed mathematically, they differ from each other. On one hand, *kinematic* methods comprise algebraic transformations in vectors, symbolizing bodies' positions and rotations. *Kinematic* can be direct, or forward, when this transformation is done overtime or they can be inverse, or backwards, when it is performed from a goal state back to the starting one (Zhao and Badler, 1994). On the other hand, *dynamics* deal with physics, more specifically, how forces result in motion and they are argued to yield more realistic results (Cerezo, Pina, and Serón, 1999; Multon et al., 1999). Akin to *kinematics*, *dynamics* can also be inverse or direct, depending on whether the final position is known or not, respectively. Additionally, there are also hybrid methods combining *kinematics* with *dynamics* (Shapiro et al., 2007; Horswill, 2009; Shapiro, Pighin, and Faloutsos, 2003), argued to yield better results than using one of the aforementioned methods individually.

Both *kinematic* and *dynamic* motions can be applied to a skeleton (Skrba et al., 2009). In animation, a skeleton (a name that comes from an analogy with anatomy), or rig (Multon et al., 1999), is a structure composed of linked bones, where each link forms a joint. These bones, in turn, apply transformations to groups of the morphology's vertices, whenever they perform a motion. Cerezo, Pina, and Serón (1999) also add that procedural animations can be driven by *controllers*. At low-level, these controllers direct the method used (whether it is *kinematic*, *dynamic* or *hybrid*) to make the necessary calculations in order to perform a given motion.

The use of controllers implies a relationship between behaviors and animations and, indeed, Cerezo, Pina, and Serón (1999) and Multon et al. (1999) argue these play a key role in agents' animations.

As a final note regarding procedural animation techniques, Multon et al. (1999) and Cerezo, Pina, and Serón (1999) cite the previous approaches as the main ones, there are still other methods, including *mesh deformation* (Cerezo, Pina, and Serón, 1999; Sumner et al., 2005).

In creatures, procedural animations can be either generalist or more geared towards specific morphologies. In case of the former, Sims (1994) evolved controllers to generate motions adapted to given environments, through the use of forward dynamics. This work inspired others, including an animation editing tool (NaturalMotion, 2008), a simulation program (Lee Graham, 2007) and the work of Wampler and Popovic (2009). In the latter, optimization methods are used to generate gaits corresponding skeletons of legged animals, though the authors concluded they were not able to capture all possible kinds of gaits. Still on the topic of controllers, Coros, Beaudoin, and Panne (2009) introduced a system allowing agents to adapt their locomotion whenever it was disrupted.

Contrary to the previous works, Spore (Maxis, 2008), a 'God videogame', makes extensive use of procedural animations albeit without the need of controllers. Hecker et al. (2008) explains the system behind Spore as one using a sequencer and reverse kinematics. By manually creating a few animations for known morphologies, the system can then generalize them so they can be sequenced accordingly. Nonetheless, Hecker et al. (2008) states this system still relies on human testing to ensure that the resulting animations work correctly.

There are other procedural animation approaches focusing on specific morphologies. Examples include the work of Cenydd and Teahan (2013) for arthropods, Coros et al. (2011) for quadrupeds, Geijtenbeek, Panne, and Stappen (2013) for bipeds and Tu and Terzopoulos (2000) for fishes. All of these works also employ a skeletal-based approach with either kinematics or dynamics. However, regarding the latter, Nava (2017) presents an alternative way, one relying on mesh deformations, to animate fishes. In Abzû (Giant Squid Studios, 2016), an aquatic exploration videogame, the fishes' morphology was deformed through mathematical equations. This way, developers were able to add a high number of fishes without the overhead of traditional skeletal-based animations.

Procedurally Generated Behaviors

There are two main approaches to develop behaviors: one, traditionally used in most videogame genres, named Agent-Based, involving the creation of agents as independent entities with individual goals and objectives. In 1984, Braitenberg (1984) introduced the concept of synthetic approach to build agents in order to model animal-like beings, named Braitenberg Vehicles. In his work, Braitenberg starts by building simple moving beings and ends by giving them the means to produce more complex constructs such as concepts.

Years later, a systematic study, in this approach, was made by Epstein and Axtell (1996) where the authors coined the method as Bottom-Up and concluded it to be suitable for studying emergent macrostructures from the interactions between individual agents and with their environment. The other approach, more used in tactical-heavy genres, is known as System-Based. Essentially, it follows the metaphor of a 'puppet master' where the master controller issues commands to groups of agents. This latter approach, however, is beyond the scope of this thesis as the Agent-Based approach is generally employed in the context of creature behavior modeling (Millington and Funge, 2009).

Grand, Cliff, and Malhotra (1997) illustrated the architecture behind the agents in Creatures (Creature Labs, 1996; Grand and Cliff, 1998). These agents, or Norns, are comprised of a multilobe neural network. Each lobe is responsible for performing one of the following tasks: decisions, perception and conceptualization. In addition, the neural network is fed by an artificial hormone system and sensors allowing them to develop instinctive behaviors and learn. In the end, Grand, Cliff, and Malhotra (1997) argue that emergent social behaviors were perceived.

In order to promote the authoring of emergent behaviors through planning, Mateas and Stern (2003) and Michael Young et al. (2004) developed a language and an architecture, respectively. On one hand, the language, ABL, supports the definition of behaviors that can even be run in parallel, all of which to achieve a given goal. On the other hand, the architecture, Mimesis, can be integrated into videogame engines so that engines generate plans according to a given goal. Nonetheless, both approaches were developed with believable agents in story driven environments in mind. The ABL language was even used in Façade (Procedural Arts, 2005; Mateas and Stern, 2005), an interactive drama videogame.

In the videogame Black and White (Lionhead Studios, 2001), the player has the possibility to influence the behaviors of autonomous creatures. This is done through a supervised learning mechanism in which players can reinforce positively, or negatively, the observed creatures' actions (Millington and Funge, 2009). Particularly, whenever a creature performs a certain action (sleeping, attacking, eating, etc.), players can pet or spank it. This, internally, creates a positive/negative bias to promote/inhibit those actions, respectively.

Santos and Roque (2010) provided some initial work in procedural behaviors with Petri Nets, where they created an architecture. These are reorganized after each playthrough in order to give birth to new behaviors, in a Pac-man-based videogame. In their work, Santos and Roque (2010) compare a strictly random method to reorganize behaviors with a data-fed technique that changes the Petri Nets' topology according to gameplay metrics. Results indicated that, in general, Petri Nets have potential for behavior modeling and, in particular, the latter technique ensued a better videogame experience.

2.1.3 Discussion

Automating videogame content creation has been a long-lasting research. As the literature suggests, it still faces several obstacles. On one hand, PCG studies still focus on more granular types of content, leaving out others, more complex compositions as future directions. From our review, in creature PCG, only one work tackled several aspects at once (Sims, 1994). In fact, even other surveys conducted stress this out (Hendrikx et al., 2013). While, from a problem-solving perspective, it seems logical to engage in such manner, proceeding towards bridging the gap between each individual components is still something that will eventually need to be solved. Nonetheless, solving this problem will lead closer to other ones PCG currently faces. So, on the other hand, even with a wide variety of algorithms for many different contexts, PCG is still considered 'uncreative' (Yannakakis, Liapis, and Alexopoulos, 2014).

By deconstructing theories behind the creative process, it becomes evident the pivotal importance of evaluation methods. However, in current research, these evaluations take on the form of merely functional ones as it is present in the works surveyed. For instance, in morphologies, outputs follow an implicit evaluation by simply adhering to the grammar's rules. Resulting creatures are then judged by either their apparent 'interesting look', or variety. The same is true even when creativity models are used (Ribeiro et al., 2003). Analogously, procedural animations are evaluated by their similarity to their real life counter-parts or if they manage to convey the actions they are trying to portray. Procedural behaviors in creatures, on the other hand, even lack any form of evaluation as emergent behaviors are merely detailed as added value.

We argue that to ultimately evaluate a creature, humans, or computers, must look beyond its elements' functionality. Firstly, creatures are more than just their individual elements. Indeed, when experiencing a creature, it is usually its whole composition and, as such, evaluation should take that into account. Secondly, in media, in general, and in videogames, in particular, they exist within a context. The latter defines their role, and interpretation, inside the media artifact. Detracting from its intended role may cause a creature to be perceived poorly. So, with this in mind, a potential evaluation method should be able to provide a context-driven framing.

We conjecture such framing is possible when we assume the creature's most nuclear role to be that of supporting the artifacts' ability to suspend the disbelief of the users consuming it. Therefore, we propose an evaluation method should indicate how creatures fare in this support or, in other words, how they are believable.

2.2 Believability

Fogg and Tseng (1999) define believability as a construct measuring the perception users have of a given artifact's credibility. By considering an artifact as a source of information, users are willing to trust its information if they perceive it to be credible, in spite of the information's actual veracity. While the authors focus their studies on computers, and their software products, we argue that any artifact capable of conveying information, regardless of its technological backbone, falls within the definition of Fogg and Tseng (1999). This, of course, includes, by definition, any type of media artifacts.

2.2.1 Believability in Media Artifacts

In media artifacts, believability is tightly coupled with the concept of 'Suspension-of-Disbelief'. This was first introduced by Coleridge, in his autobiographic reflection, where he observes this notion's importance, in literary works, when he states "(...) my endeavours should be directed to persons and characters supernatural, or at least romantic, yet so as to transfer from our inward nature a human interest and a semblance of truth sufficient to procure for these shadows of imagination that willing suspension of disbelief for the moment, which constitutes poetic faith" (Coleridge, 1817). By deconstructing this sentence, we can identify the elements necessary for 'suspending disbelief'. First, Coleridge notes the agency in the act. This suspension is not done automatically but rather, people must be willing to do it, even at a subconscious level. Second, this is framed within an artifact's bounded veracity as Coleridge states it is only required a 'semblance of truth'. Finally, the act happens in a contextualized environment, or 'moment'. It then becomes evident that 'suspension-of-disbelief' works paradoxically: we are willing to believe in something, that is unbelievable, for the sake of its context.

Tolkien (1947) elaborates this concept further. In his essay, he reflects upon the elements leading up to the acceptance of something unbelievable, in the context of a narrative. He starts by claiming literary media exist in a 'Secondary World', one "which your mind can enter" (Tolkien, 1947). Inside this world, a supposed mental model of the narrative, laws are set, by its creator, and, as long as a subject is experiencing it, he/she abides by them. However, it is possible for disbelief to happen and as "the spell is broken" (Tolkien, 1947), Tolkien argues it is then necessary to resort to suspending disbelief. In fact, he states that a well constructed 'Secondary World' does not need such suspension. He stresses this further when dealing with themes "not to be found in our primary world at all, or are generally believed not to be found there" (Tolkien, 1947), by stating the need to go beyond mere descriptions of the unreal but rather constructing a framework to support it. These themes, coined 'Fantasy', are present throughout several media, from cinema to videogames. So, it should be safe to assume, they share the common need for frameworks supporting a believable 'Secondary World'.

In psychology, studying how the previous phenomena manifests itself in technology-based media artifacts gave birth to the concept of presence. Defined as the "the perceptual illusion of nonmediation" (Lombard and Bolmarcich, 1997), presence, formally known as telepresence, is a phenomenon which causes a person to stop acknowledging the existence of a medium, responding as if it was not there. Lombard and Bolmarcich (1997) explain this illusion can occur in two ways: either the medium is considered invisible, acting as a window to the medium's content, or it can seemingly transcend its nature to something else, such as a social entity.

In their work, Lombard and Bolmarcich (1997) categorize 6 different approaches used to conceptualize presence. These include:

1. Presence as social richness - This accounts for the degree which a medium is perceived "as sociable, warm, sensitive, personal or intimate when it is used to interact with other people" Lombard and Bolmarcich (1997). Measure this is dependent on the user's subjective judgment of the medium's interaction and the task at hand.

2. Presence as realism - This relates to how a medium accurately depicts real-life elements (objects, events and people). This is divided further into 'social

realism' and 'perceptual realism'. The former deals with how accurate the medium depicts these elements while the latter has to do with how credible an event may be even if they do not exist in the non-mediated counterpart.

3. Presence as transportation - This conceptualization involves the confluence between an user and its consumed media, either by 'transporting' the user to the media's world, the media's world to the user or two communicators to a place they share.

4. Presence as immersion - In this perspective of presence, the user's senses are immersed in a virtual world, more so when the experience is compelling. This conceptualization has its roots in early virtual reality, where the users' senses are made to focus on the experience, through means of technology (head-mounted display, headphones, gloves, etc.). Lombard and Bolmarcich (1997) add this has a psychological component manifested in users feeling involved, engaged or engrossed in the media. They elaborate stating the best way to measure this is through self-reports.

5. Presence as social actor within medium - In this concept, media personalities are perceived as a social actor, regardless of the fact the 'interaction' is mediated. The authors illustrate this point with Petz (PF Magic, 1995): a videogame where players take care of virtual animals and, though they are virtual, the users act as if they are real. Lombard and Bolmarcich (1997) explain the users' perception "lead them to illogically overlook the mediated or even artificial nature of an entity within a medium and attempt to interact with it".

6. Presence as medium as social actor - Unlike the previous perspective, this one involves reactions to signals provided by the medium itself, instead of entities within it. It is, once more, suggested users disregard the mediated nature of an experience and, as such, they treat the medium as a social entity.

The concept of presence seems to indicate that, regardless of a medium's platform (and whether it is powered by technology or not), users may have the tendency to ignore its 'barriers', given the perceptions they attain from its elements, acting as if their mental model of the mediated world, and the non-mediated one, co-exist.

2.2.2 Believability in Videogames

Following the survey conducted by Lombard and Bolmarcich, 1997, Mcmahan (2003) propose presence to analyze 3D videogames. Indeed, Warpefelt, Johansson, and Verhagen (2013) argue videogames have a need to promote presence. As believability can be an element to help induce 'suspension-of-disbelief' and, consequently, presence, it is considered a crucial quality of their world's actors, leading players to accept them as 'alive and thinking' (Rosenkind, Winstanley, and Blake, 2012). This is further supported by studies who have shown believable actors are perceived to be more engaging than non-believable ones (Rosenkind, Winstanley, and Blake, 2012; Hingston, 2011; Warpefelt, 2016; Bates, 1994; Togelius, Yannakakis, and Karakovskiy, 2011; Hudlicka, 2008; Lim et al., 2009). As such, it is assumed to be one of the driving factors when crafting videogame experiences (Warpefelt, 2016). With such importance, there have been several studies in how to craft believable actors.

Believable Agents

Studies in believable actors mainly follow two directions: one studying how to create behavioral patterns which, when observed, incite believability. This is the scope of the Believable Agents field (Hingston, 2011). The other, has its roots in animation where artists are concerned with creating life-like representations of beings, then known as believable characters. Here, anthropomorphism was key to make viewers relate to them (Bates, 1994). Nonetheless, the literature behind both directions appears to focus extensively on simulating humans. On one hand, as either in-game characters, or players (Togelius, Yannakakis, and Karakovskiy, 2011) and, on the other, as compelling and emotional characters, part of a narrative (Bates, 1994). It is worth noting however, that believability is in no way a synonym for realism. In fact, some of the construct's earlier studies were conducted with cartoons (Thomas and Johnston, 1997).

However, even in the field of Believable Agents, believability, as a construct, is still in its early steps (Togelius, Yannakakis, and Karakovskiy, 2011). First, there is no actual standard definition for believable agents (Lucas et al., 2012). Second, works such as the ones proposed by Loyall (1997), Bates (1994) and Mateas (1997), transcribed their criteria from the same ideals upon which believable characteres are constructed. Specifically, they derive their foundations from the principles described in 'Illusion of Life: Disney Animation' (Thomas and Johnston, 1997). So, according to these authors, to be believable, agents must:

- Have Personality
- Have Emotions
- Have Self-Motivation
- Be able to perform Parallel actions
- Engage in social Behaviors
- Behave consistently

These factors are evident in several agent architectures and algorithms (Warpefelt, 2016; Rosenkind, Winstanley, and Blake, 2008; Riedl and Young, 2005; Parenthöen, Tisseau, and Morineau, 2002; Magerko, 2007; Umarov and Mozgovoy, 2012; Reynaud, Donnart, and Corruble, 2014; Ochs, Sabouret, and Corruble, 2010; Hudlicka and Broekens, 2009).

Another point is that, while synthetic humans, and humanoids, are abundant in videogames, they are not the only type of existing virtual 'living-entities'. There is an overwhelming focus on the first (Bates, 1994) (possibly due to the way people tend to anthropomorphise media artifacts (Reeves and Nass, 1998)), despite the existence of other animals. Even videogames focusing more on animals, follow this anthropomorphic trend (Frank, Stern, and Resner, 1997).

Despite the field's anthropocentric approach, most of the works surveyed are also biased towards generative solutions. These are offered as is without any attempts to evaluate their results: as long as they meet each criterion, it is assumed they are considered automatically believable. Still, there are some perspectives on agent believability who offer methods of validation.

Evaluating Believability

Taking into account the six conceptualizations of presence, Lombard and Ditton (2000) propose a way to measure them: since they argue that presence is subjective, deriving from an user's mental model of his/her perception, their assessment tool comes in the form of a questionnaire, following other similar studies. This work managed to group the several perspectives in a unified questionnaire prompting participants to answer their individual tendency to suspend disbelief and their responses to five, of the sixth, presence conceptualizations present in the literature: presence as social richness, realism, transportation, immersion and social actor within medium.

In videogames, Togelius, Yannakakis, and Karakovskiy (2011) offer a reflection on how to assess Believability. In their review, they distinguish three approaches in evaluating believability: either objectively, by measuring physiological responses, by collecting statistical data embedding behavioral responses during gameplay, or, subjectively, through self-reporting. While they argue believability can be assessed through the combination of the three aforementioned approaches, the remainder of their review focuses solely on the latter as they argue it can yield accurate results.

Regarding subjective approaches, they provide several suggestions. First, they claim this approach can be measured through free-responses or as questionnaires, pointing to the ability to gather rich information on the former and easiness to analyze. On the latter, the authors provide three alternatives on how data can be collected on questionnaires: as Booleans, rankings or ordered sets. They then follow-up with a reflection on when these reports can be administered. While they state there is no standard accept time window, they argue in favor of administering them after subjects have been exposed to the media. Togelius, Yannakakis, and Karakovskiy (2011) also question if whether subjects should be playing the videogame or observing it being played, arguing the former may cause subjects to sidetrack away from their goal. They also question the amount of time required between sessions to allow for an optimal assessment. Regarding this, the authors claim, it is dependent on the videogame's genre and its context. Particularly, to the evaluated agents' on-screen presence. Finally, they reflect on how believability can be best represented. While they do not provide a definitive answer on this account, they suggest states, continuous values or vectors (similarly to how emotions are represented through arousal and valence).

Using a scale to evaluate an aspect of believability has been previously done by Arrabales, Ledezma, and Sanchis (2011). In their work, they use ConsScale (Raúl, Ledezma, and Sanchis, 2010) to assess, in a competitive multiplayer videogame, whether several agents behave as if they were controlled by humans (similar to a Turing Test), considering them believable if they do so. ConsScale is a tool designed to quantify the 'level of consciousness' an artificial agent appears to have, by analyzing its architecture. It is composed of 13 levels where each contains a set of statements illustrating the cognitive skills an agent must have. These levels go from 'Disembodied' (entities lacking defined boundaries or cognitive skills) to 'Super-Conscious' (entities surpassing human-beings, capable of managing several consciousness simultaneously). By scoring the agents' cognitive functions accordingly, authors were able to compare how each agent fared and attempt to correlate this with their perceived believability. However, this work fails to range other types of believability because agents are judged as avatars for players rather than simulated living beings.

The aforementioned works appear to focus extensively on self-reporting. This is how they attempt to measure a construct. Since, by definition, these are closely tied to a person's mental model and are subjective so, measurements are done indirectly through a person's attitudes towards something. There are several tools to this extent. One of these are Likert Scales (Spector, 1992). Commonly used throughout social sciences, Likert Scales offer a means to assess people's attitudes. Spector (1992) lists several applications for these scales including measuring emotional states, personal needs, personality, job descriptions, etc. Outside of social sciences, this type of technique was proposed by Togelius, Yannakakis, and Karakovskiy, 2011 to assess believability. In fact, in videogames, Likert Scales, or their generalized form rating scales, have been present in evaluating other types of constructs, such as player experience (Wiemeyer et al., 2016).

2.2.3 Discussion

Media artifacts, such as videogames, can be considered sources of information as players interpret a message, the designers' intent, in the form of a ludic macrostructure. In order to experience it, they must be willing to accept, or believe, the information they are provided with, even though it might challenge, or even contradict, their ideal of the real world. This strengthens the idea that believability is achieved from the mental model constructed via an observer's interpretation of a given artifact. Therefore, it becomes the job of the videogame's creators to provide the necessary elements in order to incite this believability.

One such element, we argue, is the set the actors inhabiting these videogame's worlds. Indeed, there have been several works trying to use criteria in order to enhance such actors. Yet, our state-of-the-art review suggests they are mostly used at face value. This is further underlined by the extensive usage of the empirical studies summed in Thomas and Johnston (1997). The main problem with this, is that these ideals were not challenged to the point where systematic studies were made to assess how they actually measure believability or if they do it at all. While there are some proposed evaluation methods, these mainly attempt to measure the degree which an actor is capable of simulating a human playing a game, leaving us without instruments to address the believability of non-anthropomorphic creatures.

To bridge the gap towards creature design, we will aim to consider how believability can be a possible way to evaluate creatures, since as far as could be verified, from our review, there are no such studies for non-anthropomorphic ones.

Chapter 3

Research Goals

Summing up the discussion made previously in Chapter 2, our literature review brought forward opportunities, and challenges, used to focus our effort. First, we limited our research goal to the design of videogame creatures. Creatures, akin to other type of content, have had their production costs increase in tandem with their demand. While automated solutions attempt to mitigate this, creature PCG is still an underdeveloped niche. So much so, its surveyed works branch out to solve its individual elements (morphologies, animations and behaviors) without considerable attempts to link them back together. This provided us an opportunity to study how these elements could influence each other in the final product. While not directly related, the PCG solutions' perceived lack of 'creativity' gave us insights in how this study could be conducted: through an evaluation method, specifically one where a creature manifests itself as an agent and is assessed through perceived in-game phenomena.

By reflecting upon the possible roles creatures might have in a videogame artifact, we position our research in how they can support the videogame's world. This goes in line with our previously mentioned ideal as it requires creatures to be agents and evaluated at a perceived macro-level (as part of the whole play experience). One way to achieve this is through believability; though, its literature seems to be biased towards anthropomorphic agents, offering us an opportunity to lay the first steps in non-anthropomorphic beings.

All in all, our research goal lies in developing an evaluation instrument to assess creature believability and help guide the creation of future creature assets, study how this may impact a potential procedural solutions. With this in mind, we can outline our effort in the form of a question-tree.

3.1 Research Questions

Following our previous line-of-thought, we materialize the goal by asking **"How can we empower creature designers to devise (automated) creatures through an evaluation instrument, capable of measuring creature assets in a contextualizing environment?"**. To address this aim we will propose and study the Creature Believability Scale as a possible solution. We further instantiate our goal by providing a case study, through a proof-of-concept prototype, Orphibs Evolvapalooza, as a means to study the impacts the scale can have in guiding a theoretical automated design. Our initial question is divided to a greater extent, in the following question-tree:

1. **How can we empower creature design through a contextualized evaluation instrument?** - With this question, our research centers on finding ways to give

designers the authority to design creatures, even if aided by an automated tool, using an evaluation instrument, referred to as contextualized. This means, the evaluation is done taking into account the circumstances set by a videogame.

(a) **How do the creatures' components manifest in identifiable phenomena?** - As we established, creatures, as content, comprise several components (animations, morphology and behaviors). This question tries to frame how the creatures' several components can be perceived by an observer.

(b) **How to evaluate believability of videogame agents?**

  i. **Are the state-of-art criteria enough to evaluate believability?** - This question arises from the state-of-the-art's lack of a systematic study in assessing believability. While it lists several criteria, these were taken at face value from the work described in Thomas and Johnston (1997).

  ii. **What makes a creature look and feel believable?** - Following the previous question, this one focuses on non-anthropomorphic creatures.

(c) **How can the evaluation instrument guide the creation of content?** - With the intent in empowering designers, this questions frames the research in studying how an instrument can inform the design on posterior iterations.

Throughout Chapters 5 and 6, we offer a detailed description of the research effort and how each deliverable worked towards it, whilst Chapter 7 transcribes our final thoughts on each of the questions.

3.2 Methodology

Our main question ("How can we empower (automated) creature design through an evaluation instrument, capable of measuring creature assets in a contextualizing environment?") hints the need of a method (Wobbrock and Kientz, 2016) capable of either manually, or automatically, evaluate creatures. With this in mind, we proposed an evaluation instrument aided by a tool called Creature Believability Scale. The construction of the latter required an iterative process to fine tune it. In addition, we developed a prototype (Wobbrock and Kientz, 2016) to help study how to use the tool to inform new creature design exercises. Therefore, we opted to develop our research through a Design Science Research (DSR) (Vaishnavi, Kuechler, and Petter, 2004) methodology.

The DSR methodology aims at producing knowledge through design. Following a more pragmatic perspective, DSR seeks this knowledge from the analysis, and reflection, on man-made objects ('science of the artificial'), adhering to pre-specified goals, resulting from an awareness of problem. It is worth noting however, that DSR is not analogous to simply performing a design activity (routine design). Only through the "production of interesting (to a community) new and true knowledge" (Vaishnavi, Kuechler, and Petter, 2004) can design be considered DSR. Figure 3.1 illustrates a model of the process behind DSR, as detailed by Vaishnavi, Kuechler, and Petter (2004).

The methodology begins with an **Awareness of Problem** step, comprising the definition, and identification, of a potential problem. Here, multiple sources can be

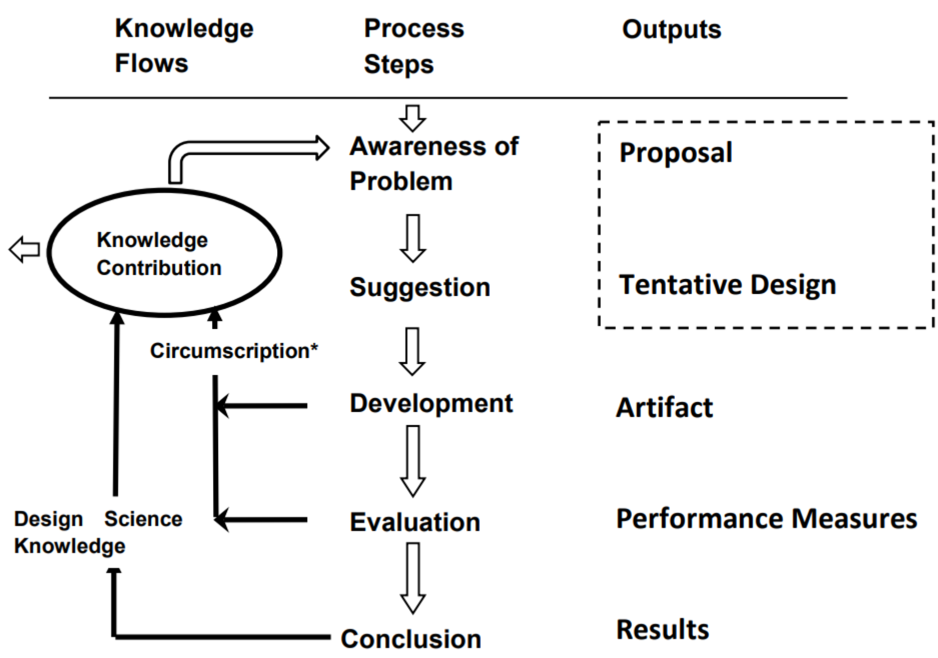


FIGURE 3.1: The Design Science Research Process (Courtesy of Vaishnavi, Kuechler, and Petter (2004)).

consulted in order to bring forth possible pitfalls or obstacles needing solutions. From this step, a **Proposal** arises, detailing a prospective research plan.

This step is followed by **Suggestion** where a possible solution is created through abduction, drawn from the knowledge gathered previously. This possible solution, or **Tentative Design**, forms the blueprints necessary to create the actual solution. In the context of Information Systems this accounts to interaction models and software architecture, or similar diagrams.

Feeding from the previous step, **Development** comprehends the solution's actual implementation. The tentative solution is transformed into a proof of concept **Artifact** and its implementation varies according to its evaluation requirements.

The next step, **Evaluation**, attempts to validate the previously created artifact. This is usually done with a desired performance model presented during the Proposal step. However, it is possible that such model does not exist yet. So, one can be developed alongside other Suggestion's artifacts. During this step, the artifact's behaviors are analyzed in an attempt to explain possible deviations from an initial hypothesis. This analysis can be input into an additional Suggestion step, forming an iterative loop as it "generates understanding that could only be gained from the specific act of construction" (Vaishnavi, Kuechler, and Petter, 2004). Loops can be created as many times as a researcher deems necessary, as long as they are justified by further analyses.

The research effort is then ended in the **Conclusion** step. This does not mean, however, the research is definitely complete but rather it is in, at least, an acceptable state to enable communication of results. **Results** are collected and summed in a statement of learning so that knowledge can contribute back to its community. This statement is materialized in the form of "constructs, models, frameworks, architectures, design principles, methods, and/or instantiations — and design theories" (Vaishnavi, Kuechler, and Petter, 2004).

3.3 Research Plan

Following the DSR methodology, we outline our research effort, taking into account its steps, as illustrated in Figure 3.2. In addition, our outline is further divided into three moments, according to their contributions. These moments are the Exploratory Phase, the Creature Believability Scale Construction Phase and the Proof of Concept, and Scale Use Study, Phase and they are detailed in Chapters 2, 4, 5 and 6. During each phase, we employed mixed method (Creswell, 2013) approaches to collect and analyze data.

In the Exploratory Phase, we conducted the initial literature review (materialized in part of Chapter 2), which helped define its foundation. In particular, we identified the opportunity of an under-explored PCG sub-field. This led to the creation of 2 prototypes, detailed in Chapter 4. These had the initial goal of producing a proof-of-concept for a system capable of generating creatures. In the end, these ultimately informed us on our initial perspective regarding the research problem.

As the previous iteration set the basis for revising the research, we conducted an additional review of the literature (Chapter 2). This introduced, and helped further understand, the concept of 'believability'. We figured this to be a direction towards the solution to our original problem. However, similarly to creature PCG, creature believability was a concept yet well addressed in the literature. As such, the second DSR iteration reported in Chapter 5 focuses on proposing an evaluation method, in the form of a psychometric scale, capable of measuring this construct. This proposal underwent 4 revisions, each accompanied by either a quantitative or a mixed-method approach. During the third revision, data was collected employing experts as test subjects. This allowed us to create an additional artifact, the creature believability spectrum, which helped further understand the role of creatures within a videogame context and how this affects their perceived believability.

With a proposed evaluation method, the third iteration of this research revolved around constructing a videogame prototype, detailed in Chapter 6, as a case study and a proof-of-concept. First, its design was guided by the previous phase's proposal. Second, its implementation incorporated the evaluation method as a game mechanic. Its conception underwent a paper prototype stage followed by a software implementation one. Each of these was succeeded by an usability test before the prototype was deployed. It was subsequently deployed online to several people. From this experiment, we collected to further assess the scale. While results suggested the scale can be further improved, the thesis concluded at this point as the data gathered throughout this research was enough to established a solid foundation for the scale's future directions (described in Chapter 7).

During each of the aforementioned phases, several intermediate outputs were created. Each of them helped build subsequent phases.

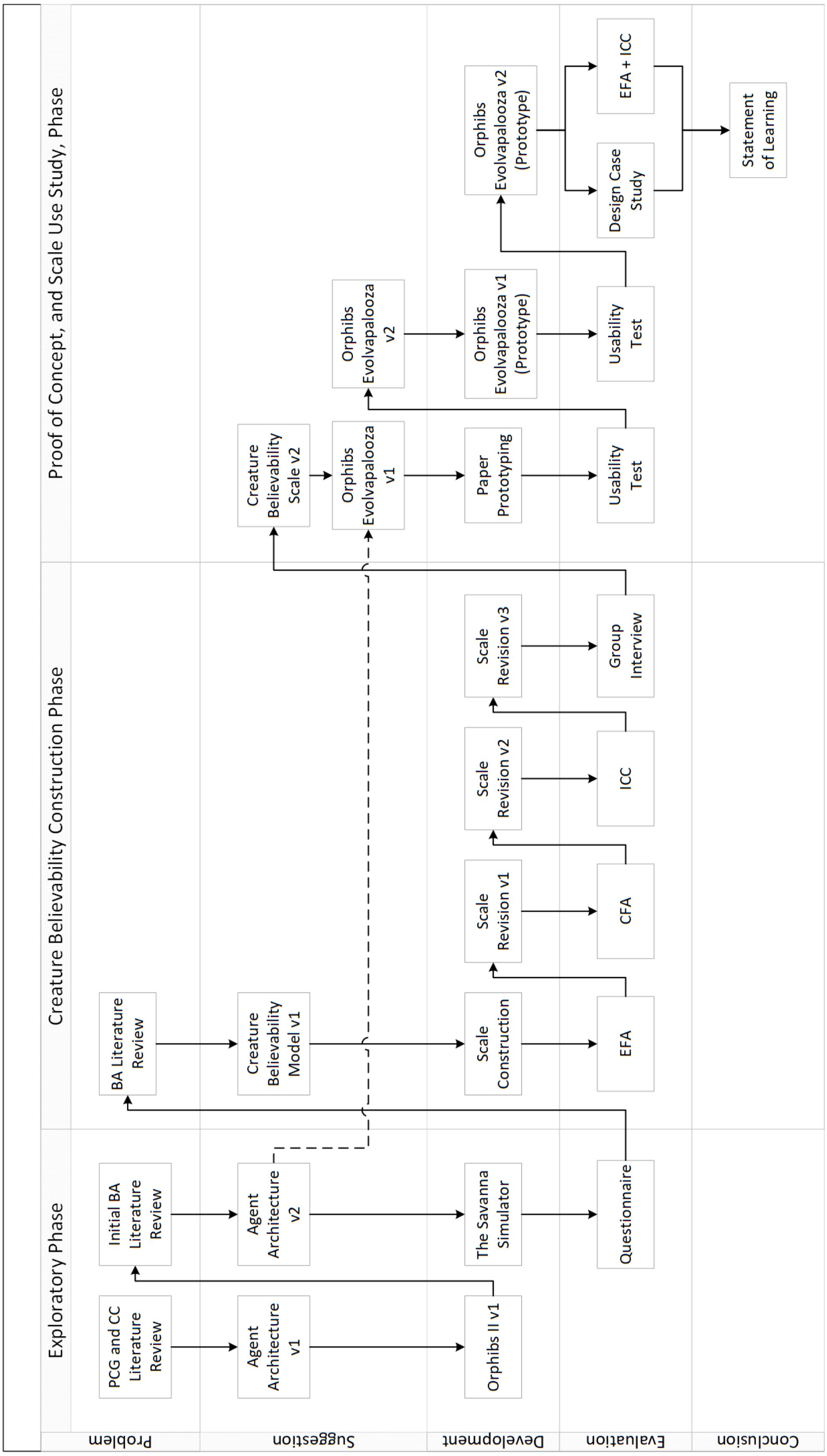


FIGURE 3.2: Plan outline with the research effort's activities framed within the DSR methodology.

3.3.1 Main Outputs of the Research

The research effort, and its encompassing phases, originated several outputs that were used either to start additional DSR iterations, or to contribute to the thesis' statement of knowledge. These include:

- **Agent Architecture** - Devised during the Exploratory Phase, this architecture was used to model the agents present in its phase's prototypes. It underwent 3 revisions: two during the first prototype and another one during the second. Though it was not directly used, this architecture would later serve as the foundation for the proof of concept's agents devised in the research' third iteration. Moreover, it was designed to support some of the elements of what we would later understand as the 'believability' construct. Evaluating its results, built the basis for the literature review which originated another DSR iteration.
- **Creature Believability Scale** - This scale's initial four iterations were constructed during Creature Believability Scale Construction Phase while it was only stabilized in the final DSR iteration. This was used to both help design and later evaluate the proof of concept later developed.
- **Proof of Concept Study** - This is another of the thesis' main outputs the final one to be originated from its research effort. The study constitutes a reflection on how to build an artifact with the goal of maximizing perceived believability in virtual creatures and how to assess this perception in players. Particularly, it studies the proof of concept created in the Proof of Concept and Scale Use Study Phase. As previously mentioned, this artifact was designed around the Creature Believability Scale and, similarly to the prototypes developed during the first DSR iteration, contains agents. Their architecture was built from the one devised in the Exploratory Phase.

Chapter 4

Awareness of Problem

While most of the work resulting from this phase did not directly contribute to this thesis' deliverables, it helped define the true nature of our research problem. Initially, our goal involved the creation of a generation system capable of producing a creature that exhibited coherence between interesting behavior and appearance. To this extent, two exploratory prototypes were developed. The problems we encountered while researching these, were what led us to reformulate our research problem.

The initial one, named Orphibs II, provided the conceptual and technical frameworks later used in further prototypes. Particularly, its agent architecture allowed the introduction of expressions. These were integral in linking animations to behaviors and brought forth believability as a potential means to evaluate creatures.

The second prototype, the Savanna Simulator, was built to extend the Orphibs II agent architecture. First, it transitioned from the original prototype's 2D space into a 3D one (as videogames' visuals have been shifting towards 3D, over the years (Roettl and Terlutter, 2018)). Second, it helped expand the concept of expressions into that of Audio-Visual performance (the audio-visual output of a given agent, or group of agents, over time). This helped better understand how we could evaluate our creatures in general. However, our initial evaluation attempt did not convey the information we needed to properly assess the creatures' quality: we concluded we were trying to assess an ill defined construct. In particular, our attempt lacked a context, a means to link expressions to believability and a way to generalize the evaluated elements so they could be used with fantastical animals. This is what further led us to research believability as a construct.

While these prototypes were not developed further, the concept behind Orphibs II and the expanded system made for the Savanna Simulator would later help shape Orphibs Evolvapalooza (Chapter 6), while addressing explicitly the scale items in the creature PCG designs. For these reasons, we detail their research process in the following sections, highlighting the insights we obtained from them.

4.1 Orphibs II

Orphibs II was built with the goal of devising a proof-of-concept to assess the development feasibility behind a system capable of both procedurally generating sequences of behaviors as well as influencing the visual aspects of the agents accordingly.

We started our development cycle by creating a modular system, capable of outputting creatures' behaviors. Following up on this, Orphibs II was created with the purpose of providing a test-bed for this system. Adopting the design philosophy set for Petz (PF Magic, 1995), we decided to develop it as a life-simulation game, allowing players to interact with pet creatures because "People understand what pets are and essentially know how to interact with them. This is critical for producing believable characters" (Frank, Stern, and Resner, 1997).

Inspired, not only by Petz (PF Magic, 1995), but also Creatures (Millenium Interactive, 1996) and The Sims (Maxis, 2000), the player takes on the role of a caretaker for alien creatures called Orphibs. Similar to these games, he/she does not control an avatar but rather is represented by a hand capable of picking-up objects, including Orphibs and scrolling through the game world. A world composed of a 2D top-down grassy environment, populated by several objects including a square straw bed, an apple tree and three toys. Orphibs can use each of these objects except for the tree. Additionally, the world is subject to a day-night cycle. A screenshot of the prototype is illustrated in Figure 4.1.

FIGURE 4.1: Screenshot from Orphibs II. It illustrates two Orphibs several of the in-game objects.

Orphibs are anthropomorphic bipedal creatures. They have internal needs: hunger, energy, fun, social and reproduce. These regulate which actions the Orphibs can perform at a given time. They can play with toys, eat apples, sleep on the straw bed, talk to each other, wander around the game world and reproduce. When they perform the latter, they lay eggs resulting in an offspring that receives genetic material from each of its progenitors. Finally, Orphibs can mature over time: they grow into a maximum height, display body discoloration and eventually die.

4.1.1 Agent Architecture

The Orphibs II prototype was created using Unity3D (Unity Technologies, 2005). Its agent architecture is based on the one presented in the The Sims (Zubek, 2010). It is illustrated in Figure 4.2.

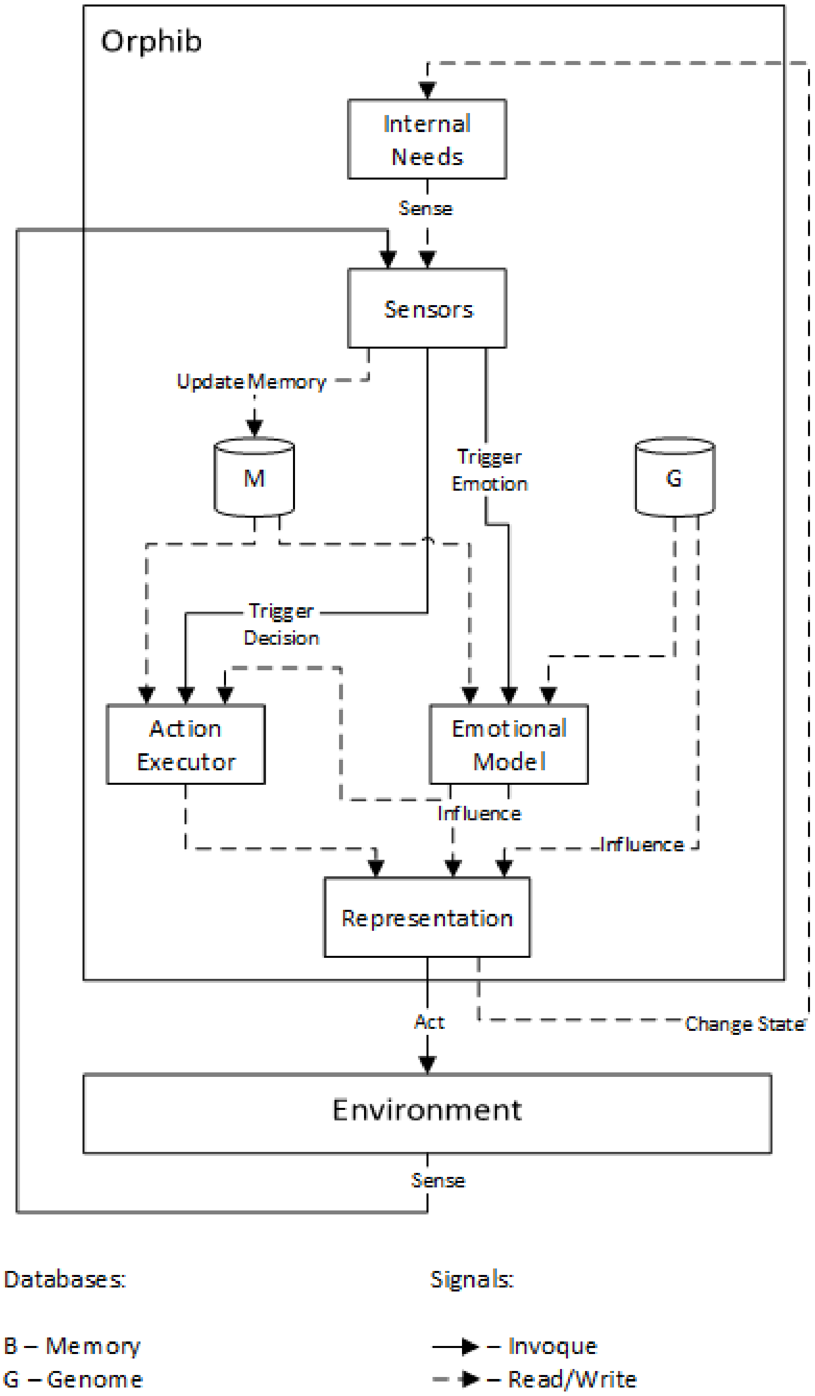


FIGURE 4.2: Orphibs Agent Architecture. Components are depicted as rectangles, while data bases as cylinders. Method Invocations and Read/Write operations are depicted as straight, and dashed, arrows respectively.

The architecture has the following components:

- **Action Executor** - Fed by the Sensor and Internal Needs modules, this is responsible for making Orphibs deliberate and execute actions. However, these can be canceled any time players pick Orphibs up.

- **Representation** – Module containing the Orphibs' visual representation: eye and body pigmentation, animations and aging. Orphibs' physical traits such as their eye and body color, are stored in the genome module.
- **Sensors** - This perceives the world's objects' external state. When an object stops being perceived, its last recorded state is moved to the Memory database. Orphibs only possess a visual sensor, i.e. virtual eyes, that varies according to day light intensity.
- **Memory** - This stores, as beliefs, the most recently perceived state of an object. These represent the object's attributes: position, existence and availability. Additionally, they are coupled with a certainty degree, ranging from 0 to 1. This decays over time when an object is no longer being perceived.
- **Emotional Model** – A module implementing the Computational Belief-Desire Theory of Emotion (CBDTE) (Reisenzein, 2012b). Emotions are generated from the Orphib's beliefs and desires.
- **Internal Needs** - A module that stores and updates the satisfaction levels of the Orphibs' needs at a constant rate during every update time frame. Each need decays at a constant rate. Reproduction and energy however, decay differently: the former only starts to decay when the Orphib reaches 10% of its total life span (to simulate puberty), with the rate of $Constant \times \frac{MaxAge}{1+(CurrentAge-0.1\times MaxAge)}$. This means that as the Orphib ages, its satisfaction decaying rate decreases. On the other hand, energy changes according to the expression $decayRate \times (1.7 - EyeSight)$, where $EyeSight$ decreases with the world's light intensity so, Orphibs spend more energy when they see less.
- **Genome** – Database storing the Orphib's complete genotype used during reproduction.

Reasoning System

The reasoning system behind the Orphibs is an implementation of the Goal-Based Behavior (Zubek, 2010). However, instead of selecting one of the top three actions that lead to the highest utility fulfillment, actions are chosen with a bias towards positive, intense, emotions.

During the execution of an action, Orphibs can perform a meta-action called 'Wait'. For instance, this happens when Orphibs want to play with toys that are being used by other Orphibs. In this case, the Orphib stops and waits for its turn. However, while waiting, the Orphib checks for an action that might provide a better emotional fulfillment than the current one. Moreover, the halted action's score decays 10% every time the Orphib performs an unsuccessful selection. This way, deterministic behaviors and racing conditions are reduced.

Finally, the Orphibs' reasoning system also implements a basic personality model. This is accomplished by each Orphib having individual emotional intensity functions. They map an emotion to a value that ranges from -1 to 1 and, as previously explained, they influence the reasoning system.

TABLE 4.1: Emotion Degrees.

Emotion Type	Emotion Degree
Happiness	+
Unhappiness	-
Fear	-
Surprise	+/-
Disappointment	-
Hope	+
Relief	+
No Emotion	0

Emotional Model

The inclusion of emotions in the Orphibs II architecture was made so that we could bridge the gap between behaviors and animations (Cerezo, Pina, and Serón, 1999; Multon et al., 1999), using expressions (detailed later in this Chapter), following the line-of-thought stating emotions play an integral part in inciting believability (Nareyek, 2007; Hudlicka, 2008; Ochs, Sabouret, and Corruble, 2010; Hudlicka and Broekens, 2009) (though we would later uncover this to be only part of a more complex problem).

The emotional component was based around the Computational Belief-Desire Theory of Emotion (Reisenzein, 2009; Reisenzein, 2012b) as it primes for simplicity. In this theory, a subset of emotions are perceived as outputs resulting from the comparisons between perceived beliefs and an internal belief-desire database. This was further expanded to include 'fantasy' emotions, through the imagination of events (Reisenzein, 2012a).

Internally, emotions are generated whenever beliefs and desires are fed into the comparators. They comprise an intensity, their absolute value; a type, which denominates what emotion was generated; and a sign, an integer value indicating if emotions are positive (1), neutral (0) or negative (-1). However, neutral degree is only used to indicate no emotions. Table 4.1 showcases the emotion types and their respective degrees.

Since more than one emotion can be generated from the comparators, whenever a comparison is made, we assume that the one originated was the most intense. Because Orphibs only express one emotion at a time, the originated one's intensity is compared with the their current emotion. If it is greater, Orphibs express the originated emotion. Additionally, expressed emotions' intensities decay, linearly, over time and when it reaches a value close to zero, the Orphibs stop feeling that particular emotion.

There are two moments when emotions can be generated: Before an action is selected (pre-effect) and during its execution (post-effect). Pre-effect emotions are equivalent to the fantasy emotions depicted in Reisenzein (2012a): since Orphibs do not know the actual outcome of their actions, they express emotions from what they assume it will be. Post-effect emotions, on the other hand, are the Orphibs' reaction to events, more specifically, their reaction to newly perceived beliefs.

Besides influencing the reasoning system, emotions are also expressed visually through facial expressions, to convey to the player what the Orphib is feeling. These

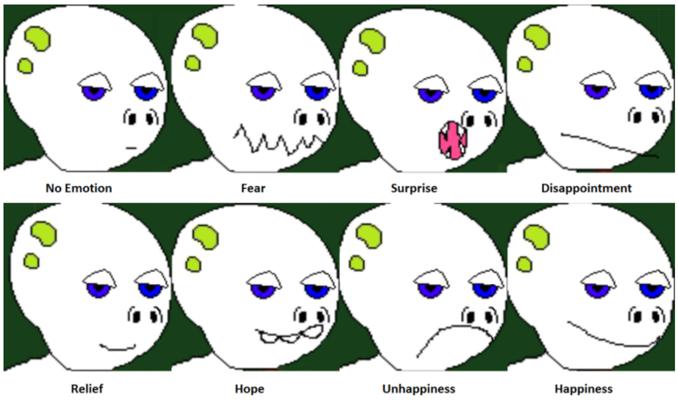


FIGURE 4.3: The Orphibs' facial expressions. From left to right and, top to bottom are depicted: no emotion, fear, surprise, disappointment, relief, hope, unhappiness and happiness.

are depicted in Figure 4.3.

This change also accommodated the aforementioned genetic personalities. Since the emotions' intensity was also mapped to functions similar to the utility ones, these were also included during the recombination, and mutation phases present in the creatures' reproduction.

Evolutionary Mechanism

Since Orphibs support reproduction, they have an internal evolutionary mechanism consisting in a simplified Genetic Algorithm (GA) (Eiben and Smith, 2007). Their fitness is represented by their ability to survive long enough to reproduce. In addition, parent selection and survivor selection are not performed in a system-centric manner but rather by the Orphibs' internal reasoning system and maturing mechanism respectively.

Algorithmically, the Orphibs' reproduction works in the following manner:

1. Approach potential mate.
2. Perform mating dance.
3. Create egg.
4. Inseminate egg.
5. Recombine the genetic material from both progenitors.
6. Fulfill the reproduction need of both Orphibs, by the same value.
7. Hatch the egg after 10 seconds have passed.

For recombination purposes, it is considered that the Orphib that engages in the reproduction action is the alpha progenitor, while the other one is the beta.

Genome

The Orphibs' genotype is composed of a data structure containing several data types:

- **Maximum Size** – Floating point representing the Orphibs' maximum size.

- **Maximum Age** - Floating point representing the Orphibs' maximum age.
- **Maximum Vision** - Floating point representing the farthest an Orphib can see under ideal conditions.
- **Minimum Vision** - Floating point representing the farthest an Orphib can see under non-ideal conditions.
- **Eye Color** – RGBA color representing the Orphibs' eye color.
- **Body Color** - RGBA color representing the Orphibs' body color.
- **Personality** – Vector of data point vectors. This represents the Orphibs' personality.

Additionally, these attributes are grouped so that the recombination operator is applied to each group independently. The groups are: one containing maximum size and maximum age, one composed of maximum, and minimum vision, one comprising the eye color, one containing the body color and one for the personality.

Variation Operators

The Orphibs' evolutionary mechanism uses both a crossover operator and a mutation operator. Recombination is applied to each group of the genome independently with a probability of 0.99. This means that each group has a 1% change of being cloned from the alpha progenitor's genome. The recombination algorithm used for most groups was Whole Arithmetic Recombination (Eiben and Smith, 2007) with an alpha factor of 0.5 while the personality group uses Uniform Crossover (Eiben and Smith, 2007) with a probability value of 0.5. Unlike crossover, mutation is applied to each attribute, with a probability of 0.3 (this value was chosen in order to provide some local variability). The operator used was Non-Uniform Mutation (Eiben and Smith, 2007) with a Gaussian Distribution (Eiben and Smith, 2007) with the previous value of the genome attribute as means and standard deviation of 1.

Genetic Personalities

To support procedural behaviors, the system includes genetic personalities. Since Orphibs can reproduce, they have an internal evolutionary mechanism consisting of a Genetic Algorithm (Eiben and Smith, 2007) albeit without an explicit fitness function. The progenitors' utility functions could then be recombined, and mutated, giving place to offspring.

These personalities derive from the Orphibs' emotion intensity functions which map the emotions' intensities to the values resulting from the model's computations (Reisenzein, 2012b). These functions are transmitted from parents to offspring using variation operators. Because altering them leads to different emotional states, they influence the Orphibs' resulting actions.

4.1.2 Discussion

As previously mentioned, the main goal behind the creation of Orphibs II was to devise a proof-of-concept to assess the development feasibility behind a system capable of both procedurally generating sequences of behaviors as well as influencing the visual aspects of the agents accordingly. While a simulation was run to validate the game's agents' autonomy, this prototype was still not ready for user evaluation.

The prototype was a low fidelity one, as it was built using 2D and Orphibs had an anthropomorphic design, both due to their cost simplicity and our development experience. Yet our research goal focused on 3D environments (the reason for this, has to do with the fact that videogames' visuals have been, over the years, shifting towards 3D (Roettl and Terlutter, 2018)) and non-anthropomorphic creatures. Building upon the knowledge gained from Orphibs II systems, we created an additional prototype in an attempt to overcome the aforementioned obstacles.

4.2 The Savanna Simulation

As mentioned, this prototype builds upon Orphibs II. The first main change was the addition of a third spatial dimension. This required developing new graphical assets. The second change involved switching from anthropomorphic creatures to non-anthropomorphic ones. This increased the works' complexity as lack of human features, namely a verbal language, made it difficult for humans to understand the creatures' motivations, emotions and other traits. To this extent, Mateas (2003) introduces expressions, a possible solution to close the gap between behaviors and animations.

Studies corroborate the importance of animal expressions for communication in real life animals (Menzel and Fischer, 2011). For instance, as early as the 19th Century, Darwin introduced his insights regarding expressions in humans and animals (Darwin, 1889). In his work, he describes several accounts of observed expressions in 'lesser animals' during different scenarios. This list includes: vocalization of sounds; hair, fur, or dorsal crest erection; body part dilation; and ear movements. Following his listing, Darwin adds that some expressions are only made possible by specific animals due to their morphological and anatomical nature.

From this, we can understand expressions, of a given agent, as the motions created in response to a specific stimuli, either from the environment or the agent itself. Motions can be, for instance, of muscular nature resulting in what we understand as facial expressions or vocalizations of sounds. Observing such motions conveys information regarding the agent's state. This can be further elaborated with the introduction if two key expression-characterizing elements: the set of actions constituting the motions and their respective dynamics. As expressions are perceived overtime, dynamics are an important factor in distinguishing between two similar sets of motions. For example, the speed at which an actor performs a given motion can be a cue for an observer to interpret the resulting expression as having been performed calmly or nervously.

Adding to the previous notion, we can append behaviors to derive an expanded concept, we defined Performance. This can be viewed as the audio-visual output of a given agent, or group of agents, over time. This translates not only to what the agents do, their behaviors, but also how they do it, their expression. From a videogame perspective, this is done through sequencing animations and sounds. In performances, observable meaning is then accomplished through the combination of both elements and it should not necessarily be bound to the agent's deliberation (e.g. tremors caused by stressful situations).

With this in mind, our new prototype involved a simplified recreation of an African Savanna, as depicted in Figure 4.4. We decide to use replications of realistic animals due to their documented expressions. This, in theory, would simplify our evaluation process later on, before moving to more fantastic creatures. Therefore, our simulation included hyenas, gazelles and lions (Estes, 2012).

The Orphibs II architecture was recreated for this prototype albeit with a small change: with the transition to 3D, we saw this as an opportunity to create a system capable of affecting the animations to a larger extent than it previously did. As such, the system's emotional component was integral in blending animations from a database of poses. This way, the animation system is able to add one, or more, expressions on top of 'regular' animations.

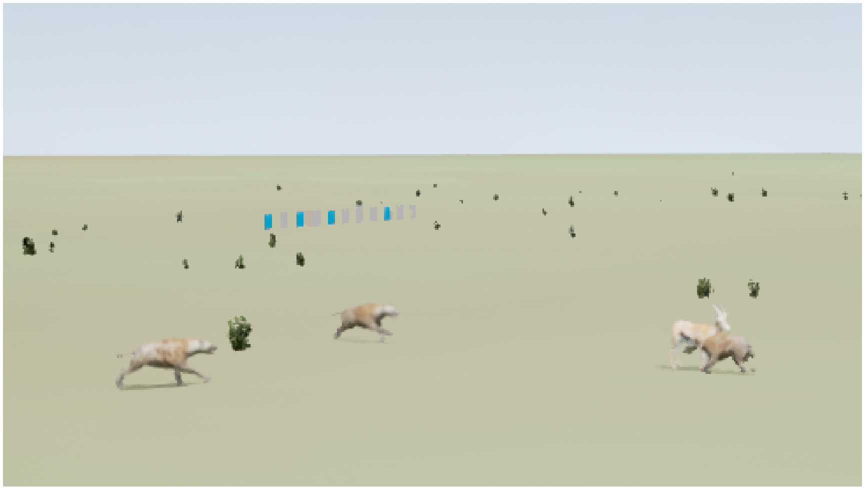


FIGURE 4.4: Screenshot from the Savanna Simulation. It depicts three hyenas chasing a gazelle.

With the prototype created, we ran several deterministic scenarios. The main goal was to verify if the agents were able to perform as intended given little to no stochastic elements. Having been successful, this set grounds to our actual evaluation activity, that of the creatures themselves.

4.2.1 Survey

As the literature highlighted 'expressions' as an important characteristic, we devised a survey to assess one of the simulations creature's expressions. It consisted in showing users a 34 second video rendering of the scenario simulation, followed by a survey (transcribed in Appendix A.1). This survey was divided into an open ended question were subjects wrote a short essay detailing their observations, followed a multiple-choice part where they had to select the expression closest to their interpretation.

The video in question was comprised by a hyena, showing a single different expression every 6 seconds. Particularly,

- From **00:00** to **00:06**: Exhibited nervousness by shaking its body.
- From **00:07** to **00:13**: No expression was exhibited.
- From **00:14** to **00:20**: Exhibited a submission expression by bowing its head.
- From **00:21** to **00:27**: Exhibited an aggressive expression by lowering its shoulders, poised to strike.

- From **00:28** to **00:34**: Exhibited a curiosity expression by tilting its head sideways.

The survey was published online and had a total of 32 participants, with 78% in the '18 - 30' age section whilst 28% were in the '31 - 59' section. Regarding videogame contact, 13% had 0 - 1 weekly hours, 66% had 2 - 10 weekly hours, 14% had 10 - 20 weekly hours, and finally, 6% had more than 20 weekly hours.

Each of the survey's parts were analyzed separately. The short essays were encoded using a two-step approach: first, we extracted key moments corresponding to the video's segments (for example, one participant answered 'First it looked happy, then sad, then angry, then curious and finally happy again'. This was divided into 5 moments: 'First it looked happy', 'then sad', 'then angry', 'then curious' and 'finally happy again'). Second, these were coded through semantic encoding to extract adjectives (in the previous example, we were able to retrieve 'happy', 'sad', 'angry', 'curious' and 'happy'). These adjectives are listed in Table 4.2 with their respective relative frequency. However, an additional category named 'Undefined' was added to group situations in which we were unable to extract adjectives. For the analysis, each moment's adjectives were compared against the intended expressions for their semantic similarity. In multiple-choice answers on the other hand, one of each of the questions' answers had the expression we were trying to convey. This is depicted in Table 4.3.

TABLE 4.3: Responses gathered from the multiple-choice part of the survey. This table indicates each of the selected answers' relative frequencies, means and standard deviation. Most frequent answers are highlighted in bold and underlined.

Question	A1 (%)	A2 (%)	A3 (%)	A4 (%)	Mean	Std. Dev.	Expected Answer
1	34.38	<u>**59.38**</u>	3.13	3.13	1.75	0.66	**A1**
2	<u>**75.00**</u>	9.38	3.13	12.50	1.53	1.03	**A1**
3	3.13	21.88	<u>**65.63**</u>	9.38	2.81	0.63	**A3**
4	<u>**84.38**</u>	6.25	9.38	0.00	1.25	0.61	**A1**
5	12.50	9.38	<u>**78.13**</u>	0.00	2.66	0.69	**A3**

Results showed that, on the first part of the survey, less than 50% of the participants failed to correctly list the intended expressions in moments 1 (only 12.50% identified 'Nervous') and 3 (31.25% identified 'Submissive') whilst in the remaining ones, more than 50% of the subjects were able to correctly identify the intended expressions, or semantically similar ones: in moment 2, we extracted the adjectives 'Neutral' (43.75%) and 'Relaxed' (9.38%). In moment 4 participants answered 'Agressive' (81.25%) and 'Angry' (3.13%). Finally, we obtained the adjectives 'Curious' (59.38%) and 'Inquisitive' (3.13%).

Conversely, when prompted to answer the multiple choice portion of the survey, more than 50% of the participants correctly chose the intended expression.

These results indicate that, though the evaluation determined whether the expressions were correctly communicated to the participants, it had shortcomings:

1. The information we could gather from this survey informed more about the expressions themselves than the creatures. Even if participants were able to

TABLE 4.2: Responses gathered from the essay part of the survey. This table indicates the frequency of each keyword, on each part of the clip. Values of 0% were replaced by '-' for readibility.

Expressions	**Video Segments**				
	(00:00 - 00:06)	(00:07 - 00:13)	(00:14 - 00:20)	(00:21 - 00:27)	(00:28 - 00:34)
Relaxed	9.38%	9.38%	-	-	3.13%
Attentive	9.38%	9.38%	12.50%	-	-
Neutral	37.50%	43.75%	6.25%	-	12.50%
Happy	6.25%	6.25%	-	-	-
Hunger	3.13%	3.13%	-	-	-
Nervous	12.50%	-	-	-	-
Enraged	3.13%	3.13%	-	-	-
Menacing	3.13%	3.13%	-	-	-
Tired	6.25%	9.38%	3.13%	-	-
Calm	3.13%	3.13%	-	-	-
Sleepy	-	-	3.13%	-	-
Submissive	-	-	31.25%	-	-
Sad	-	-	12.50%	-	-
Distracted	-	-	3.13%	-	-
Resigned	-	-	3.13%	-	-
Observant	-	-	6.25%	-	-
Exhausted	-	-	3.13%	-	-
Secure	-	-	3.13%	-	-
Aggressive	-	-	-	81.25%	-
Angry	-	-	-	3.13%	-
Excited	-	-	-	3.13%	-
Anticipative	-	-	-	3.13%	-
Playful	-	-	-	3.13%	-
Defensive	-	-	-	-	-
Curious	-	-	-	-	59.38%
Friendly	-	-	-	-	3.13%
Cute	-	-	-	-	3.13%
Inquisitive	-	-	-	-	3.13%
Confused	-	-	-	-	3.13%
Docile	-	-	-	-	3.13%
Downcast	-	-	3.13%	-	-
Undefined	6.25%	9.38%	9.38%	6.25%	6.25%

correctly identify the expressions, there was not enough information to assess, for instance, if they were consistent with the creature.

2. Since the creature was evaluated completely deprived of context (e.g, not part of ongoing narrative, activity or game), we conjecture this lack of context did not clarify some expressions that were potentially ambiguous (e.g., bowing the creature's lead to participants identify expressions such as sleepy and sad).

In the end, this only lead us to reflect and propose a different evaluation method.

4.3 Discussion and Summary of the first DSR Iteration

The prototypes developed during this phase provided important insights in building the knowledge required to propose the Creatrure Believability Scale (detailed in Chapter 5). At first, the extension of the Orphibs II agents' architecture allowed the introduction of *expressions*, a concept integral in connecting animations to behaviors. This unraveled believability as a potential way to evaluate creatures, at a macro-level.

Despite that, the evaluation method we employed proved fruitless in informing our initial intent. In the essay part of the evaluation, we noticed some expressions were mistaken by visually similar ones (i.e. a submissive expression was mistaken by a sad one). Since the evaluated creature was displayed in a blank space, the expressions themselves were not sufficient to convey a message. So, we theorized a key element was missing during the evaluation: a frame setting a stage, an environment, and a time-line filled with perceivable events from which an observing player can derive meaning from them or, in summary, a context.

Other factors were absent from both our method and the works of Nareyek (2007), Hudlicka (2008), Ochs, Sabouret, and Corruble (2010), and Hudlicka and Broekens (2009). The link between expressions and believability, or other keywords cited by the authors such as 'realism' proved to be vague. First, our evaluation does not hint at any of these. We could not derive any sense of believability by having subjects confirming whether agents are expressive or not.

Second, our hyena had its features implemented from an ethology guide (Estes, 2012) yet, had we used expressions from any other animal, or even invented our own, would it be believable? Had we used a fantastic animal, instead of a hyena, would it be believable? We could not find any answers for these questions, from the results alone.

Third, the idea that only a few factors are capable of addressing an ill defined construct, such as believability, seem counter-intuitive. But, upon a closer inspection on the works cited, it becomes apparent the implicit anthropomorphism, coupled with pre-defined expectations, fills in for the lack of additional factors. These problems set the grounds for revising the current *believability* construct into one capable of serving as a basis for an assessment tool.

Chapter 5

The Creature Believability Scale

The Creature Believability Scale (CBS) is a tool to assess believability on non-human creatures. In the literature, autonomous agents are sometimes catalogued as Non-Player Characters (Nareyek, 2007; Hudlicka, 2008; Ochs, Sabouret, and Corruble, 2010; Hudlicka and Broekens, 2009), further cementing their anthropomorphic stance and, while creatures may not necessarily be considered characters *per se*, they, as autonomous agents, share some common grounds. Namely, both of are meant to support game worlds and, in some cases, their narratives. So, as believability is considered a beneficial characteristic in Non-Player Characters, we assume believability to be beneficial, also, for the aesthetic appreciation of creatures, thus contributing to the acceptance of the fictional contexts where they play a meaningful role. Furthermore, such characteristic goes beyond the simple functionality assessments used in PCG approaches, also present in the literature (the works of Ribeiro et al. (2003), Coros, Beaudoin, and Panne (2009), and Grand, Cliff, and Malhotra (1997) to name a few). It is, foremost, contextualized and requires an assessment at a macro-level.

Taking into account the construct nature of believability, our proposal comes in the form of a rating scale (Spector, 1992), following a similar structure to other evaluation methods present in the literature (Kidd, 2003; Kidd and Breazeal, 2004; Lombard et al., 2000). The scale's most recent revision (its fourth one) is transcribed in Appendix D.

5.1 Scale Construction

The initial construction of the scale proposal underwent a three-step process, inspired by the methods described by Spector (1992) and Hinkin, Tracey, and Enz (1997). Firstly, we defined what we would consider creatures of interest for this study, establishing a division between humans (and humanoids) and the entities under our study. Based on the definition proposed by Whithlatch (Whitlatch, 2015b) (which went in the desired direction) we defined *creatures* as previously stated:

Zoomorphic entities, inspired by (contemporary or otherwise) living or fictional beings, that are not identifiable as human, nor fundamentally human-like, despite whatever anthropomorphic characteristics they might, or might not, have.

We then inductively framed *creature believability* as the maintained perceptual illusion that a creature is a natural part of its (virtual) environment, perpetuating a state of Presence, even if the environment, the creature, or both are not factual or even realistic. This happens when it provides audio-visual feedback the audience expects, or comes to expect. So, we postulate that, since Believability is closely tied to perception and the suspension of disbelief, the measure of Believability is the inverse of the measure of how conscious the audience is of perceived errors in the representation of the creature, independently of the subjective nature of those errors.

This was followed by defining a list of underlying constraints of our scale:

1. Unlike humans, and humanoids, who have a distinct (limited) set of characteristics, our definition includes markedly heterogeneous beings, ranging from insectoid to mammal-like. With this in mind, our first concern was to create a sufficiently broad set of identifiable elements to evaluate the wide variety of creatures.

2. Instead of considering believability as a binary factor, we opted to work with a fuzzy set. This would allow us to better quantify the qualitative, and variable, nature of the believability a given creature may have, as well as better identify which elements contribute to that perception.

3. As believability is intrinsically perceptual (Fogg and Tseng, 1999), we considered limiting the evaluation of the creatures to perceivable (phenotypic) elements.

4. Whilst the creature definition given includes entities inspired by both living and fictional beings, the everyday experience of a human is with living beings. Therefore, the starting point for the evaluation elements was plausible, naturally occurring, characteristics.

5. Finally, since our objective is to validate and, subsequently, use this scale with players, which may or may not have a scientific background, the language was made accessible, deprived of technical terminology.

After analyzing the constraints, we constructed several candidate statements. This was done by deconstructing the vague definition, proposed for believable agents, ‘alive and thinking’ (Rosenkind, Winstanley, and Blake, 2012). To this extent, we surveyed characteristics present in living beings (Kadhil, 2005) and how cognitive processes, colloquially known as ‘thinking’, could be achieved. Both were linked together, taking into account how they could be perceived, through induction on examples and with the support of extant literature. For the latter we surveyed multiple study fields such as, but not limited to, believable agents, ethology, biology, human perception of living-beings, illustration and animation. The 46 statement-long list (revised once for form simplification) is depicted as follows:

1. The creatures move by themselves
2. The creatures’ motions are fluid
3. The creatures’ motions reflect their weight/size
4. The creatures’ expressions anticipate their actions
5. The creatures make several simultaneous motions
6. The creatures’ actions involve more than one step
7. Each of the creatures’ body parts have inertia
8. The creatures interact with the environment
9. The creatures react to stimuli
10. The creatures controls their body

11. The creatures recognize themselves
12. The creatures' behaviors differ from other creatures of the same species
13. The creatures focus on stimuli
14. The creatures direct their behaviors towards targets
15. The creatures locate objects in the environment
16. The creatures have diverse priorities
17. The creatures alternate between tasks
18. The creatures show positive (or negative) emotions towards objects, or events
19. The creatures show expressions to known stimuli
20. The creatures express moods through their body
21. The creatures feel empathy towards other creatures
22. The creatures' same-stimuli reactions change over time
23. The creatures learn from past events
24. The creatures learn through imitation
25. The creatures are able to apply old behaviors to new, similar, situations
26. The creatures use objects from the environment
27. The creatures make tools
28. The creatures work with other creatures for a common goal
29. The creatures coordinate with other creatures
30. The creatures communicate with other, same-species, creatures
31. The creatures communicate with other, different-species, creatures
32. The creatures absorb substances/energy from the environment to survive
33. The creatures expel material
34. The creatures have different life-stages
35. The creatures change the way they look with age
36. The creatures change the way they sound with age
37. The creatures change the way they behave with age
38. The creatures engage in reproductive acts
39. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.
40. The creatures have traits particular to their sex
41. The creatures' postures and expressions are coherent with their behavior
42. The creatures' actions are appropriate to their context
43. The creatures' body are adapted to their habitat

44. The creatures' bodies and behaviors are consistent
45. The creatures play by themselves
46. The creatures play with others

Some works in believable agents reference the existence of some degree of consciousness. To this end, we used ConsScale (Raúl, Ledezma, and Sanchis, 2010) to derive some of our scale's items. Firstly, we used only a subset of the scale, each of the scale's levels' biological phylogeny working as a heuristic to observe constraints 1 and 4. Thus, only level 2 (viruses) through 8 (primates) were used. Subsequently, we rewrote the levels' items to enumerate the phenotypic traits corresponding to the respective statements. The resulting items accounted for reaction to the environment, intention (Menzel and Fischer, 2011) and display of emotions and sociability (both supported by ethological studies (Lorenz, 2010; Adkins-Regan, 2005; Darwin, 1889; Dawkins, 2000)). It is important to note, however, that regarding intentionality, it may be argued that some creatures, such as insects, work similarly to automata, reacting only to stimuli as they appear. Though, attributing motivations/intentions to living-beings (and even inanimate objects) is a behavior common to humans (Dennett, 2009; Braitenberg, 1984), as an observing subject can treat a creature "as if it were a rational agent who governed its 'choice' of 'action' by a 'consideration' of its 'beliefs' and 'desires'" (Dennett, 2009). Several items were gathered from these works, they include items 8 through 11, 13 through 18 and 21 through 31.

We also added an item for personality, as it is a trait that has been identified in animals (Gosling, 2008). Defined as "those characteristics of individuals that describe and account for temporally stable patterns of affect, cognition, and behavior" (Gosling, 2008), we can measure if a creature has personality by comparing it with others of the same species in similar situations. For example, frightening two dogs may cause one to flee and the other to retaliate. Item 12 takes this information into account.

From a different perspective, Thomas and Johnston argue that cognitive processes can be illustrated through expressions (Thomas and Johnston, 1997). They explain, that workers at Disney, throughout its early years, extensively studied animals, concluding they "communicate their feelings with their whole body attitude and movement" (Thomas and Johnston, 1997) (a statement also observed as early as the 19th century by Darwin (Darwin, 1889)). This reflection was the origin of the 12 Principles of Animation, 9 of which were incorporated into our scale: every action is preceded by an expression (Lorenz, 2010; Reynaud, Donnart, and Corruble, 2014), living-beings usually perform several movements simultaneously and these have both an acceleration, and a deceleration, and are performed in circular motions. Finally, body parts and other appendages follow the law of inertia. The remainder principles, the ones detailing how to make appealing characters, 2D animations and narratives, were beyond the scope of our assessment. To synthesize these aspects we considered items 2 through 7, 19 and 20.

In the case of biology, living beings are capable of autonomous movement and viewed as open systems (Kadhil, 2005): they retrieve matter and energy from the environment and, in return, produce waste material. Both can be perceived directly through actions. Though, the latter may be omitted due to age restrictions. Nonetheless, observing waste objects in the game world may have the same effect as observing the act (McCloud, 1994). From these notions we extrapolated items 1, 32 and 33 from this.

Living entities are also expected to, at the very least, grow in size and number of cells (Kadhil, 2005). If they have encompass a life-cycle, observing some form of growth can be as simple as witnessing the being's various stages (maturing, etc.) or observing several different creatures during their various life stages (McCloud, 1994). In nature, there are visual cues that help identify animals during their different life stages, including difference in size, limb and facial feature proportions, etc. (Lorenz, 2010). Some of these are even tied to the creatures' biological sex (Lorenz, 2010). Items 34 through 37, 40 were derived from these concepts.

Living entities can also reproduce sexually or asexually to generate offspring (Kadhil, 2005). This can be measured directly by observing the act of reproduction. However, such acts may impact the age rating of video games. Nonetheless, indirect measurements can be accounted for: if we are able to observe growth, it is then implied that infant creatures were created through reproduction. Items 38 and 39 were written from this.

Still regarding biology, an aspect that the ethology sub-field refers is the creatures' (mostly mammals) ability to play (Lorenz, 2010). This can happen throughout their life-time and can impact how some behaviors, such as hunting, are learned. From this knowledge, we wrote items 45 and 46.

Within the scope of illustration, Whitlatch identifies the importance of internal and external coherence (Whitlatch, 2015b; Whitlatch, 2015a). The former revolves around the coherence between a creature's behavior and their body's design ("why a creature looks the way it does" (Whitlatch, 2015b)). This means the creatures' body parts are integral to hint some of the behaviors a creature may have. If damaged, or absent, it is expected that they do not behave as properly, or at all (i.e. a creature without eyes, must not see, etc.). External coherence, on the other hand, measures the relationship between the creature's body and its habitat ("the anatomical structure supports and makes possible the lifestyle" (Whitlatch, 2015b)). This requires an assessment, not only of the creature, and its design, but also of the environment it is in. This was translated into items 41 through 44.

On a final note, as a starting point, the items were accompanied by a 5-point agreement Likert Scale (Brown, 2010) ('Strongly Disagree', 'Disagree', 'Undecided', 'Agree' and 'Strongly Agree'). So, with a first proposal made, we deployed the scale to validate how it fares when measuring 'believability'.

5.1.1 Administering the Creature Believability Scale v0.1

To administer the initial iteration of the scale, we designed a questionnaire (transcribed in Appendix A.2) where participants, after filling out demographic data, are shown 28 (McLeod, 2009) clips, 20 to 30 seconds long, obtained from various videogame sources. These clips all included at least one creature engaged in a specific activity. We used short clips because we meant to incite immediate responses rather than evaluating the participant's recollection of the clips (which could be contaminated by memory inaccuracy). It is worth noting that, we used videos in lieu of allowing users to actually play the videogames because Togelius, Yannakakis, and Karakovskiy (2011) explains the latter may cause subjects to sidetrack away from their goal.

After viewing each clip, subjects were prompted to:

1. Describe how many creatures were present in the video. This was included as a fail-safe measure (Mason and Suri, 2012; Kittur, Chi, and Suh, 2008) to allow deployment to Mechanical Turk.

2. Score 6 statements, taken from our believability items, through a 5-point Likert Scale. These items were displayed in a randomized order. Moreover, to reduce confirmation bias, 3 of the 6 items refer to elements absent in the respective clip. There were two reasons behind showing only a subset of the scale. The first involves user fatigue as administering the complete questionnaire would account for 1288 items (46×28). While one could argue that increasing the video length would potentially reduce the number of clips and subsequently the items' total, it would force the users to recollect more information, something we previously stated we wanted to avoid. The other reason was due to the fact that, to our knowledge, not every videogame has creatures with the characteristics present in the scale's items. However, grouping creatures from different videogames together would, on one hand, cause the issue we previously discussed and, on the other, potentially introduce a bias due to the change of context between videogames.

3. Rate the clip's creatures' believability using a 10-point Differential Semantic Scale with a Non-Believable-Believable pair. This was meant to assess the presence of correlations between Likert items and the creature's believability.

4. Similarly rate the clip's setting's believability. This was to reveal whether or not (and to what extent) the setting's and the creature's perceived believability are correlated.

Choosing Content for Evaluation

The videogames, included in this experiment, were chosen, by taking into account the following factors: first, we selected ones where creatures had an extensive on-screen presence, as we assume that, in these videogames, creatures have an additional development effort that would not be justified otherwise. As such, most of the videogames we considered are ones with open-world elements and life-simulation videogames. Finally, we chose to consider videogames made in the last 15 years. The main reason behind this lies in our belief that such videogames incorporate recent technology. This way, we mean to reduce any bias which could arise from notorious technological limitations.

This selection process resulted in several creatures from 19 videogames:

- Hyenas and cheetahs from Afrika (Rhino Studios, 2008)
- D-Horse and D-Dog from Metal Gear Solid V: The Phantom Pain (Kojima Productions, 2015)
- The EyePet from EyePet (SCE London Studio and Playlogic Game Factory, 2009)
- The dog from Fable 2 (Lionhead Studios, 2008)
- Dogmeat from Fallout 4 (Bethesda Game Studios, 2015)
- A rhinoceros from Far Cry 4 (Ubisoft Montreal, 2014)

- An Adamantoise from Final Fantasy XIII (Square Enix 1st Production Department, 2010)
- Chop from Grand Theft Auto V (Rockstar North, 2013)
- A black panther and a Bengal tiger from Kinectimals (Frontier Developments, 2010)
- A Rathian and a Rathalos from Monster Hunter Freedom Unite (Capcom, 2008)
- The Artic Fox from Never Alone (Upper One Games, 2014)
- A dog from Nintendogs (Nintendo EAD, 2005)
- Red Pikmins from Pikmin (Nintendo EAD, 2001)
- Dogs and a cat from The Sims 2 (Maxis, 2005)
- A sabertooth tiger from The Elder Scrolls V: Skyrim (Bethesda Game Studios, 2011)
- Chaos from Sonic Adventure 2: Battle (Sonic Team USA, 2002)
- Chaos from Sonic Adventure DX: Director's Cut (Sonic Team USA, 2003)
- Creatures from Spore (Maxis, 2008)
- Trico from The Last Guardian (genDESIGN and SIE Japan Studio, 2016)

Before deploying to a larger population, the survey then underwent a pilot testing process with 5 test subjects, correcting typos and other errors.

5.1.2 Survey Deployment Results

Our survey was deployed to Mechanical Turk, where 43 users participated (32 Males and 11 Females) with an average age of 31 $\pm$ 6. Regarding education, 19% of the participants had an Highschool degree, 70% had a Bachelor's degree whilst 12% had a Master's degree. Finally, 35% of these users had a weekly exposure to media (videogames, movies, tv) of up to 20 hours, while others had 20 to 40 weekly hours (42%), 40 to 60 weekly hours (16%), or 60 to 80 weekly hours (7%). Though Hinkin, Tracey, and Enz (1997) and Spector (1992) propose a minimum sample size between 100 and 200, for this initial item analysis step, we did not have the resources to recruit that many participants.

The results' analysis was performed under a two-step method. Firstly, we analyzed the items on a per-clip basis. This allowed us to study how the believability scores could correlate with the clip's items and remove the ones who did not. This is explained in subsection 5.1.2. Secondly, we grouped the items together and performed, on the questionnaire as a whole, a reliability analysis followed by an Exploratory Factor Analysis (EFA), as depicted in subsection 5.1.2.

Reliability Analysis

Before analyzing the items on a per-clip basis, we first studied the reliability of the believability semantic differential scales, which will be henceforth referred to as believability ratings. Specifically, we grouped them together and then calculated their Cronbach Alpha coefficient. As expected, results show a value of 0.96. While this is

TABLE 5.1: The extracted components using a PCA, with a Varimax Rotation and a stop criterion of Eigenvalue >= 1. A dash line is set to separate the group of values above a 50% Cumulative Variance

Component	Eigenvalue	% of Variance	Cumulative %
1	9.067	23.859	23.859
2	3.551	9.346	33.205
3	3.118	8.205	41.410
4	2.765	7.276	48.686
5	1.923	5.061	53.747
6	1.875	4.933	58.680
7	1.693	4.455	63.135
8	1.471	3.870	67.005
8	1.393	3.667	70.672
10	1.306	3.437	74.109
11	1.119	2.945	77.054

considered redundant (Steiner, 2003), it is as predicted since the group consisted in the same question across all clips.

Having an indication that the believability ratings were internally consistent, we ran a Principal Components Analysis (PCA) on each group of items (the 3 non-control items plus its corresponding believability rating). By fixing one factor, we considered to be Believability, we used the believability ratings as a control value: items were kept for the next analysis if both they and the believability rating loaded on the factor, as they were assumed to measure the same construct. Furthermore, we used cut-threshold loading value of 0.4. The resulting items are indicated in the column 'Analysis 1' of Table 5.2.

As depicted in the table, most items were able to load alongside the believability ratings. In fact, only 6 items were left out because their loadings were inferior to 0.4. These were item 16 (The creatures have diverse priorities), item 21 (The creatures feel empathy towards other creatures), item 26 (The creatures work with other creatures for a common goal), item 32 (The creatures absorb substances/energy from the environment to survive), item 40 (The creatures have traits particular to their sex) and item 46 (The creatures play with others).

Exploratory Factor Analysis

Having now a filtered list of items, we grouped the remainder 40 items together and analyzed them as a whole questionnaire. First, we analyzed the group's internal consistency by calculating its Cronbach Alpha coefficient. With a value of 0.9, between the accepted values (Steiner, 2003), we did not, at this point, remove any further item.

We then performed an EFA. The technique used was PCA with a Varimax Rotation, utilizing *Eigenvalues* $>= 1$ as a stopping criteria. This resulted in 11 components, illustrated in Table 5.1. However, the number as being too many for practical application of the scale. In order to find a more satisfactory solution, one with less factors, we established an additional criteria: a total variance explained minimum value of around 50%. As expressed in Table 5.1, this accounts for retaining either 4 or 5 components.

TABLE 5.2: The CBS analysis. The first column depicts the original scale's items. The second column displays which items were kept during the per-clip PCA with one fixed factor and a loading cut-threshold value of 0.4. The third column shows the loadings obtained after the second analysis performed using a PCA, with 4 fixed factors, Varimax Rotation and a loading cut-threshold value of 0.4. Values are omitted when their loading fails to meet the threshold. The items kept on the final iteration of the scale are depicted in bold.

Item	**Analysis 1**	**Analysis 2**			
	(Passed)	1	2	3	4
1. The creatures move by themselves	✓	0.442	**0.513**		
2. The creatures' motions are fluid	✓			**0.516**	
3. The creatures' motions reflect their weight/size	✓		**0.464**		
4. The creatures' expressions anticipate their actions	✓				**0.539**
5. The creatures make several simultaneous motions	✓		**0.714**		
6. The creatures' actions involve more than one step	✓	0.622			
7. Each of the creatures' body parts have inertia	✓				
8. The creatures interact with the environment	✓	**0.617**			
9. The creatures react to stimuli	✓		**0.735**		
10. The creatures controls their body	✓	**0.587**			
11. The creatures recognize themselves	✓				
12. The creatures' behaviors differ from other creatures of the same species	×				
13. The creatures focus on stimuli	✓		**0.620**		
14. The creatures direct their behaviors towards targets	✓	**0.763**			
15. The creatures locate objects in the environment	✓	**0.514**	0.454		
16. The creatures have diverse priorities	×				
17. The creatures alternate between tasks	✓			0.580	
18. The creatures show positive (or negative) emotions towards objects, or events	✓				**0.593**
19. The creatures show expressions to known stimuli	✓	0.416			**0.691**
20. The creatures express moods through their body	✓				
21. The creatures feel empathy towards other creatures	×				

22. The creatures' same-stimuli reactions change over time	✓			**0.668**	
23. The creatures learn from past events	✓			**0.544**	
24. The creatures learn through imitation	✓				**0.620**
25. The creatures are able to apply old behaviors to new, similar, situations	✓			**0.444**	0.564
26. The creatures use objects from the environment	×				
27. The creatures make tools	✓				
28. The creatures work with other creatures for a common goal	×				
29. The creatures coordinate with other creatures	✓		**0.744**		
30. The creatures communicate with other, same-species, creatures	✓		**0.723**		
31. The creatures communicate with other, different-species, creatures	✓			**0.519**	
32. The creatures absorb substances/energy from the environment to survive	×				
33. The creatures expel material	✓	**0.632**	0.449		
34. The creatures have different life-stages	✓				
35. The creatures change the way they look with age	✓			**0.682**	-0.425
36. The creatures change the way they sound with age	✓			**0.583**	
37. The creatures change the way they behave with age	✓			**0.724**	
38. The creatures engage in reproductive acts	✓		**0.539**		
39. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	✓		**0.495**	0.433	
40. The creatures have traits particular to their sex	×				
41. The creatures' postures and expressions are coherent with their behavior	✓	0.621			
42. The creatures' actions are appropriate to their context	✓	**0.654**			
43. The creatures' body are adapted to their habitat	✓				**0.694**
44. The creatures' bodies and behaviors are consistent	✓	0.702			
45. The creatures play by themselves	✓				
46. The creatures play with others	×				

Our next step involved choosing between 4 or 5 factors. To this extent, we ran two additional PCA, with the same rotation method as before, one by fixing 4 factors and, the other, 5. However, this time, we used a cut threshold loading factor value of 0.4. From observing the resulting loaded items, we concluded that using 4 factors, over 5, grouped items with similar underlying semantics and thus, would facilitate the process of naming/categorizing those factors. The factor loadings, resulting from a PCA with 4 fixed factors, are described in Table 5.2.

Next, three of the researchers analyzed independently the factor loadings in order to find categories to represent each factor grouping. This process underwent as follows:

1. Each of the three researchers individually studied the obtained factors and corresponding loadings and came up with naming proposals which would explain most, if not all, of the loaded items. This included deciding in which factors cross-loading variables would be kept.
2. We then gathered to discuss our proposals. During this step, we considered discarding items which deviated from our proposed semantics.
3. The process ended when we reached a consensus.

This process originated the "Relation with the Environment", "Biological/Social Plausibility and Sociability", "Adaptation" and "Expression" concepts for explaining factors 1 through 4 respectively: besides removing items with factor loadings below 0.4, an additional 5 items were removed. This included item 6 (The creatures' actions involve more than one step) and 17 (The creatures alternate between tasks) because their underlying concept did not align with the other factor-adjacent items; items 19 (The creature show expressions to known stimuli) and 41 (the creatures' postures and expressions are coherent with their behavior) who appeared to be better suited for loading with the Expressions factor; and, finally, item 44 (the creatures' bodies and behaviors are consistent) which we considered to belong to the Biological/Social Plausibility factor.

5.2 First Revison

The revised scale, and encompassing items, are then as follows:

Relation with the Environment - This category corresponds to the items related to environment interactions, ranging from reactions to environmental cues or directed behaviors to systemic exchanges. The originated items are then as follows:

1. The creatures interact with the environment
2. The creatures controls their body
3. The creatures direct their behaviors towards targets
4. The creatures locate objects in the environment
5. The creatures expel material
6. The creatures' actions are appropriate to their context

Biological/Social Plausibility - Corresponds to the creature's plausibility as a biological organism. This is demonstrated by showing autonomy and reactivity to its surroundings. Additionally, this category also encompasses the creature's ability to interact with other creatures. The items are as follows:

7. The creatures move by themselves
8. The creatures' motions reflect their weight/size
9. The creatures make several simultaneous motions
10. The creatures react to stimuli
11. The creatures focus on stimuli
12. The creatures coordinate with other creatures
13. The creatures communicate with other, same-species, creatures
14. The creatures engage in reproductive acts
15. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.

Adaptation - This category involves learning behaviors and growing (which we considered to be an adaptation at the biological level). The originated items are as follows:

16. The creatures' same-stimuli reactions change over time
17. The creatures learn from past events
18. The creatures are able to apply old behaviors to new, similar, situations
19. The creatures change the way they look with age
20. The creatures change the way they sound with age
21. The creatures change the way they behave with age

Expression - Expression encompasses the elements wherein creatures use their body as a means to communicate, learn or survive. The items are as follows:

22. The creatures' expressions anticipate their actions
23. The creatures show positive (or negative) emotions towards objects, or events
24. The creatures show expressions to known stimuli
25. The creatures learn through imitation
26. The creatures' body are adapted to their habitat

As a final step, we performed an additional reliability test on the remainder items as a confirmation. By calculating the Cronbach Alpha coefficient, it yielded 0.88 which is inside the acceptable range (Steiner, 2003).

The previous EFA suggested how the scale could be divided among several dimensions, the resulting theoretical model still needs confirmation. To this extent, the revised scale underwent an Confirmatory Factor Analysis (CFA), using a setup similar to the previous one.

5.2.1 Administering the Creature Believability Scale v0.2

We designed a second questionnaire (transcribed in Appendix A.3) where participants were shown ten 40 to 60 seconds-long clips, from various videogame sources, each with at least one creature engaged in a specific activity. This duration allowed them to engage in more activities or at least, in more time consuming ones (i.e. learning). Finally, the clips' order was randomly sorted to avoid bias. After viewing each clip, subjects were prompted to answer a fail-safe question (Mason and Suri, 2012; Kittur, Chi, and Suh, 2008), and to score each of the scale's items.

Choosing Content for Evaluation

We followed similar criteria to the ones in the previous experiment: we considered recent videogames, to reduce any bias arising from notorious technological limitations, and ones where creatures had an extensive on-screen presence, to facilitate the perception of their activities.

This process resulted in the following creatures:

- Radstags from Fallout 4 (Bethesda Game Studios, 2015)
- Wolves and tigers from Life of Black Tiger (1Games, 2017)
- Chop from Grand Theft Auto V (Rockstar North, 2013)
- D-Horse from Metal Gear Solid V: The Phantom Pain (Kojima Productions, 2015)
- Aliens from No Man's Sky (Hello Games, 2016)
- A sea vulture from Risen (Piranha Bytes, 2009)
- Wolves from The Elder Scrolls V: Skyrim (Bethesda Game Studios, 2011)
- a deer from theHunter (Games, 2005)
- Trico from The Last Guardian (genDESIGN and SIE Japan Studio, 2016)
- Antilopes from The Witcher III: The Wild Hunt (CD Projekt RED, 2015)

5.2.2 Survey II Deployment Results

Our survey was deployed online, where 19 users participated (12 Males and 7 Females) with an average age of 31.5 ± 11.9. Regarding education, 11% of the participants had a Highschool degree, 37% had a Bachelor's degree, 47% had a Master's degree and 5% had a Doctorate degree. These degrees were in the fields of Science, Technology, Engineering and Math (68%), Humanities (Literature, Social Sciences, etc.) (26%) and Arts (Illustration, Music, etc.) (5%). Finally, 32% of these users had a weekly exposition to media (videogames, movies, tv) of up to 20 hours, while others had 20 to 40 weekly hours (42%), 40 to 60 weekly hours (16%), or > 60 weekly hours (11%).

Confirmatory Factor Analysis

We considered each <subject, videogame> tuple as a separate answer because each clip contained its own context, creature(s) and activities. We were then able to use 190 entries (19 subjects $\times$ 10 clips) in our analysis. While the overall number of entries is above the 100, proposed by Hinkin, Tracey, and Enz (1997), the number of different subjects, 19, is well bellow it. This was due to the same reason as before: lack of resources to recruit additional participants.

The data underwent a CFA using a Maximum-Likelihood Path Analysis in order to test the scale's theoretical model's goodness-of-fit. To compare model data, we used the Chi-Square, Chi-Square/Degrees of Freedom ($\frac{\chi^2}{df}$), Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA) indexes.

A first structured equation models was considered. This model (4F), transcribed the original scale. In order to do this, the scale's four dimensions were converted to latent variables, each with a covariance link to the remainder ones. Additionally these dimensions were linked to their respective items, measured variables.

After conducting the CFA, the indexes, listed in Table 5.3. By observing the values, we can conclude that, firstly, the null model, as expected, provides a bad fit as indicated by the indexes. Specifically, the $\frac{\chi^2}{df}$ value (9.768) is above 5, the CFI (0.000) values is below 0.8 and finally, RMSEA (0.215) is one decimal above 0.1, all values lie within a bad fit on their respective ranges.

Similarly, the original model (4F), also suggests a bad fit. While its $\frac{\chi^2}{df}$ value was 4.594, below 5, yet above 2, the others, CFI (0.630) and RMSEA (0.138), however, was above 0.6 and 0.1 respectively, suggesting a bad fit.

TABLE 5.3: Goodness-of-fit indexes for Alternative models for the CBS. (n=190)

Step		χ^2 (df)	$\frac{\chi^2}{df}$	CFI	RMSEA
	Null Model	3174.497 (325)	9.768	0.000	0.215
1F	1 Common-Factor	1082.293 (291)	3.719	0.722	0.120
4F	4 Factors (without adjustments)	1345.917 (293)	4.594	0.630	0.138
4FW	4 Factors (with adjustments)	852.059 (264)	3.215	0.795	0.111

Because our original model suggested a bid fit for the data, we devised additional alternatives.

One of the alternatives, involved a model (1F) which, unlike the previous one, hypothesizes only 1-Common Factor, or dimension (Believability), to explain the items' variances. Surprisingly, results are slightly superior to those in 4F. First, its $\frac{\chi^2}{df}$ (3.719) is also below 5, but above 2. However, comparatively, its CFI (0.722) value is close to 0.8, the threshold for a reasonable fit, yet still below it. Finally, its RMSEA (0.120) value, while near 0.1, it is still above it, making it a bad fit.

In addition, we adjusted model 4F using hints from estimates' modification indexes. This model was named 4FW and its adjustments were as follows:

- Similarly to the 1F model, we added covariances between residuals.
- Item 25 ("The creatures learn through imitation") was changed to explain the dimension Adaption.
- While the removal of several items was suggested (considering the conventional cut threshold of 0.4), we opted only to remove item 16 ("The creatures' same-stimuli reactions change over time"). The main reason for this was due to the item's wording generated confusion.

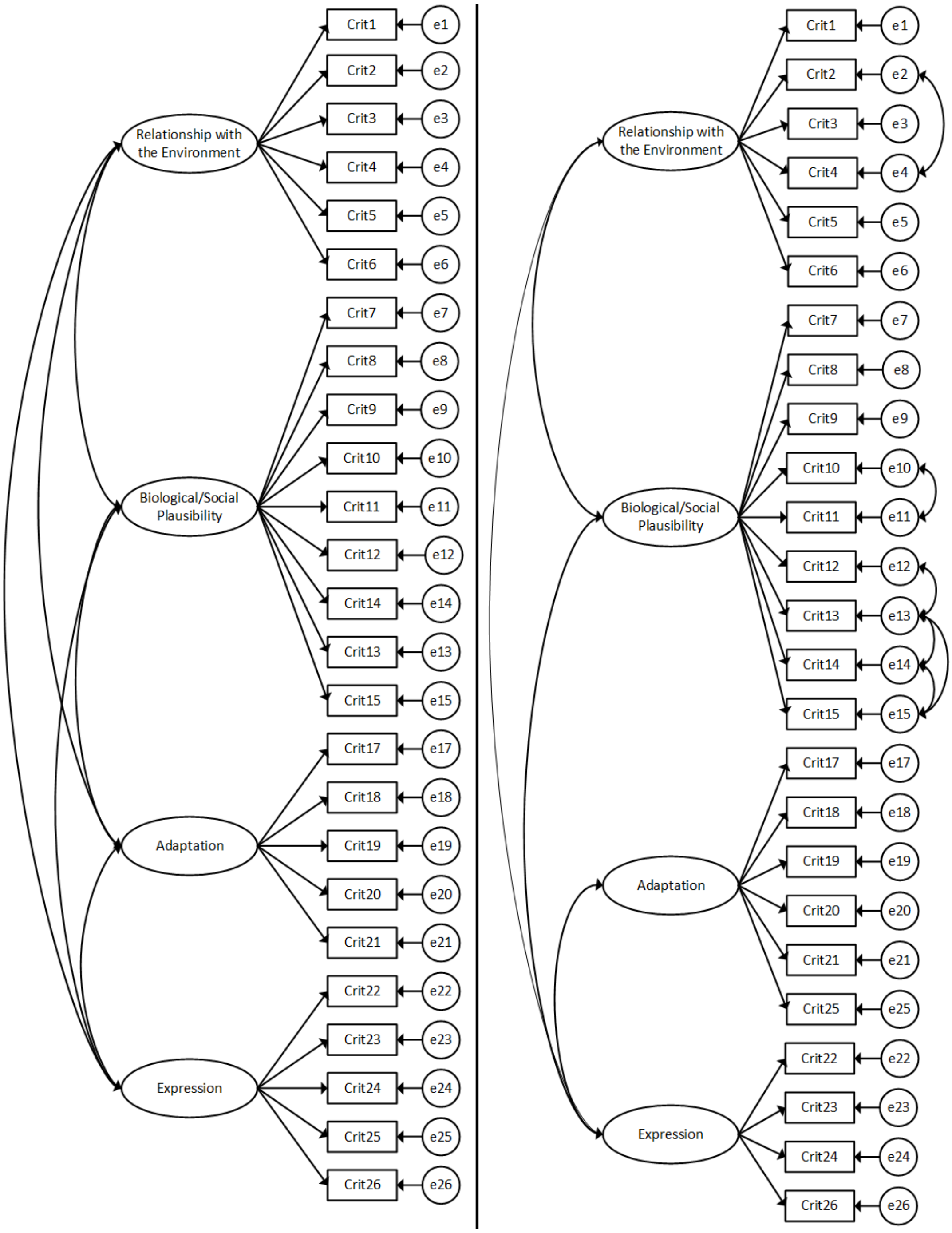


FIGURE 5.1: The Structured Equation Models of model 4F (left) as well as model 4FW (right). The scales' construct's dimensions, or latent variables, are depicted as ovals. Items, or measured variables, as rectangles and residuals are illustrated as circles. Moreover, single-headed arrows detail how the measured variables' variance is explained and double-headed arrows show covariances.

TABLE 5.4: The revised scale with the items' respective estimated loadings. The first column depicts the scale's items while the second one displays the loadings obtained after the analysis performed using a CFA.

Factor	Item	Loading
Relationship with the Environment	1. The creatures interact with the environment	0.765
	2. The creatures control their body	0.787
	3. The creatures direct their behaviors towards targets	0.633
	4. The creatures locate objects in the environment	0.783
	5. The creatures expel material	0.074
	6. The creatures' actions are appropriate to their context	0.642
Biological/Social Plausibility	7. The creatures move by themselves	0.733
	8. The creatures' motions reflect their weight/size	0.605
	9. The creatures make several simultaneous motions	0.551
	10. The creatures react to stimuli	0.646
	11. The creatures focus on stimuli	0.520
	12. The creatures coordinate with other creatures	0.256
	13. The creatures communicate with other, same-species, creatures	0.280
	14. The creatures engage in reproductive acts	0.139
	15. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	0.014
Adaptation	~~16. The creatures' same-stimuli reactions change over time~~	-
	17. The creatures learn from past events	0.376
	18. The creatures are able to apply old behaviors to new, similar, situations	0.534
	19. The creatures change the way they look with age	0.963
	20. The creatures change the way they sound with age	0.968
	21. The creatures change the way they behave with age	0.981
	25. The creatures learn through imitation	0.629
Expression	22. The creatures' expressions anticipate their actions	0.667
	23. The creatures show positive (or negative) emotions towards objects, or events	0.775
	24. The creatures show expressions to known stimuli	0.794
	~~25. The creatures learn through imitation~~	-
	26. The creatures' body are adapted to their habitat	0.335

- We also removed data outliers according to their Mahalanobis Distance p1 and p2 values (when they were both equal to zero), reducing the entries from 190 to 182.
- Finally, we removed covariances between the latent variable pairs "Relation with the Environment" and "Adaptation", and "Adaptation" and "Biological/-Social Plausibility", as they had an estimated covariance close to zero.

By comparing fit indexes with the other alternative models, it is clear 4FW yields better results: Its $\frac{\chi^2}{df}$ is the highest with 3.2, within the reasonable fit range. Additionally, the CFI (0.795) is close to 0.8 and it can be argued this value borderlines a reasonable fit. Finally, the RMSEA value (0.111) is around 0.1. However, while this model fares better than the previous ones, each of the alternatives can be either, considered bad fits or, borderline a reasonable fit, due to how their values lie either inside the ranges of a reasonable fit, or close to them. Therefore, additional tests need to be conducted to confirm these results. Taking this into account, we considered model 4FW for a revision proposal.

The original model (4F) and the one we considered (4FW) are illustrated in Figure 5.1 whilst Table 5.4 lists the revised scale and the items' respective loadings.

5.3 Second Revision

After the previous revision process, out of the original 26-item scale, 1 item was removed and another one transited to a different dimension. We then obtained:

Relation with the Environment

1. The creatures interact with the environment
2. The creatures controls their body
3. The creatures direct their behaviors towards targets
4. The creatures locate objects in the environment
5. The creatures expel material
6. The creatures' actions are appropriate to their context

Biological/Social Plausibility

7. The creatures move by themselves
8. The creatures' motions reflect their weight/size
9. The creatures make several simultaneous motions
10. The creatures react to stimuli
11. The creatures focus on stimuli
12. The creatures coordinate with other creatures
13. The creatures communicate with other, same-species, creatures
14. The creatures engage in reproductive acts
15. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.

Adaptation

16. The creatures learn from past events
17. The creatures are able to apply old behaviors to new, similar, situations
18. The creatures change the way they look with age
19. The creatures change the way they sound with age
20. The creatures change the way they behave with age
21. The creatures learn through imitation

Expression

22. The creatures' expressions anticipate their actions
23. The creatures show positive (or negative) emotions towards objects, or events
24. The creatures show expressions to known stimuli
25. The creatures' body are adapted to their habitat

Up to this point, results from the previous validation have shown, the scale still needed further research. Because the model was deemed a reasonable fit, additional studies had to be conducted as an attempt to understand why this has happened. We pondered upon several alternatives: repeating the CFA with additional subjects. A sample pool of between 300 to 500 entries would suffice. Unfortunately, due to the lack of resources, this was unfeasible. Nonetheless, an alternative would be to revise the items' statements by interviewing several subjects and try to understand if they are similarly interpreted across interviewees. Another option would be to revise the test setup altogether; for instance, using actual gameplay as an alternative to video clips. We opted to go with the latter two alternatives.

5.3.1 Administering the Creature Believability Scale v0.3

Unlike the previous two setups, for data collection on the new revised scale (v0.3), we used actual gameplay sessions in lieu of videos. This way, subjects could experience, first hand, how the creatures behaved in the videogames' worlds, try their available interactions and even change the camera perspective (when possible) to better understand any unfolding events. We took this opportunity to analyze the scale's agreement between evaluators (i.e. experts) and analyze potential problems if it yielded a 'poor' agreement. While this step is not actually contemplated on the process detailed by Spector (1992) and Hinkin, Tracey, and Enz (1997), we deemed it necessary to validate how the scale fares as an assessment method because, without a consistent measurement across evaluators in similar circumstances, its utility would be diminished.

So, the experiment was designed as follows:

1. The group of experts were prompted to play a randomly sorted set of videogames.
2. Subjects played each videogame in 30m to 1h sessions. These were monitored and recorded, where subjects were:
   (a) Explained the videogames' controls and possible interactions.

   (b) Explained the constraints of the experiment: they were not allowed to advance the videogames' narratives or watch cutscenes. They only had to interact, and observe creatures, using the videogames' available means.
   (c) Appointed which creature, or creature species, they had to pay attention to.
   (d) Prompted to rate the creature, or species, using the CBS.
   (e) At any time, subjects could reset the videogame to a previous state whenever it failed to provide creatures, regardless of the reason.

3. After the session, subjects were inquired to rate the creatures' perceived believability on a scale from 0 (Non-Believable) to 10 (Believable).
4. In the end, after the data was analyzed, subjects were brought back for a follow up interview.

We recruited experts on gaming who were 'hardcore gamers' (Poels et al., 2012). With an extensive contact, and knowledge of videogames, we conjecture this would reduce any potential entry barriers, when experts had no previous contact with a particular videogame. This way, they could focus on evaluating the creatures, rather than trying to understand the intricacies of the videogames first.

Choosing Content for Evaluation

During this experiment, we altered some of the criteria used to select videogame content. While we kept the same rule of choosing recent games (for the reasons previously explained), this time, we decided to lax on the creatures' on-screen presence time to better understand whether there were differences in the perception of the scale's items. Because of this, we included other genres not contemplated, in the previous two experiments (e.g. First Person Shooters), as creatures, in these genres, usually have a small on-screen presence. In accordance to these new criteria, the list of videogames was as follows:

- Trico from The Last Guardian (genDESIGN and SIE Japan Studio, 2016)
- The EyePet from EyePet (SCE London Studio and Playlogic Game Factory, 2009)
- Striders from Horizon Zero Dawn (Guerrilla Games, 2017)
- Armadillos from Everything (David OReilly, 2017)
- Threadfin Butterflyfishes from Abzû (Giant Squid Studios, 2016)
- Cifiarae Vasarku from No Man's Sky (Hello Games, 2016)
- Botamon from Digimon World: Next Order (B.B. Studio, 2017)
- Cats from The Sims 3 (Maxis, 2009)
- Squirrels from SimAnimals (Electronic Arts, 2009)
- Raptors from Turok (Propaganda Games, 2008)
- Alligators from Mafia III (Hangar 13, 2016)
- Boars from S.T.A.L.K.E.R.: Call of Pripryat (GSC Game World, 2010)

- Pigs from Minecraft (Mojang, 2011)
- Lyotes from StarCraft II: Legacy of the Void (Blizzard Entertainment, 2015)
- Deers from The Witcher III: The Wild Hunt (CD Projekt RED, 2015)
- A deer from theHunter (Games, 2005)
- D-Horse from Metal Gear Solid V: The Phantom Pain (Kojima Productions, 2015)
- A Rasteirist from Spore (Maxis, 2008)
- Rhinoceros from Far Cry 4 (Ubisoft Montreal, 2014)
- Lizards from Rain World (Videocult, 2017)

5.3.2 Gameplay Sessions' Results

The data gathered was analyzed through means of an Intraclass Correlation Coeficient (ICC). We first predefined a baseline agreement value of 0.6 (The threshold value for 'Good' (Cicchetti, 1994)). Then, having selected 20 videogames, we recruited 6 experts (3 above the minimum sample size required for achieving a theoretical baseline value of 0.6, with 20 observations per subject (Bujang, 2017)). The main reasoning behind using creatures as the analyzed classes was due to the fact that, in case of a 'poor' agreement value, isolating issues pertained to the creatures (e.g. being designed in such a way that hinders perceiving the scale's items) would be facilitated.

As a final note, this analysis also helped guide the follow-up interviews detailed in sub-chapter 5.3.2.

Expert Agreement Analysis

The ICC values were obtained using a Two Way Mixed approach and these are listed in Table 5.5. As can be observed, 13 out of 20 creatures obtained an ICC value within the baseline of 0.6. Furthermore, out of the former 13, 6 had values above 0.75, accounting for an 'Excellent' agreement correlation (Bujang, 2017). From the total's remainder 7 creatures, 3 had values above 0.4 falling within the 'Fair' group (Bujang, 2017) (though 2 of them were close to the 0.6 threshold) whilst the other 4 were considered 'Poor', with an ICC below 0.4 (Bujang, 2017).

Having calculated the ICC for the evaluated videogames, our next step involved understanding potential issues with the scale.

Follow-up Interviews

With such a small group of experts, we opted to use descriptive statistics, namely box-plots, to visualize potentially problematic items. These are referenced in Appendix B.1. Our initial speculation revolved around three categories:

1. Items were not interpreted correctly.
2. Experts failed to perceive an item, on a particular creature.
3. Interactions between experts differed so much, they were not able to experience the same things.

TABLE 5.5: ICC values calculated for each videogame, evaluated by 6 experts using the Creature Believability Scale. Values above the baseline 0.6 ICC are depicted in bold, and underlined.

Creature (Videogame)	**ICC**
Trico (The Last Guardian)	0.644
Eyepet (Eyepet)	0.634
Strider (Horizon Zero Dawn)	0.763
Armadillo (Everything)	0.782
Threadfin Butterflyfishes (Abzû)	0.748
Cifiarae Varsaku (No Man's Sky)	0.744
Botamon (Digimon World: Next Order)	0.288
Cat (The Sims 3)	0.575
Squirrel (SimAnimals)	0.318
Raptor (Turok)	0.318
Alligator (Mafia III)	0.781
Boar (S.T.A.L.K.E.R.: Call of Pripryat)	0.691
Pig (Minecraft)	0.571
Lyote (StarCraft II: Legacy of the Void)	0.327
Deer (Witcher III: The Wild Hunt)	0.694
Whitetail Deer (theHunter)	0.659
D-Horse (Metal Gear Solid V: The Phantom Pain)	0.745
Rasteirist (Spore)	0.673
Rhinoceros (Far Cry 4)	0.413
Lizard (Rain World)	0.689

As experts had a 'Good' agreement when scoring 13 of the creatures, our analysis primarily fell on the remaining 7. Nonetheless, we considered the former ones as well, whenever it would help our assessment. With this in mind, we first considered items with an InterQuartile Range (IQR) above 3 (i.e. high response spread and potentially low consensus), then items with more than 2 outliers and, finally, items with an IQR of 2.

While this accounts for every item in the scale, it is worth noting that some of these issues tied to certain creatures (Squirrel and Rasteirist), or certain experts (Expert 4). Following this, the next step involved comparing each of the experts' individual responses against potentially problematic items. These were grouped by their respective experts and listed in Table 5.6.

Afterwards, the experts were interviewed by first asking them to state a particular item and its interpretation followed by inquiring the reason behind giving a certain score on a particular creature.

TABLE 5.6: Potentially problematic items grouped by (Expert, Creature) tuples.

	E1	E2	E3	E4	E5	E6
Whitetail Deer		3, 13, 18, 19, 20			18	
Pig	3	3, 12, 22	3, 22	3, 22, 23, 24	3, 5, 26	3, 26
D-Horse	3	2, 3, 7	3, 9	3, 7, 15	3, 7	3, 12, 13
Alligator	18, 19, 20		22	23, 24	23	
Botamon		13, 15	2	7, 19	7	15
Cifiarae Varsaku	3, 14, 16, 19	4, 14, 15		3, 23, 24	1, 3, 6, 14, 26	3
Threadfin Butterflyfish	23	1, 22		23, 24	13, 15	23, 24
Trico		14, 15, 26	5, 26			
Raptor	23	6, 12	2		12, 23	
Rhinocerous	13	12		26		
Lizard	24	8			6, 8, 22, 23, 24, 25	
Cat		3, 5	18, 19, 20	23, 24		3
Boar	24	10, 13		23, 24	6, 11	
Strider	26	22	9	23, 24	12	
Eyepet	5	16, 18, 19, 20		14, 18, 19	1, 5, 6	
Squirrel	18, 19, 20	12, 18, 19, 20	18, 19, 20	18, 19, 20	8, 18, 19, 20	18, 19, 20
Lyote	3, 24	3, 6	3	3	1, 3, 4, 6, 8, 9, 11, 26	3
Armadillo	6	8			6	

Deer	3	3, 12	3	3, 15	1, 3, 4, 11, 12, 16	3
Rasteirist	14, 15, 18, 19, 20	13, 14, 15, 18, 19, 20	14, 15, 18, 19, 20	14, 15, 18, 19, 20	5, 14, 15, 18, 19, 20	14, 15, 18, 19, 20

From the interviews, it became evident that, on certain videogames, experts did not entirely perceive items 3, 14, 15 and 18, 19 and 20. This was because they had different gameplay experience. For instance, when evaluating the Pig creature from Minecraft (Mojang, 2011), E3 stated item 3 meant "exactly what it is written... they move based on targets... either in their direction... or they steer clear away from them". When asked to explain the reasoning behind his answer, he replied "I did not see pigs reacting to what I did. I killed one... They did not run, they didn't do anything...". However, while E4 explained the statement meant "every time a creature does something, or has a reaction, it is directed towards on an object or a specific situation", he offered a different account of his experience: "In Minecraft, they directed themselves towards each other... to me... to the objects I laid on the floor... to fences".

Experts also failed to perceive items 4, 16 and 17. Though, the reason revolved around actually perceiving the activities themselves (this was even stressed in items related to learning activities).

The remainder items were considered either vague or confusingly worded, the experts' impressions are summed as follows:

1. **The creatures interact with the environment** - Expert 5 had difficulty in discerning which elements could be considered 'environment'. When prompted to explain some of his answers in this particular item, he mentioned "It was difficult to understand if creatures interacted with the 'environment' or just specific objects". This response made us speculate the word 'environment' might be too vague and invoke some ambiguity.

2. **The creatures control their body** - Experts 2 and 3 were unsure whether the sentence meant if the creature was presented in such a way it appeared to be controlled by an exterior entity (akin to a puppet) or, lacking motor coordination.

5. **The creatures expel material** - Expert 1 claimed this item lacked an indirect counterpart, similarly to the reproduction ones: He stated he was uncertain to score this whenever he saw expelled material but not the act itself.

6. **The creatures' actions are appropriate to their context** - When asked what the statement meant, Expert 5 stated "I think the actions define the context". Similarly, when Expert 1 was asked to clarify why he gave a particular rank to this statement, on a particular creature, he responded "The creature only wanders around and I can't understand if this is sufficient taking into account the videogame's gameplay mechanics". From these statements, we discerned both experts were unable to decide whether it meant the environment surrounding a creature or the circumstances surrounding the actions.

7. **The creatures move by themselves** - Experts 2, 4 and 5 inquired if this sentence took into account following behaviors, claiming these would require mimicking a target's trajectory, or if the movement should be independent of other entities' trajectories.
8. **The creatures' motions reflect their weight/size** - Experts 2 and 5 considered this sentence's wording confusing.
9. **The creatures make several simultaneous motions** - There were two moments where experts failed to understand this item. First, Expert 3 did not correctly understand what it meant to make 'simultaneous motions'. Second, Expert 5 found it difficult to perceive these motions when observing the creature's animations.
10. **The creatures react to stimuli** - Unlike the previous sentences, the main problem with this item was caused by videogame limitations. Because these are not programmed to account for every possible situation, and interaction, there were times where creature reacted to some stimuli but not others. This was particularly noted by Expert 2.
11. **The creatures focus on stimuli** - Here, Experts 2 and 5 argued that 'focusing' was a type of reaction. As such, this lead to difficulties in distinguishing between one and the other.
12. **The creatures coordinate with other creatures** - Experts 1, 2 and 5 did not comprehend to what extent some behavior could be considered coordinating. Furthermore, they felt confused with the inconsistent expressions. Namely, this item states 'other creatures' whilst the following one declares 'same-species, creatures'.
13. **The creatures communicate with other, same-species, creatures** - See item 12.
22. **The creatures learn through imitation** - The main difficulty in this item laid in the extent required to considered something was learned. Furthermore, Expert 2 was unable to decide if imitating once was sufficient or if creatures had to reproduce what they had imitated in a new context.
23. **The creatures show positive (or negative) emotions towards objects, or events** - Here, Experts 1, 3, 4 and 6 were under the erroneous impression this only accounted for facial expressions. To counter this, Expert 5 suggested clarifying the types of expressions considered (for example, body language, vocal sounds, etc.).
24. **The creatures show expressions to known stimuli** - Same as 23.
25. **The creatures' expressions anticipate their actions** - Same as 23.
26. **The creatures' body are adapted to their habitat** - This item was also considered confusing by Experts 2, 3 and 6. Moreover, Experts 1, 4 and 5 alleged they lacked knowledge in biology to be able to correctly assess the creatures' adaption to a given habitat.

During this phase, additional feedback was also provided: Experts 1 and 5 found some statements required an extensive amount of time to assess whether they happened, or not (namely anything related to learning). This was hampered by the lack of any clear indication that a given expert maxed out every possible situation/interaction the game supported.

Another issue found, during the setup, involved the vocabulary chosen for the scales. Up to this point, the scale used an Agreement semantic (Brown, 2010) yet, Expert 6 suggested either a 'Not Applicable' option or a more fitting semantic as they struggled mapping any occurrences with the amount of agreement.

With the impressions given by the subjects during the interviews, it became apparent that some of the scales' sentences were still not as accessible as we originally anticipated. Therefore, an additional effort was made as a means to clarify the items. This resulted in a third revision of the scale.

5.4 Third Revision

From the knowledge gathered, the items listed previously were rewritten. An additional change was centered around the scales' points. We decided to use a Likelihood semantic (Brown, 2010) - 'Not at all', 'Slightly', 'Somewhat', 'Moderately' and 'To a Great Extent' - instead of the previous one. With these points, we conjectured they provided a more direct mapping between occurrences and the items, as well as not needing an additional 'Not Applicable' option. With this in mind, the items were also revised to better fit the scale's new points. The revision is as follows:

Relation with the Environment

1. The creatures interact with surrounding inanimate elements
2. The creatures have a functional motor coordination
3. The creatures direct their movement towards/away from targets
4. The creatures are able to find objects in the environment
5. The creatures are able to expel organic waste
6. There are signs of previous expelled organic waste, such as feces, urine, goo, etc.
7. The creatures' actions are appropriate to their circumstances

Biological/Social Plausibility

8. The creatures are able to move independently
9. The creatures' agility reflect their weight
10. The creatures' body parts move concurrently during a motion
11. The creatures react to environmental triggers
12. The creatures are aware of environmental triggers
13. The creatures can perform tasks by coordinating with other creatures, or humans
14. The creatures audio-visually communicate with other creatures, or humans
15. The creatures engage in reproductive acts
16. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.

Adaptation

17. The creature learns better responses for repeated events
18. The creature uses past experience to respond to novel events
19. The creatures change the way they look with age
20. The creatures change the way they sound with age
21. The creatures change the way they behave with age
22. The creatures can imitate others to develop their own behaviors

Expression

23. The creatures' body language anticipate their actions
24. The creatures show positive (or negative) emotions towards objects, or events
25. The creatures' body language change due to known stimuli
26. The creatures' bodies seem adapted to their habitat

Before deploying this revised scale, a second interview was conducted to validate its statements. This was meant to confirm if a different population sample was able to interpret, as intended, the revised items.

5.4.1 Group Interview

We conducted a group interview following the steps detailed in Krueger, 2002. To this extent, six people were recruited, with an additional one to moderate it. Moreover, we recorded the interview for later analysis. Its process was as follows:

1. Participants were introduced to the context behind the scale: its objectives and underlying research scope. Furthermore, they were explained how the interview was going to be conducted.
2. Participants were introduced to each other.
3. Each of the scale's statements was introduced, one at a time.
4. On each statement:
   (a) The moderator started by reading it and showing an example video, whenever it helped clarify the statements' intent.
   (b) Participants were prompted, one at a time, to list concrete situations they would associate with the statement at hand.
   (c) The moderator then sought, when needed, to clarify participant interventions on their meaning.
   (d) When there were conflicting views, the moderator promoted debate between the participants.
5. The recorded session was analyzed as presented and the information it provided was used to revise the scale.

Group Interview Analysis

During the session, the group was able to associate, with most items, examples close to their intended interpretation. However, there were 6 items considered to be of a more difficult interpretation:

1. **The creatures interact with surrounding inanimate elements** - Most participants failed to understand the intention behind the expression 'inanimate elements' as they associated this with objects (i.e. a bone), instead of linking it with the environment itself. In fact, they promptly discarded 'environment', as they claimed it wasn't present in the statement, focusing instead on objects with an utilitarian value for the creature. This interpretation lead to discard objects such as grass, which is part of an environment, but, in the participants opinion, does not have an immediate utilitarian value. We decided to revert the expression to 'environment' yet, we kept the adjective 'surrounding' for clarity.

2. **The creatures have a functional motor coordination** - While, initially, participants were able to identify situations where a creature could have motor coordination, the word 'functional' caused them to reconsider their initial impressions. The debate then shifted towards the semantics behind a functional motor coordination. Because of this, that word was removed.

8. **The creatures are able to move independently** - Similarly to the reactions given in the previous interview, in SubChapter 5.3.2, participants were unable to decide when does a creature move 'independently'. In particular, the word chosen to qualify the movement incited participants to think of following behaviors, going beyond the intended scope of the statement (the creature's ability to move without the aid of external forces). For this reason, we chose to replace 'independently' with 'autonomously'.

9. **The creatures' agility reflect their weight** - While participants successfully identified situations within the statement's intent, they deemed it incomplete without including 'size', or a similar concept. We decided to replace it with 'volume', following the suggestion of one of the participants.

10. **The creatures' body parts move concurrently during a motion** - In general, akin to the previous interviews, participants did not understand the meaning behind the sentence. However, after some explanation, they were able to identify situations in line with the intention of the item. With this in mind, we decided to simplify the wording of the statement.

12. **The creatures are aware of environmental triggers** - Changing the word 'focus' to 'aware' yielded the same results: participants identified it as a type of reaction. With this in mind, we decided to remove this item altogether.

From the group interview, it became apparent that the previous iteration of the scale did not completely clarify all of its sentences. So, a fourth revision was put in place, again with the goal to simplify the items further, towards a more accessible language.

5.5 Fourth Revision

The previous revision process resulted in the following scale:

Relation with the Environment

1. The creatures interact with surrounding environment
2. The creatures have motor coordination
3. The creatures direct their movement towards/away from targets
4. The creatures are able to find objects in the environment
5. The creatures are able to expel organic waste
6. There are signs of previous expelled organic waste, such as feces, urine, goo, etc.
7. The creatures' actions are appropriate to their circumstances

Biological/Social Plausibility

8. The creatures are able to move independently
9. The creatures' agility reflect their volume
10. The creatures react to environmental triggers
11. The creatures can perform tasks by coordinating with other creatures, or humans
12. The creatures audio-visually communicate with other creatures, or humans
13. The creatures engage in reproductive acts
14. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.

Adaptation

15. The creatures learn from past events
16. The creatures use past experience to overcome novel situations
17. The creatures change the way they look with age
18. The creatures change the way they sound with age
19. The creatures change the way they behave with age
20. The creatures can imitate others to develop their own behaviors

Expression

21. The creatures' body language anticipate their actions
22. The creature shows signs of emotional reaction to objects, events and other creatures
23. The creatures' bodies seem adapted to their habitat

After conducting the initial EFA, followed by a CFA and 2 additional validation processes, we assumed this fourth revision to be stable enough to deploy for further validation, on a practical context, that will be presented in Chapter 6.

5.6 Creature Believability Spectrum

Up until this point, we have been discussing the CBS and, while it outputs a believability index to help assess creatures, this can be supplemented by a more visual framework. As an example, the latter was provided in the context of uncanny valley (Edward, Yifan, and Shanshan, 2007) where the authors argue their study can "help understand how virtual characters fit in Mori's context of human-like appearance and familiarity" (Edward, Yifan, and Shanshan, 2007). Such frameworks can help designers understand how to improve preexisting designs, but also compare their own designs against the frameworks' elements.

We theorize one such artifact can aid in interpreting how each dimension, and statement, contribute to overall believability perception. So, during the second revision of the CBS (subsection 5.3), the data collected was additionally used to create this supplementary artifact. It came in the form of a *believability spectrum*, a referential list to help compare videogame creatures. Besides providing a visual framework, these artifacts also supply a possible use case for the CBS and similar studies can give interesting perspectives when analyzing the state-of-the-art.

5.6.1 Creature Believability Spectrum Construction

Similar to the spectrum concept presented by Edward, Yifan, and Shanshan (2007), the *Creature Believability Spectrum* consists in an axis ranging from 0 (Non-Believable) to 10 (Believable) where creatures are sorted according to their perceived believability index. A single believability index can be calculated by the weighted sum of the rates obtained after applying the CBS, whilst the weights are extrapolated from the items' loadings, obtained from the previous CFA. These are listed in Table 5.7.

Indexes can be calculated for either their overall 'believability', in which case all of the items are used, or, for a certain dimension where, conversely, only its items are considered.

5.6.2 Spectrum Construction Results

As mentioned, the data used to develop this spectrum was collected during the experiments conducted in subsection 5.3, where 6 experts evaluated 20 creatures, using the CBS, each from a different videogame source. For each creature, the median of each of its items' rates were calculated. This was used to obtain the weighted sum, first according across all of the scale's dimensions, then per each one. This allowed us to understand how creatures fared according to each dimension and how the dimensions contributed to the creatures' overall indexes.

We also plotted the comparison between the best, worst and medium creatures as well as their individual indexes, per dimension. Though these were not directly used in this analysis, as an example, Figure 5.2 illustrates the comparison between the best, worst and medium creatures, according to their indexes. The complete list of graphs are listed in Appendix B.2.

TABLE 5.7: The CBS v0.2 items' weights. Weights are transcribed from the loadings calculated in Table 5.4

Item	Weight
1	0.765
2	0.787
3	0.633
4	0.783
5	0.074
6	0.642
7	0.733
8	0.605
9	0.551
10	0.646
11	0.520
12	0.256
13	0.280
14	0.139
15	0.014
16	0.376
17	0.534
18	0.963
19	0.968
20	0.981
21	0.629
22	0.667
23	0.775
24	0.794
25	0.335

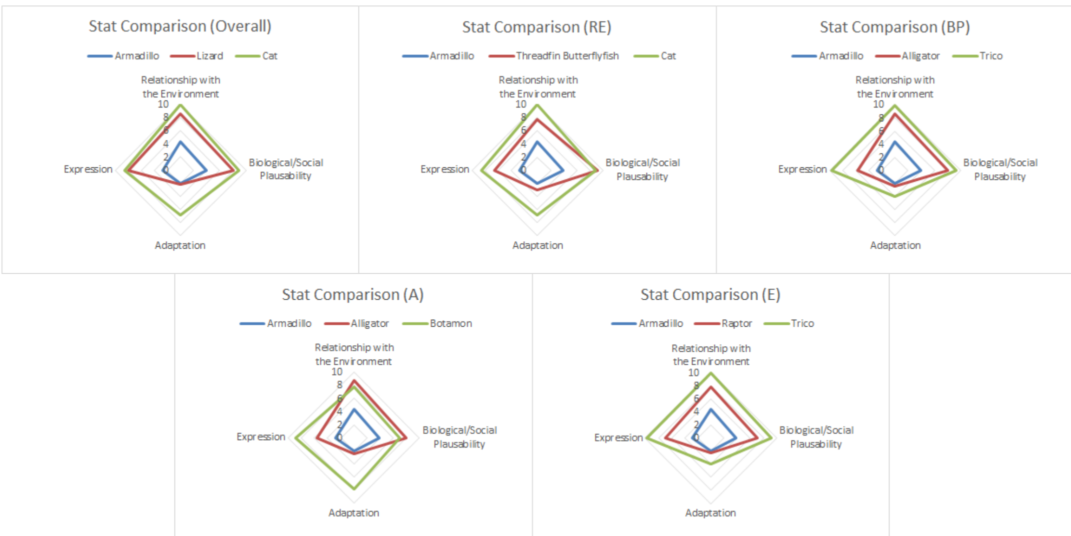


FIGURE 5.2: Comparison between the best, worst, and medium creature. From left to right, and top to bottom, these take into account their overall 'Believability', 'Relationship with the Environment', 'Biological/Social Plausibility', 'Adaption' and 'Expression' indexes respectively.

As the creatures, used in the study, play different roles, it is important to first understand these roles as they may shed light upon their perceived believability indexes:

- **Trico (The Last Guardian)** - The Last Guardian is an action-adventure videogame, with puzzle elements, where the player takes on the role of a boy stranded on a mysterious, hostile, land where he meets a gargantuan griffin, named Trico. In this videogame, it helps the player during most of the videogame, either by assisting him solve puzzles, traversing paths or defending him from attacking monsters. Furthermore, the videogame's narrative is built towards cultivating a friendship between it and the players' avatar.
- **Eyepet (Eyepet)** - This is an augmented reality life simulation videogame. In it, players must take care of an alien mammal, named Eyepet. To do this, they are tasked to feed and play with it.
- **Strider (Horizon Zero Dawn)** - This is an Action Role-Playing Game (RPG), with open-world elements, where players take on the role of a girl named Alloy in a viking-inspired post-apocalyptic world. In this world, machines have replaced most of its fauna and Striders are horse-like herbivores the player can hunt for parts, or mount for transportation.
- **Armadillo (Everything)** - Everything is a life simulation videogame, with a procedurally generated universe. While the player can control any object, including creatures, as he pleases, the videogame itself can 'play on its own', if left unattended for a period of time. This is because the videogame is "about the things we see, their relationships, and their points of view. In this context, things are how we separate reality so we can understand it and talk about it with each other" (Muncy, 2016). Creatures, in general, and armadillos, in particular, wander, by rolling, around the videogame's several planets.
- **Threadfin Butterflyfishes (Abzû)** - An adventure videogame focusing on scuba diving, Abzû follows the journey of an unnamed diver exploring the ocean. During his/her journey he/she stumbles upon several species of fish, one of them being the Threadfin Butterflyfish. These fishes decorate the videogame's environment.
- **Cifiarae Varsaku (No Man's Sky)** - No Man's Sky is an action-adventure survival videogame. As players explore planets, they encounter several procedurally generated creature species inhabiting them, one of them being the 'Cifiarae Varsaku'. These herbivore aliens can be hunted for materials, or befriended, in which case they expel minerals the player can use.
- **Botamon (Digimon World: Next Order)** - Based on the Digimon franchise, Digimon World: Next Order is a Japanese RPG. This videogame combines RPG elements with life simulation so, players are meant to raise digimon creatures by feeding, playing, evolving and training them. The latter two are used to pin your pet against hostile digimons.
- **Cat (The Sims 3)** - The Sims 3 is also a life simulation videogame. While its primary purpose is to control, or influence, the lifes of virtual human-beings, there is an expansion that introduces animals. In The Sims 3, animals can be controlled by the player, similar to their human counterparts yet, they can also 'live' autonomously, without any kind of player intervention.

- **Squirrel (SimAnimals)** - Another life simulation videogame. Though, the main difference between this one and the latter is that this one focuses on raising animals in a forest environment. The player can feed and engage several types of animals, one of them being squirrels.
- **Raptor (Turok)** - Turok is a First Person Shooter (FPS) where players assume the role of a native-american marine, stranded on an unknown planet, populated by Dinossaurs. In this videogame, Raptors are presented as targets the player must hit in order to progress.
- **Alligator (Mafia III)** - A third-person shooter set on a fictional American city, based on New Orleans, during the 1970s. Taking on the role of an African-American criminal, seeking revenge for the murder of his father figure, players must build a crime empire by developing various rackets. The only creature present in this videogame is an alligator and it only serves two purposes: to help the player dispose the bodies of fallen enemies and as a 'barrier' between water channels, hindering players who attempt to cross them without a proper transportation.
- **Boar (S.T.A.L.K.E.R.: Call of Pripryat)** - This is a FPS with RPG elements, based on a science-fiction Russian novel. In S.T.A.L.K.E.R.: Call of Pripryat, players are put in a dystopian Chernobyl, populated by several mutated monsters, in the aftermath of its real-world counterpart nuclear disaster. Boars are one of the monsters that roam the countryside, attacking the player, or any other being, unprovoked.
- **Pig (Minecraft)** - Minecraft is a survival videogame. As players explore a block world, they can raise pigs similarly to their real life counterparts and feed, breed and even slaughter them.
- **Lyote (StarCraft II: Legacy of the Void)** - A real-time strategy (RTS) videogame where players take on the role of either a human general, or an alien one, capable of creating, and commanding, troops to fight in skirmishes. While most of the videogame's creatures are present as one of the factions' army units, there are others, such as Lyotes, who play a pure decorative role, meant to bring life to the videogame's environments.
- **Deer (Witcher III: The Wild Hunt)** - This is another Action RPG. In this one, players embody Geralt, a witcher (monster hunter), in a Slavic fantasy world. As players roam through this world, they can encounter deer and hunt them for crafting materials.
- **Whitetail Deer (theHunter)** - theHunter is a hunting simulation videogame. Players are tasked to hunt several animals using real-life approaches (baiting, tracking, etc.). Whitetail Deer, much like the other animals in this videogame, are one of the targets the player must hunt.
- **D-Horse (Metal Gear Solid V: The Phantom Pain)** - In Metal Gear Solid V: The Phantom Pain, the players embody a special operations expert, named Big Boss. During his missions, he is able to deploy one of several companions, to help him. One of these companions is D-Horse. It is used to help Big Boss quickly traverse the videogame's terrain and serve as distraction for enemies.
- **Rasteirist (Spore)** - Spore is another RTS videogame, with life-simulation elements. Players take on the role of a deity responsible for evolving life, going

from individual cells to multicellular beings capable of space exploration. During the videogame's 'prehistoric' phase, the players encounter several other species of creatures, one of them being the Rasteirist. These creatures are herbivores and can befriend the players' creatures, if engaged through dance.

- **Rhinoceros (Far Cry 4)** - Another FPS with RPG elements, set in a fictional Himalayan territory, governed by a dictator. Caught in the middle of a power struggle, players must join either of the conflict's factions. The videogame recreates the Himalayan biome with several fauna, include a species of Rhinoceros. These roam the world, are territorial and can be hunted, for crafting materials, or baited to attack the player's enemies.
- **Lizard (Rain World)** - Rain World is a survival platformer. As players take on the role of a bipedal creature, called 'Slug-cat', they must survive in a hostile world while searching for a way to get back to 'Slug-cat's' lost family. During this journey, they encounter several carnivore creatures, some of them known as Lizards. These roam through the world, feeding on any living being they may encounter or even engaging each other in territorial behaviors.

Spectrum and Chart Analysis

Figure 5.3 illustrates the believability spectrum, taking into account the weighted sum across all dimensions while Figures 5.4, 5.5, 5.6 and 5.7 show the distribution of indexes along each value spectrum for the dimensions 'Relationship with the Environment', 'Biological/Social Plausibility', 'Adaptation' and 'Expression' respectively.

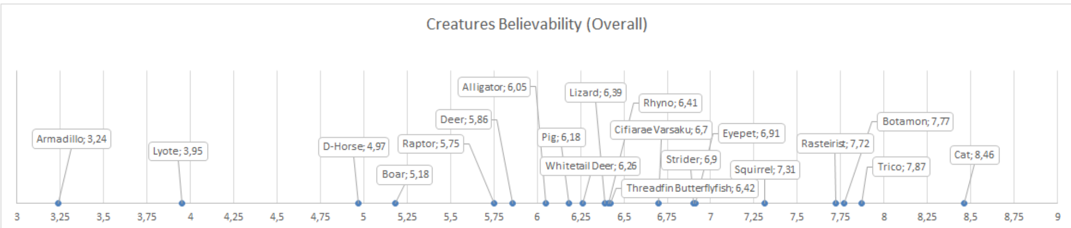


FIGURE 5.3: Creature Believability Spectrum calculated from the data gathered from 6 experts.

From a first analysis, 17 creatures (85%), have believability indexes above 5, with 9 of them (approximately 50%) scoring between 6 and 7. It appears that, in most of these videogames, developers have made some effort to make their creatures seem, at least, somewhat believable, regardless of their role:

- **Cat** (8.46) - In The Sims 3 Maxis, 2009, cats are considered family members. Players get to interact extensively with them as they would with any other Sim. Cats are able to interact with virtually any object, and animal, or Sims, in the videogame, learn tricks, procreate and expel materials. Furthermore, these creatures have an extensive on screen presence.
- **Trico** (7.87) - Trico is present throughout most of the videogame (though it is absent during a couple of sections), cooperating with players to help them progress. Players are able to feed it, teach it tricks and communicate with it, issuing commands. During the course of the videogame, Trico is able to focus on particular environmental objects, prompting players to pay attention to them, and interact with them. Additionally, it is able to expel materials.

Near the end, it is also shown Trico was able to procreate. Visually, Trico is agile, capable of creeping, jumping and pouncing (something consistent with its feline inspiration). Similarly to the previous videogame, Trico has a substantial on-screen presence.

- **Botamon** (7.77) - Botamon is one of the companions players can have and caretake. To achieve the latter, they can feed it, lay it to rest and take it to a bathroom. Players are also able to communicate with it, issuing orders, and use it on fights. During these fights, Botamon autonomously fights its assailants, taking into account the players' suggestions. However, outside of fights, it follows players around. Botamon, similarly to the other companions has a limited life-span so, when it dies, players must choose a different companion. Because of this, players end up, on average, with several different companions, during their playthrough. Therefore, this creature is meant to have a somewhat medium on-screen presence.
- **Rasteirist** (7.72) - Rasteirists are one of several species the player can either befriend or antagonize. Opting the former route, causes them to appear more frequently to help players and, eventually, get assimilated by them. The latter path however, leads to the species' extinction. They react to players and predators, can procreate and communicate through dancing. Additionally, they wander around their nest and occasionally graze. As such, the creatures' on-screen presence is, at most, extensive though, this depends on the players' actions.
- **Squirrel** (7.31) - Similar to cats, and the Eyepet, squirrels are one of the animals players must take care of, during their play session. They can be fed and are able to react to players, other animals and the objects in the environment. They can also procreate. As a result, their on-screen presence is substantial.
- **Eyepet** (6.91) - The whole videogame centers around taking care of an Eyepet. Players can feed, wash and play with them. The videogame comes with several toys however, Eyepets can also react to players' gestures. Additionally, they are able to mimic players' drawings. They are the most persistent element on screen.
- **Strider** (6.9) - Striders play a passive role in this videogame. They wander through its environment, occasionally grazing in herds. They react to sounds and the players' presence. If players are in plain sight, they will attack them, otherwise they will search through the last location they spotted them. Players can mount or hunt Striders for materials. Their on-screen presence is medium.
- **Cifiarae Varsaku** (6.7) - These creatures are present in one of No Man's Sky's planets. While they do not procreate, they come in four varieties: adult, cub, female and male. They are passive and wander around the environment. They react to players or predators, fleeing in both cases. Yet, they can also befriend players if they are fed by them. In these situations, they will expel organic material as an appreciation gift. Their on-screen presence can go from small to extensive, depending on how the player interacts with them.
- **Threadfin Butterflyfishes** (6.42) - These fishes in Abzû are present throughout the initial sections of the videogame. They swim in schools-of-fish and flee if players approaches them. However, whenever players use a special swimming

technique, they will mimic them. Since species get replaced throughout sections of the videogame, they have a small on-screen presence.

- **Rhinoceros** (6.41) - These animals are one of the creatures present throughout the videogame's environment. These wander passively yet, they will attack when provoked or whenever humans comes into their line-of-sight. Nonetheless, rhinoceros can flee if players scare them with loud gun shots. As such, their on-screen presence is small to medium.
- **Lizard** (6.39) - Lizards roam aimlessly the videogame's world, relying on their hearing and eye-sight to search for food. While they can be killed, their strength and resistance makes it difficult. As such, players are meant to hide/flee from these creatures as soon as they see them. If they spot players, they spend some time tracking them. Nonetheless, players are not the only entities Lizards react to. They can hunt other creatures or go into territorial fights with other Lizards. Furthermore, they can flee from larger predators. Their on-screen is thus medium to large.
- **Whitetail Deer** (6.26) - These creatures, as any other in this videogame, are only meant to be hunted. Nonetheless, they are only visible after a player successfully tracks them through footprints, or other signals. When this happens, players are prompted to kill them. However, players can observe these creatures, in which case, they will roam, occasionally grazing and reacting only to gun shots and the players' sight. If players take too much time killing them, they will disappear through the forest, appearing moments later. Their on-screen presence is mall to medium.
- **Pig** (6.18) - In Minecraft, pigs are passive blocky creatures that roam the videogame. Pigs can wander, be fed and procreate, producing piglets. Moreover, they can follow players if they are lured with food. They can also be herded and slaughtered for food. Their on-screen presence between medium and large.
- **Alligator** (6.05) - Unlike most of the creatures present throughout this list, alligators play the role of barriers hindering players from crossing bogs/swamps without proper transportation. Additionally, they serve a gameplay purpose: they facilitate disposing bodies a player might have. They stand idle, most of the time, wandering occasionally. However they will follow, and attack players, if they swim by their line-of-sight. They have a limited amount of reaction targets, reacting only to bullets, boats, dead bodies and players. With this in mind, alligators have a small on-screen presence.
- **Deer** (5.86) - Deers in Witcher III: The Wild Hunt are environmental props. Though they can be hunted for materials, they usually roam the videogame's world passively, fleeing as soon as players approach them. They also herd with other deer and occasionally graze. Their on-screen presence is small.
- **Raptor** (5.75) - Raptors are a weak enemy type players are force to kill in order to proceed with the videogame. They stay idle and pounce at players as soon as they see them. They'll keep pouncing until they are dead. Due to their low health, these creatures have a small on screen presence.
- **Boar** (5.18) - Similarly to Raptors, Boars are an enemy type in S.T.A.L.K.E.R.: Call of Pripryat. This means they are meant to be killed as soon as they attack players. They roam the videogame's world, when unprovoked, and will charge

at anything that crosses their line-of-sight. They only stop when killed or when their target leaves their line-of-sight. Unlike the previous creatures, they have a small to medium on-screen presence.

- **D-Horse** (4.97) - This creature is one of the companions players can take on missions. Because of this, D-Horse's on-screen presence is intended to be, at most, substantial. During a mission, D-Horse follows players around. It can react to enemies, their vehicles and players. In addition, players can mount it and ordered it around. This includes telling it where to go, to stay in place and defecate.
- **Lyote** (3.95) - Lyotes are mere props, occasionally appearing in some of the videogame's levels. They wander around the level and can be killed by players. Their on-screen presence is small, mostly because they are placed outside of the levels' main path and don't offer much interactions.
- **Armadillo** (3.24) - Armadillos are one of the creatures that roam the universe inside Everything. They hop around on alternate sides, individually or in formations, and players can follow them. Their on-screen presence can be, at most, large.

The main exceptions in this assessment are D-Horse and the armadillo who were ranked 4.97 and 3.24, respectively. This was in spite of their on-screen presence. Though the latter might be explained by the author's intent to provide a philosophical experience rather than an immerse one (Muncy, 2016).

By analyzing each of the dimensions individually, it becomes apparent the creature's aspects are tackled heterogeneously: even though a creature may score high on overall believability, this was not necessarily true for each of the scale's dimensions. This is supported by the weights in Table 5.7.

The 'Relationship with the Environment' dimension, listed in Figure 5.4, is where creatures fare better, with 16, out of 20, ranking above 7. This seems to suggest that the interactions a creature has, with their environment, are almost a minimum requirement for creature implementation. Nonetheless, this effort appears to rise as creatures play a more prominent role, such as Trico, or have an extended on-screen presence.

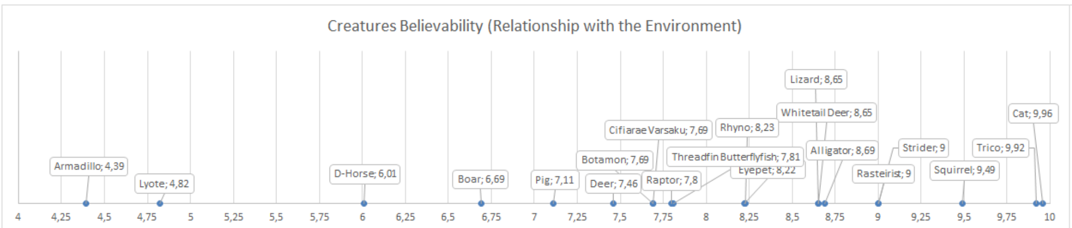


FIGURE 5.4: Creature Believability Spectrum, sorted in respect to the 'Relationship with the Environment' dimension, calculated from the data gathered from 6 experts.

Similar to the previous dimension, in 'Biological/Social Plausibility' (Figure 5.5), 19 creatures (95%) rank above 9, with 10 rated between 8 and 9. From these results, it appears this dimension is another key aspect in creature designs, with an incidence proportional to the creatures' roles and/or on-screen presence.

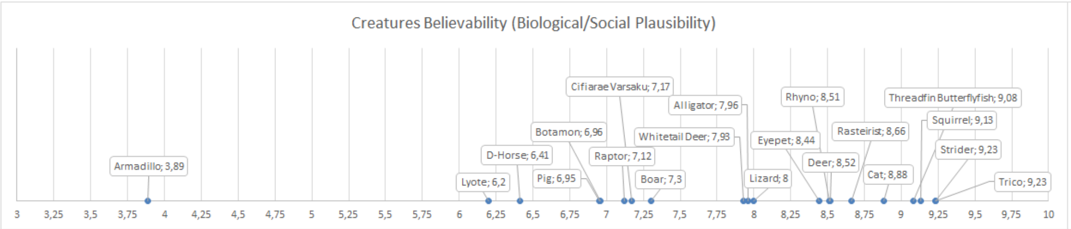


FIGURE 5.5: Creature Believability Spectrum, sorted in respect to the 'Biological/Social Plausibility' dimension, calculated from the data gathered from 6 experts.

Unlike the other dimensions, 'Adaptation', as illustrated in Figure 5.6, is where most creatures fail to be perceived as believable. Even creatures as high ranking as 'Trico' scored an index below 4. From a development perspective, it seems logical that providing creatures with adaptation elements is an added effort, not justifiable when they are mere decorations, targets, or the videogame's temporal context is not lengthy enough to show them mature. However, when simulated ecosystems play an integral part of the videogame, rather than a supporting one, creatures fare better. Furthermore, this increased ranking is also present in videogames where 'Adaptation' elements play an essential role, such as Digimon World: Next Order (B.B. Studio, 2017).

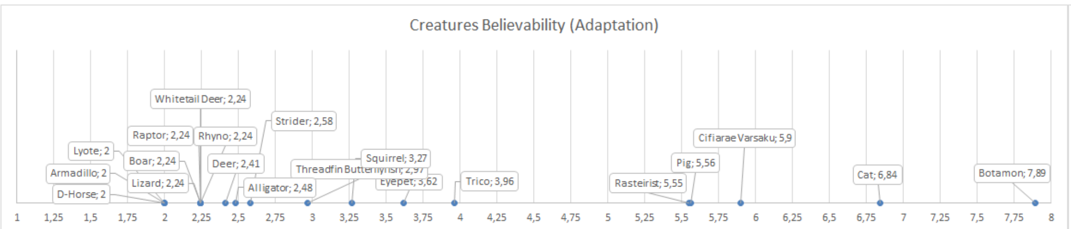


FIGURE 5.6: Creature Believability Spectrum, sorted in respect to the 'Adaption' dimension, calculated from the data gathered from 6 experts.

In the 'Expression' dimension (depicted in Figure 5.7), there are 17 creatures whose indexes span between 5 and 10. This is unlike the previous ones, where values were more concentrated in fewer ranks. By looking at the high scoring creatures, we are lead to conjecture that the ones who meant to be viewed as pets, or even friends, are more expressive than the ones who are not. This seems to be in line with the literature's perspective on believability (Thomas and Johnston, 1997; Frank, Stern, and Resner, 1997; Warpefelt, Johansson, and Verhagen, 2013; Rosenkind, Winstanley, and Blake, 2012). As an example, Fumito Ueda, The Last Guardian genDESIGN and SIE Japan Studio, 2016 videogame designer, explained, in an interview, that his team underwent an intense iteration phase to make the creature look as expressive as it could (Gigazine, 2017). In our experiment, the videogame's creature, Trico, scored the highest value: 9.74. Hence, it is our understanding that, ultimately, the amount of expressiveness a creature may have, can be strongly related with the videogame design intent..

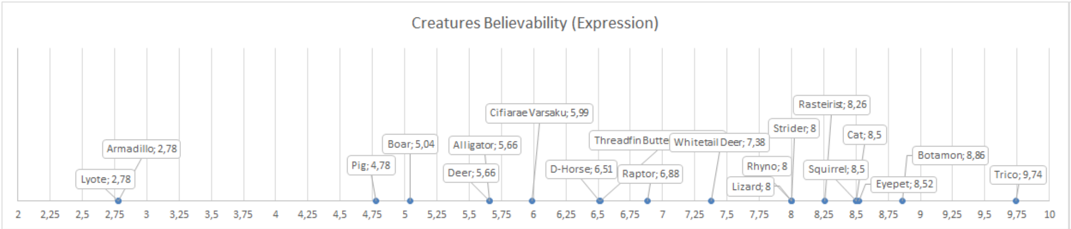


FIGURE 5.7: Creature Believability Spectrum, sorted in respect to the 'Expression' dimension, calculated from the data gathered from 6 experts.

Another point in this analysis is that, by plotting each of the believability's dimensions' indexes and sort them by their overall believability rate, additional interesting properties can arise. Figure 5.8 shows that creatures scoring high in 'Adaptation', score low in the remainder dimensions. This is true for the pig, Cifiarae Varsaku, Botamon and, to some extent, the cat. One possible explanation for this may be due to, given the videogame's design intent, the additional development required for the former justifies the negligence of the latter. Another possible reason may be that the videogame's design drives the players' attention towards aspects of the creatures' life, or apparent cognitive abilities.

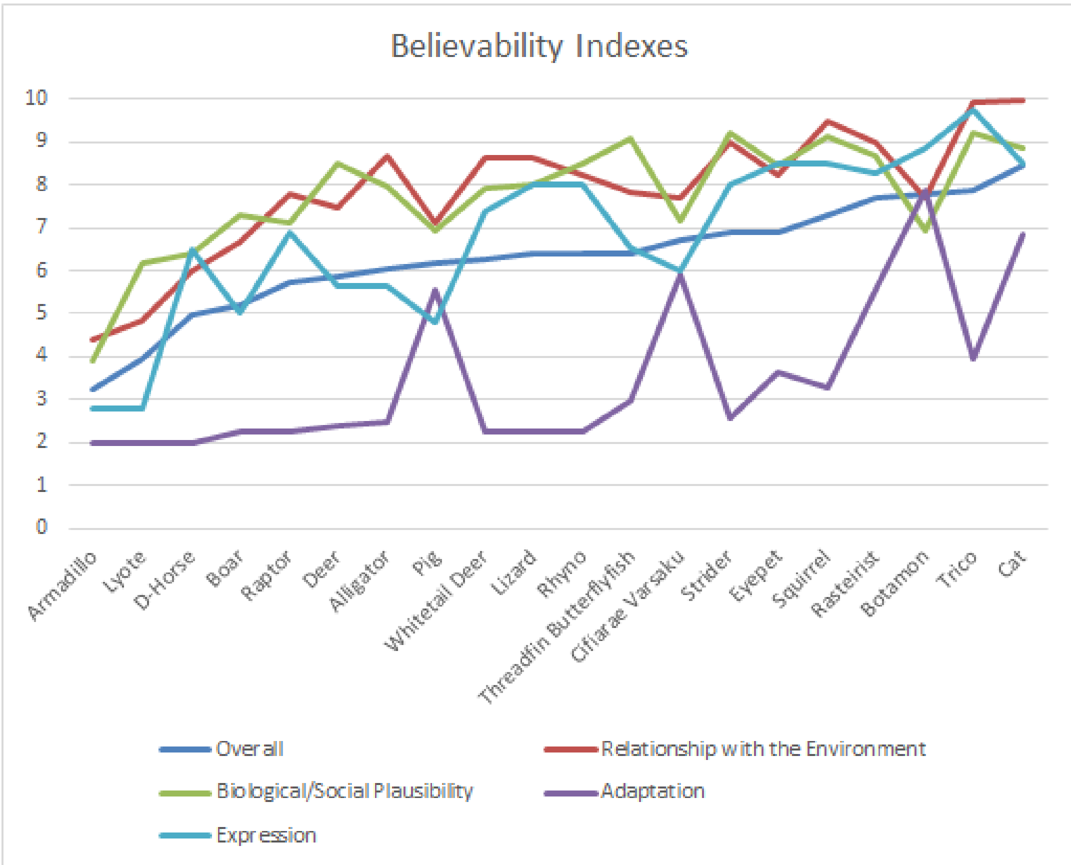


FIGURE 5.8: Believability Indexes for each of the videogames' creatures across each of the scales dimensions. Values are sorted in respect to their overall believability index.

As a final point, it is important to understand how the creatures' overall believability

TABLE 5.8: Perceived believability scores given, by 6 experts, to 20 videogame creatures. Additionally, two further columns are present: one showcasing the median calculated from these values and the other listing the values calculated from the median of the experts' CBS responses, rounded to the nearest integer.

Creature	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	Median	Calculated
Trico	10	10	9	8	10	10	10	8
Eyepet	8	9	8	7	7	8	8	7
Strider	7	7	8	5	8	9	8	7
Armadillo	8	5	1	0	3	0	2	3
Threadfin Butterflyfish	10	9	8	6	10	7	9	6
Cifiarae Varsaku	4	6	3	2	1	5	4	7
Botamon	8	5	4	5	10	8	7	8
Cat	9	8	9	7	9	9	9	8
Squirrel	9	7	9	7	8	8	8	7
Raptor	9	6	5	3	7	4	6	6
Alligator	8	7	5	4	8	8	8	6
Boar	8	6	5	5	6	4	6	5
Pig	7	7	2	3	9	7	7	6
Lyote	7	5	2	5	7	2	5	4
Deer	7	7	6	5	7	3	7	6
Whitetail Deer	10	6	4	7	9	7	7	6
D-Horse	9	8	6	5	9	6	7	5
Rasteirist	1	7	4	6	10	8	7	8
Rhyno	4	8	8	7	8	7	8	6
Lizard	8	7	7	1	9	7	7	6

indexes correlate to the experts' subjective perception of believability. By calculating the median between the scores they gave, when asked to rate the creatures' believability (Table 5.8), we were able to compare these against the calculated values used throughout this analysis. Using a Spearman's Correlation calculation, both vectors have a moderate correlation of 0.54. Though the sample's size is too small to provide any generalized conclusions, the spectrum created with the data collected from this sample, appears to be similar to the median of those values.

All in all, this case study allowed to analyze creatures from a believability perspective. Analysis such as these may provide insights in state of the art reviews and help understand how the different approaches in both development, and design, may be perceived and assist in identifying potential pitfalls.

5.7 Discussion and Summary of the second DSR Iteration

After establishing believability as a possible construct for evaluating creatures, we proposed an evaluation tool, in the form of a scale. During its construction several elements were considered. These went beyond the ones present in Believability literature. After a first validation attempt, results proposed that, on one hand, the scale was considered internally consistent (items appeared to measure the same construct) and, on the other hand, from the original 46 statements, 26 seemed to have some correlation to what people thought to be beliveability. This expands upon the literature, by showing additional elements not originally considered in its studies. Nonetheless, we performed a second validation process which showed the scale's underlying theoretical model to be, at best, a reasonable fit. Given that this was a first attempt towards making this kind of scale, we argue that while this second result is not as positive as expected, it is nonetheless a step in that direction, as, due to its infancy, the scale lacks support in previous systematic studies.

Still, when using the scale to measure the believability of the same creatures, across different evaluators, their responses were shown to agree, in 65% of the cases. For the remainder cases, we were able to identify potential interpretation problems and iterate the scale, for an additional two revisions, to clarify its statements. We now have a tool that can be tested for usefulness and further refined. While we cannot argue this revision to be the scale's final form as the "development of a scale is an ongoing process that really never ends" (Spector, 1992), we deemed it to be stable enough for a more practical exercise, detailed in Chapter 6.

A complementary contribution resulting from this process was the Creature Believability Spectrum. While this illustrates how the scale can help assess creatures, in different videogame contexts, we still needed to verify how this fared in guiding a creation process.

Chapter 6

Orphibs Evolvapalooza

With an evaluation method defined, we had met the conditions to revisit our original effort in developing a generation system capable of producing a creature that exhibited coherence between interesting behavior and appearance. Following the work done in Orphibs II (see Chapter 4), we reiterated it, using the CBS as a guideline to design the generative system. This was meant as a proof-of-concept of the CBS underlying believability construct. In addition, we used this as an opportunity to perform, with a new sample, a reliability and an EFA (Hinkin, Tracey, and Enz, 1997). As such, we decided to design a videogame where assessing creatures was part of its gameplay. Specifically, the CBS was integrated as a game mechanic. This was because, in our previous tests, subjects had to evaluate individual creatures in several different games. However, in Orphibs Evolvapalooza, the number creatures they had to evaluate was superior, in order to increase data variability. So, by having the evaluation as part of the gameplay, players wouldn't have to break their game play sessions every time that had to rate a creature. To this purpose, we developed a game prototype: first, in paper form (Roque, 2010) and then, in digital form. Both these forms underwent usability tests. The next section details the prototype's original game design and the subsequent ones list the changes made, informed by the tests. Though, the final game design document, post-revisions, can be found in Appendix D.

6.1 Original Game Design

Following the design philosophy of Petz (PF Magic, 1995), also underlain in Orphibs II, Orphib Evolvapalooza is a cartoon life-simulation videogame, where players take on the role of an animal raiser. It is inspired by the comedic interactions (e.g. by using unusual objects, such as toilet paper, to pet the animal) made available in Macho Cat: Pat the Virtual Cat (Sia Ding Shen, 2017) and the pet raising game mechanics (such as being able to pet, feed and watch the animal grow) in Creatures (Millenium Interactive, 1996), the Petz (PF Magic, 1995) franchise and Sonic Adventure's (Sonic Team USA, 2003) Chaos Garden mini-game.

Aesthetically, the videogame is meant to invoke calm, and playful, feelings, so that players feel invited to interact, and play, with the creatures.

6.1.1 Game Progression

While the videogame's objective is for players to evaluate creatures of 5 different generations, players are communicated that their main objective is to raise a generation of the most believable creatures, according to the CBS. In order to raise them, players can feed, pet and spank them. Furthermore, they are rewarded for interacting with the creatures. This comes in the form of points that can be used to evaluate creatures (Evolution Points), create a batch of creatures (Egg Points) and to buy raffles and win additional items (Money).

Players are also communicated that, in order to create better creatures, they must evaluate them, causing the next generation to evolve from these creatures. However, internally, each creature is predetermined and does not depend on the players' choices. Players are given the illusion of evolution as the generated they look, sound or behave differently even though they share common traits with the previous generations.

The videogame's progression flowchart diagram is depicted in Figure 6.1.

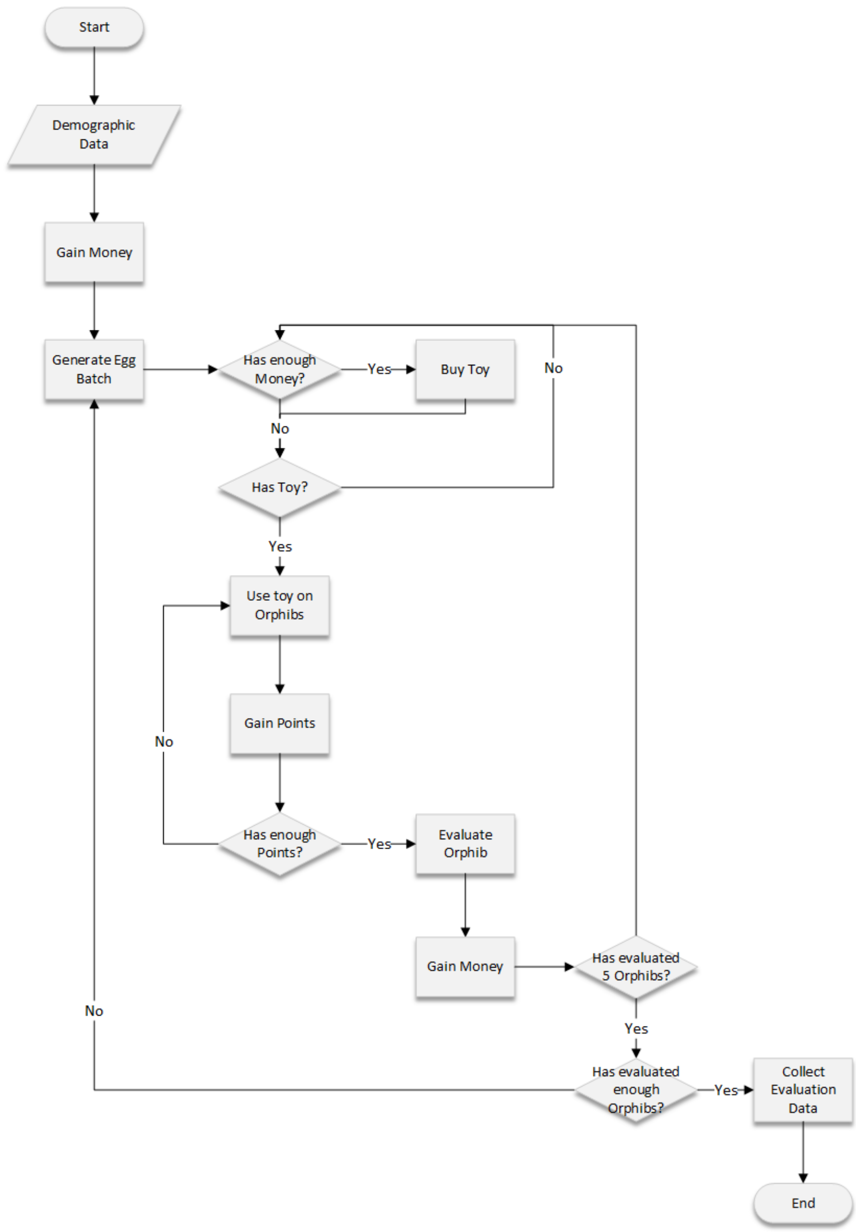


FIGURE 6.1: Orphibs Evolvapalozza Game's Progression Flowchart diagram.

The videogame starts by collecting the players' demographic data and then greets them with a small text explaining the purpose of the videogame/experiment and what is expected of them. Then, they are given money and the actual game loop starts: players are expected to generate a batch of eggs. At this time, the generated eggs hatch and they can either use items to interact with creatures, or buy new ones, if they have enough money. Every time an item is used on a creature, the player gains points. With enough of them, one of the creatures can be evaluated and money is awarded for this action. This is meant to drive the players' attention towards exploring interactions with the creatures, so that evaluation is based on exposure to them.

When this cycle is repeated 5 times, players are prompted to generate a new batch, repeating the game loop. However, once enough creatures are evaluated, the loop ends and the player's responses are collected, completing a gameplay level or iteration.

Game Mechanics

The videogame uses a top-down perspective, tilted at 45°, where players can scroll through the world. In addition, they can do four main actions:

1. They can interact with objects and creatures. This comes in the form of both neutral/friendly interactions such as grabbing items, pressing buttons or petting Orphibs and more aggressive ones (spanking Orphibs and tossing items).
2. They can buy items to play with the creature (e.g. stuffed toys, rolling pins, toilet paper, etc.). However, instead of directly choosing an item, players are awarded with a random one, from a pool of items. Unlocking tiers increases the number of available items. To buy items, players must have money.
3. They can play with creatures. When playing, players can use the items they bought, to pet creatures. As they do this, evaluation points are awarded. Moreover, the item they use, will break after many uses. They can also spank creatures when they dislike a given behavior.
4. They can evaluate creatures. Before being able to acquire a new batch of creatures, players must evaluate every creature in a given batch. Additionally, they must have enough points to do so. Evaluation is done using by rating the Orphib according to the CBS.

The videogame awards several types of points for these actions: evaluation points, money and egg points. The former are awarded when players pet creatures with an item. Each item awards a total of 50 Points: Afterwards it breaks and is rendered unusable. Money, on the other hand, is gained every time a creature is evaluated. This is done at a constant rate, increasing as the number of generated creatures increase. Furthermore, egg points are also awarded after an evaluation. These, unlike money, can be used to generate a new hatch of orphibs..

A final mechanic is related to the Orphib's actual evolution. While players are told they are choosing creatures to evolve, evolution is actually deterministic and does not depend on the players' choices. Although players are given the illusion of progress (Duarte and Carriço, 2013), this strategy is followed in the current prototype because the main goal is to collect creature evaluation data along the way.

6.1.2 Game World

The videogame is set in a small world so players can focus on interacting with Orphibs. It is a grass field, with some rocks and logs, surrounded by trees, who act as barriers. In the middle of it, lies a water fountain. Standing near that, is an egg dispenser machine. In the world's top-right corner stands a small cave formation with a mound of dirt nearby. In addition, there is a bird creature roaming around the world. This is illustrated in Figure 6.2.

FIGURE 6.2: A screenshot showing the Orphibs Evolvapalooza world. The world objects are displayed in the center of the image. From top to bottom: the cave, the egg dispenser, the fountain and the bird.

Objects

The videogame's world has several man-made objects, the player can interact with. The first two are part of the world: the fountain and the egg dispenser. When the player interacts with the former, the fountain starts producing water, Orphibs can then drink while the latter creates a new batch of eggs. In addition, there are several objects, that act as toys, the Orphibs can play with. Furthermore, there are mushrooms, the Orphibs are able to eat, and these grow in random spots along the videogame's world. These mushrooms, similarly to the aforementioned toys, can also be interacted by the player. Furthermore, there is one additional item: the clipboard. This special item can only be interacted with by the player and it allows he/she to evaluate a given Orphib.

Orphibs

The world is populated by Orphibs, herbivorous sexuated alien-like creatures. The player can play with 5 at a time. Their behavior is similar to that of most mammals, though they lay eggs (which only monotreme mammals do).

Guided by the CBS, Orphibs can autonomously (item 8 - "The creatures are able to move autonomously") perform the following actions:

- **Eat** - Orphibs walk up to a mushroom and eats it. This was taken from items 1 ("The creatures interact with surrounding environment") and 4 ("The creatures are able to find objects in the environment").

- **Play** - Orphibs can either fetch toys for the player or, follow him/her if he/she is grabbing a toy. This was based on items 1 ("The creatures interact with surrounding environment") and 3 ("The creatures direct their movement towards/away from targets").
- **Hunt** - Orphibs can stalk the bird present in the videogame's world and, when close enough, bite it. This was inspired by items 1 ("The creatures interact with surrounding environment") and 3 ("The creatures direct their movement towards/away from targets").
- **Excrete** - Orphibs ceremoniously expel organic materials and bury them. This was based on items 1 ("The creatures interact with surrounding environment"), 5 ("The creatures are able to expel organic waste") and 6 ("There are signs of previous expelled organic waste, such as feces, urine, goo, etc.").
- **Reproduce** - Unlike the other actions, this is not directly seen but rather, hinted at. When Orphibs want to reproduce, they move towards the cave and it closes. During this time, hearts are shown around the cave. After reproducing, one of the orphibs becomes pregnant and, eventually, lays an egg. These were based on items 13 ("The creatures engage in reproductive acts") and 14 ("There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.").
- **Learn** - There are three moments where Orphics can learn. First, whenever a player pets/spanks an orphib after an action, this action is reinforced/inhibited respectively. Second, when Orphibs observe the player performing some actions (interacting with the fountain), they learn it and can do it themselves. Third, Orphibs can learn to use the same actions in different contexts: after being able to fetch toys, Orphibs learn to hunt using the same basics. This was inspired by items 15 ("The creatures learn from past events"), 16 ("The creatures use past experience to overcome novel situations") and 22 ("The creatures can imitate others to develop their own behaviors").
- **Drink** - If the fountain has water on it, Orphibs can move towards it and drink its water. The reason behind this, was to account for item 10 ("The creatures react to environmental triggers").
- **Interact with the Fountain** - Orphibs can interact with the fountain, to make it pour water, akin to the player, for themselves, or other Orphibs. This was based on items 1 ("The creatures interact with surrounding environment") and 11 ("The creatures can perform tasks by coordinating with other creatures, or humans").
- **Sleep** - Orphibs move towards the mound of dirt and sleep on it. This was also to account for item 4 ("The creatures are able to find objects in the environment").
- **Talk** - Orphibs move towards other Orphibs and bark at each other. This was so that "The creatures audio-visually communicate with other creatures, or humans" (item 12).
- **Grow** - As time passes, Orphibs grow in size and change the color of their bodies, toward a greyish tint, to indicate they are maturing. Additionally, their voices' pitch lowers and they predisposition for certain actions reduces. This was based on items 17 ("The creatures change the way they look with age"), 18

("The creatures change the way they sound with age") and 19 ("The creatures change the way they behave with age").

- **Express emotions** - Orphibs have a specific body language to indicate emotions. First, they are able to pulse a glowing light to indicate they are expressing emotions. The light's colors are warm upon positive emotions whilst cold upon negative ones. Additionally, Orphibs have also special poses for this extent. Additionally, their ears and tails act differently. These elements was based on item 21 ("The creature shows signs of emotional reaction to objects, events and other creatures").

Visually, we decided to design Orphibs as fictional creatures though, with some familiar elements. This still falls within our definition ("Zoomorphic entities, inspired by (contemporary or otherwise) living or fictional beings (...)". The main reason behind this, is to allow players to derive some meaning from the familiar aspects while, at the same time, make the creature different enough from any existing species, as to avoid any expectations associated to them. As a result, Orphibs can be seen as some form of hybrid between a kangaroo and a dinosaur, as depicted in Figure 6.3. In addition, these animal inspirations, and the orphibs' color palette, were also selected to account for item 23 ("The creatures' bodies seem adapted to their habitat"), due to the game's forest environment.

FIGURE 6.3: A rendering of an Orphib.

6.1.3 Orphibs Evolvapalooza User Interface

There are two types of graphical user interfaces in Orphibs Evolvapalooza. The first one, used in the videogame's main screen, consists in superimposed navigation buttons. The other, used every other context, encompasses pop-up windows. The videogame's original Navigation Diagram is illustrated in Figure 6.4.

After going through the main menu, players are prompted to enter demographic data. Afterwards, it transits to the main interface. As previously mentioned, this comprises superimposed buttons over its main screen. These include three buttons on the lower right corner, and one on the upper right corner: one for the clipboard, one for the inventory, the other for the item store and one for the egg machine. Pressing them will recall the clipboard to the player's position, open the inventory, the item store and teleport the player to the egg machine, respectively.

When players drop the clipboard on Orphibs, a popup message appears stating the scale's item and a star mapping to a 5-point Likert rank: 1 (Not at all), 2 (Slightly), 3 (Somewhat), 4 (Moderately) and 5 (To a great extent). This pop-up window disappears after pressing the exit button.

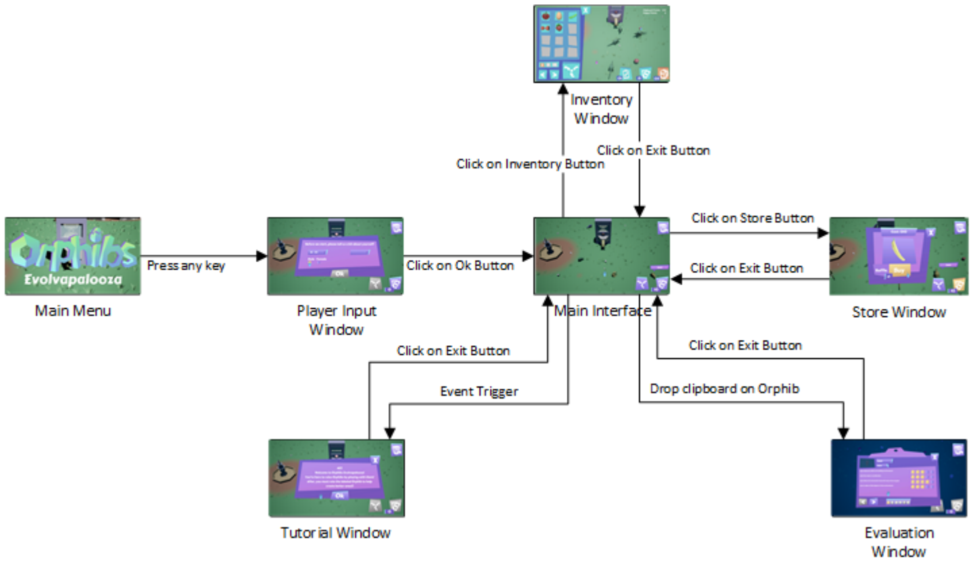


FIGURE 6.4: Orphibs Evolvapalooza's initial navigation diagram.

Regarding the inventory, this comprises a grid, illustrating the items a player has, and their respective amount. The player can click on each to make it spawn. There is also an exit button, on the upper right corner of the menu and a recall button to force spawned items to teleport to where the player is on the screen.

Similar to the other windows, this one is superimposed upon the playing-phase screen. It has a circular slideshow showing a subset of all of the available items, a text with the cost of buying the next raffle, how many tries the player has with his current currency and a button stating "Buy" (which is disabled if the player cannot buy a raffle). Additionally, an exit button is present on the upper right corner of the menu. When pressing the buy button, the slide show begins to stop until it completely slows down. At this point, the player receives the item present on the image frame.

The final type of window the player can navigate to, is the tutorial window. Unlike the other windows however, this one pops-up during special events, or after particular triggers. This is used to explain some aspects of the videogame. Every time one pops up, the videogame's logic is paused and the player must acknowledge, with the press of a button, to resume normal play. These tutorials teach the following:

- What to do with creatures
- How to evaluate
- How to win evaluation points
- How to win money
- How to win egg points
- How to buy items
- How to get new creatures

6.2 Implementation

After designing the videogame, we proceeded to its implementation. During this stage, it first underwent a Paper Prototype phase to evaluate the game's interface and, to some extent, its flow. This informed our design and, following a revision, the game was actually implemented as a software piece. The digital prototype also underwent a usability test and the resulting revision was deployed.

6.2.1 Paper Prototype

The paper prototype recreates some of the key aspects present in the videogame's design: its game flow and interface navigation. Particularly, this simulates one iteration of its loop: this takes place after players have generated an initial batch of Orphibs. Here, players can buy items, interact with the Orphibs and evaluate them. The materials used to create this prototype are depicted in Figure 6.5.

FIGURE 6.5: Paper Prototype of Orphibs Evolvapalooza. It lists window pop-ups for the item store, Orphibs' evaluation and inventory. The cutouts represent items and Orphibs.

As shown, both the items and the orphibs were represented as cut-outs (it is worth noting the, during the time of this test, the Orphibs did not have a finalized design so, they were represented as an assortment of animals and other creatures). These cutouts are placed in a board with a frame, with trees drawn, to represent the videogame world. There are two buttons, in this board, to mimic the inventory and store buttons. Due to the board's small size, we decided to omit the clipboard recall button.

To illustrate the points a player can have, we used three rotating dials, as depicted in the Figure's top left border. The remainder interface elements are also cutouts representing the inventory, the store and the evaluation windows. For the store window, there is a dice to simulate the store's randomness. To make the evaluation window more engaging, as studies suggest likert scales fail to captivate test takers (Leutner et al., 2017), it has text boxes, one per each item, succeeded by either star ranks, or a slider.

Paper Prototype Usability Test

The main objective behind this test was to refine the videogame's user interface and understand if its main loop was being correctly communicated to the players. The test's setup was made in a lab environment, and participants' actions were observed during the test. To this extent, we recruited a sample of 5 end users (Virzi, 1992).

The test's methodology is as follows:

1. Participants were explained the context behind the test.
   (a) They were also explained to use a finger to simulate a cursor and state which mouse button were using
   (b) A researcher would act as a mechanical turk to simulate the videogame's responses to player input by showing/hiding its windows, updating points or deploying items to the board.
2. The board was filled with the creature cut-outs
3. Each participant then underwent one iteration of the videogame's main loop:
   (a) An initial text box was shown to the participants explaining the videogame's objectives
   (b) They were then prompted to buy items and interact with the videogame's creatures
   (c) When enough evaluation points were gathered, participants were prompted to evaluate Orphibs
      i. Before this evaluation took place, participants were asked to select either star ranks, or sliders, as the evaluation's interface's interactive element.
4. The test ended when participants were able to evaluate every creature in the board
5. In the end, participants were asked to state any difficulties, or suggestions, they might have

The observed actions/suggestions are listed in Table 6.1. From the problems discovered, it was suggested that some of the game mechanics were not being correctly communicated to players so, we decided to increase the amount of tutorials messages to explain the context behind the videogame's possible interactions and each of the points' objectives. Regarding the interface itself, some adjustments were made in their final design to make the items' inventory more readable (e.g. changing the items' labels' locations) and to include consistent exit buttons. Finally, regarding the evaluation's interactive elements' design while most participants preferred the star ranks over the slider, the sample was not big enough for a statistical significant choice so, ultimately, we decided to go with the star rank design as it was more in line with the videogame's playful intent.

6.2.2 Orphibs Evolvapalooza Prototype v0.1

An initial software prototype was developed, following the paper prototype's usability test. It was developed in Unreal Engine 4.19 (Epic Games, 2012). The prototype is illustrated with a screenshot in Figure 6.6.

TABLE 6.1: Observations and Suggestions collected during the Paper Prototype's usability test.

Subject	Observation / Suggestion
1	The items' quantity labels, in the inventory window, were not readable due to their distance to the point icons Did not understand how to close some windows, due to a lack of consistent exit icons Did not understand the usage of each of the points
2	Did not understand how to interact with Orphibs Did not understand what points were for Preferred star ranks, instead of a slider, for the evaluation
3	Did not understand the difference between evaluation points and money Did not understand what points were for Did not understand how to evaluate Preferred slider, instead of a star ranks, for the evaluation
4	Did not understand what points were for Preferred star ranks, instead of a slider, for the evaluation
5	Did not understand between evaluation points and money Did not understand how to interact with Orphibs Preferred star ranks, instead of a slider, for the evaluation

FIGURE 6.6: A screenshot depicting Version 0.1 of the prototype's User Interface.

In this prototype, Orphibs were implemented as autonomous agents, while this portion of the software reused code from the Savanna Simulation (See Chapter 4), the overall agent architecture was redesigned to accommodate for the videogame's design, guided by the CBS, as described next.

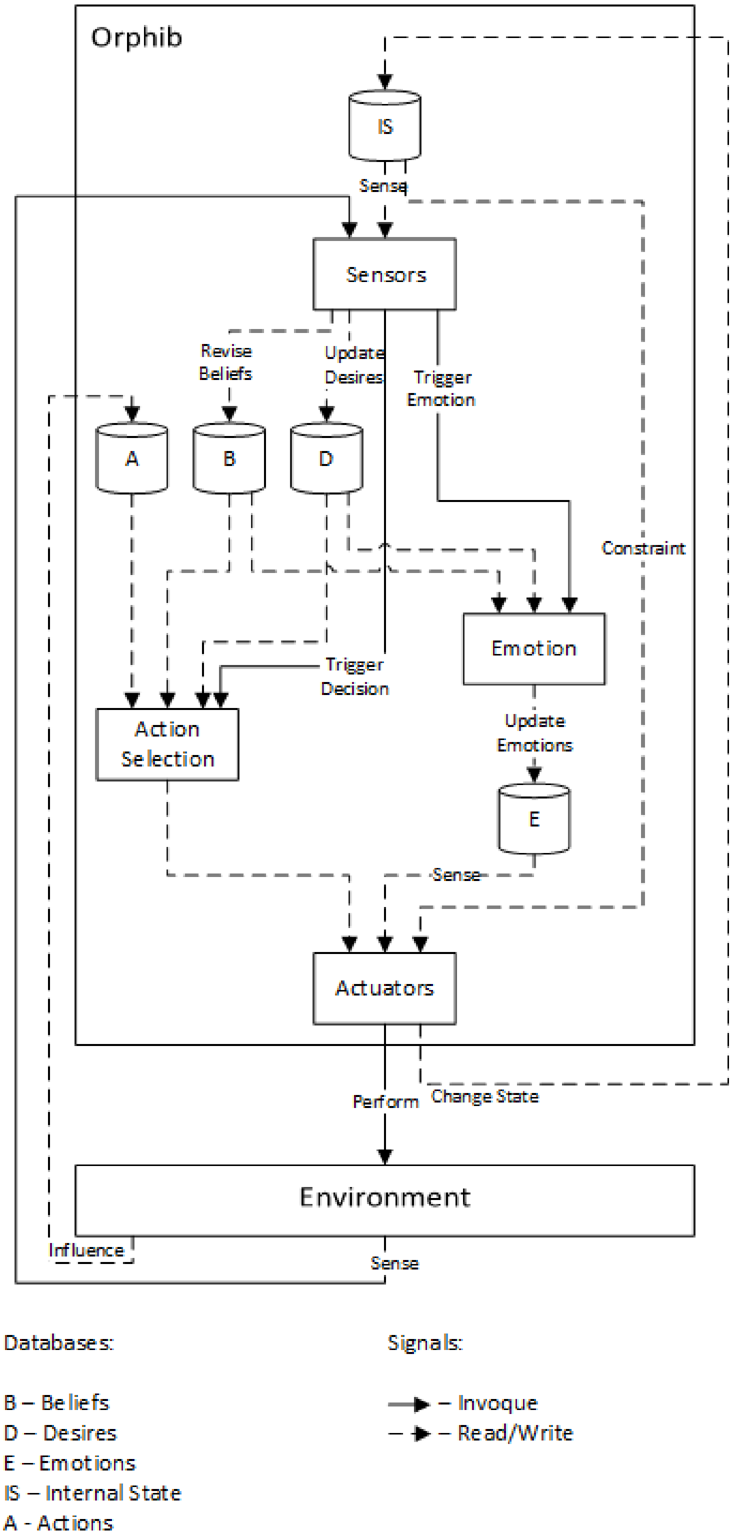


FIGURE 6.7: Orphibs Agent Architecture. Components are depicted as rectangles, while data bases as cylinders. Method Invocations and Read/Write operations are depicted as straight, and dashed, arrows respectively.

Agent Architecture

To develop the agent's architecture, we decided to devise a simple one, capable of supporting the videogame's design elements. As Orphibs were designed to perform simple actions, react to environmental stimuli (including the player), express emotions and, to some extent, learn, the proposed agent architecture is depicted in Figure 6.7.

As illustrated, the architecture has 5 databases:

- **Internal State** - The Orphibs' physiological needs, and other variables. To accommodate for its actions and mechanics, we considered the following: pain levels, hunger, thirst, energy, bowel capacity and the needs for playing, socializing and reproducing. These elements are stored as floating point values associated to an ID.
- **Beliefs** - This accounts for the Orphibs' mental model regarding the world. It includes its objects, other Orphibs and the player. This mental model is updated by the Orphibs' sensors. Similarly to the previous database, this one stores elements as floating point values associated to an ID.
- **Desires** - Desires derive from the Orphibs' Internal State. Specificaly, these convert the latter into a normalized value ranging from -1 to 1, representing the amount of inclination an Orphib has to fulfill/neglect a given desire. The elements in this database are likewise stored to the previous two.
- **Emotions** - The Emotions database contain the Orphibs' possible emotions. This stores emotion comparators (Reisenzein, 2012b), associated to an ID.
- **Actions** - This contains the Orphibs' action repertoire and their bias. The latter is used to influence the Action Selection component to prefer some actions over others, possibly more rewarding, ones. As such, these biases are used in the Orphibs' learning mechanism, detailed in the Action Selection component. In addition, this database also contains both reactive and fallback actions, used in the aforementioned component. This stores a behavior tree, and a floating point value, associated to an ID.

As for the components, there are 4:

- **Sensors** - Perceives both the internal, and external, factors to the Orphibs. The latter involves polling the Orphibs' internal states' database while the former polls the game's world for objects (this includes any item, Orphib, or the player him/herself). This information is filtered to simulation a vision sense. As such, only objects whose position have a dot product between 0.3 and 1 are considered. The inclusion of the Sensors component allowed supporting the CBS items 1 ("The creatures interact with surrounding environment"), 3 ("The creatures direct their movement towards/away from targets"), 4 ("The creatures are able to find objects in the environment"), 11 ("The creatures can perform tasks by coordinating with other creatures, or humans") and 22 ("The creatures can imitate others to develop their own behaviors").
- **Emotion** - This component, deriving from the one in Orphibs II (see Chapter 4), processes and outputs emotions, to the emotion database. While it uses the mechanisms present in the Computational Belief-Desire Theory of Emotion (Reisenzein, 2012b), we only considered two of its emotions: Happiness and Unhappiness. Moreover, emotion values are updated constantly so they tend to 0 as time elapses. The inclusion of this component was so the architecture

supported item 21 ("The creature shows signs of emotional reaction to objects, events and other creatures").

- **Action Selection** - The action selection component also derives directly from the one used in Orphibs II (see Chapter 4) as it already supported item 7 ("The creatures' actions are appropriate to their circumstances"). Internally, this uses a Goal-Based Behavior approach (Zubek, 2010):
  - Each Orphib has a main Behavior Tree (illustrated in Figure 6.8), used to select which action to perform next: Orphibs can perform 3 types of actions: Reactions (only performed after external triggers such as being hit), Actions (the Orphibs' main action repertoire) and Fallback Actions (actions, such as staying idle or wandering, performed when the selection algorithm fails). Reactions have precedence over Actions which, in turn, have precedence over Fallback Actions. The former was included by taking into account the CBS item 10 ("The creatures react to environmental triggers").
  - Whenever a reaction is not being performed, it uses a Goal-Based Behavior algorithm:

```
actionToConsider = null;
bestScore = 0;
for object in beliefs:
    for action in object.actions:
        score = GetScore(object, action)
        if(score > bestScore)
            bestScore = score;
            actionToConsider = action
if(actionToConsider)
    PerformAction(actionToConsider)
else
    index = random(0, actions.fallback.Count)
    PerformAction(actions.fallback[index])
```

  - The previous algorithm selects the highest scoring (action, object) tuples. When scoring an object, the 'GetScore' method uses the following formula: $\rho \sum_{x \in \{internalvariables\}} \frac{objectScore^2}{1+t} Desire(x)(\frac{1}{x} - \frac{1}{x+\Delta x})$. The ρ coefficient represents the Orphibs' action reinforcement. It increases/decreases as players pet/spank an Orphib after an action was performed, respectively. The expression $\frac{objectScore^2}{1+t}$ is obtained from the Orphibs' sensors. The actual $objectScore$ variable is calculated from the object's distance to the Orphib and t represents the time the object was last perceived. The $Desire(x)$ function returns the desire value the Orphib has to fulfill a given internal variable x. The remainder expressions are taken from the goal-based behavior's description in (Zubek, 2010). 15. The creatures learn from past events 16. The creatures use past experience to overcome novel situations
  - Similarly to the main controller, the actions themselves are coded as sub-behavior trees composed of several tasks.
- **Actuators** - The actuators component comprise a group of processes responsible for animating, moving Orphibs and environmental exchanges. It reads an action from a sequence of smaller, atomic, activities, stored in the Actions database

and orchestrates these resulting in a performance. This component is fed by both the internal state and the emotions database.

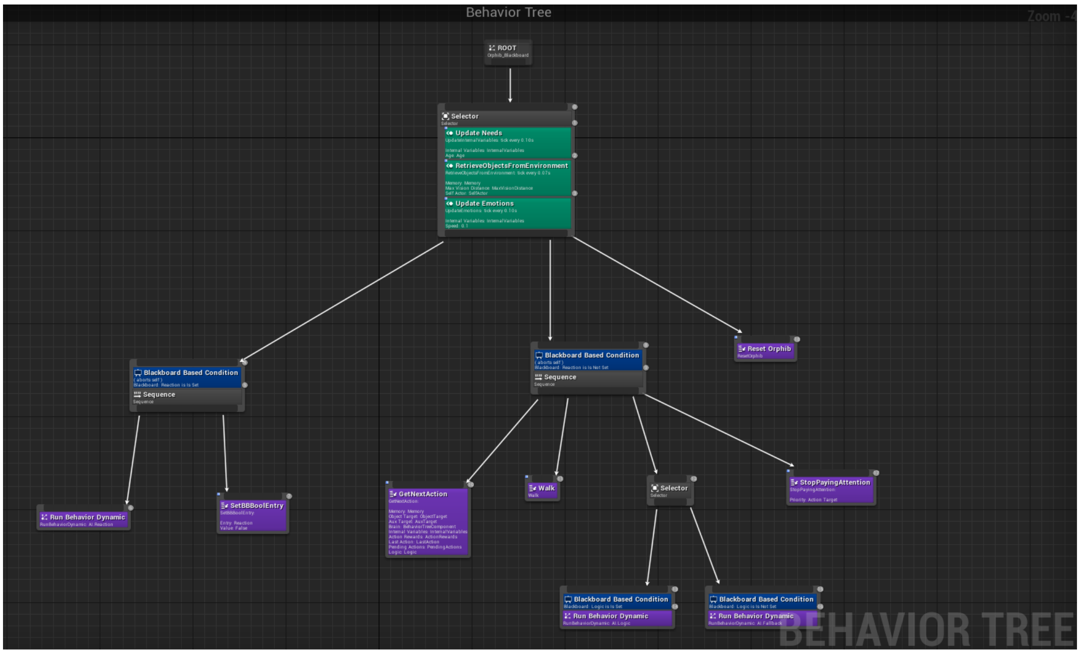


FIGURE 6.8: The Orphibs' main controller behavior tree. The root's selector node has 3 services to update the Orphibs' internal state, emotions and gather objects from its sensors. The two leftmost subtrees compute the algorithms for performing reactions, actions and fallback actions. These actions are inject into the 'Run Behavior Dynamic' nodes so, they get executed as a regular behavior three when those nodes execute.

Mapping Behaviors to Representation

One of the key findings in Chapter 4 was the concept of performance (the audio-visual output of a given agent, or group of agents, over time). Indeed several of the scale's items also hint at this connection.

The way we designed our agent architecture to support these items was through the actuators. As mentioned, some of them are responsible for sequencing animations. Internally, these rely on animation state machines, *blended animations* and *morph targets*. On one hand, the former consists in a directed graph structure where states correspond to animations. In our architecture, transitions are linked to Internal State variables. On the other hand, the latter two techniques are layered on top of the state machine, as illustrated in Figure 6.9.

The main difference between the two is that the first one, *blended animations*, encompasses skeletal-based animations whilst *morph targets*, a type of vertex-animation, correspond to a snapshot pose. In both these techniques, the animations are interpolated to a base one. In our architecture, the internal state and the emotions database influence the interpolation weight.

To support item 21 ("The creature shows signs of emotional reaction to objects, events and other creatures"), emotions play an integral part in influencing the actions: each emotion is associated to a weight that blends in mood-related animations (e.g. tail

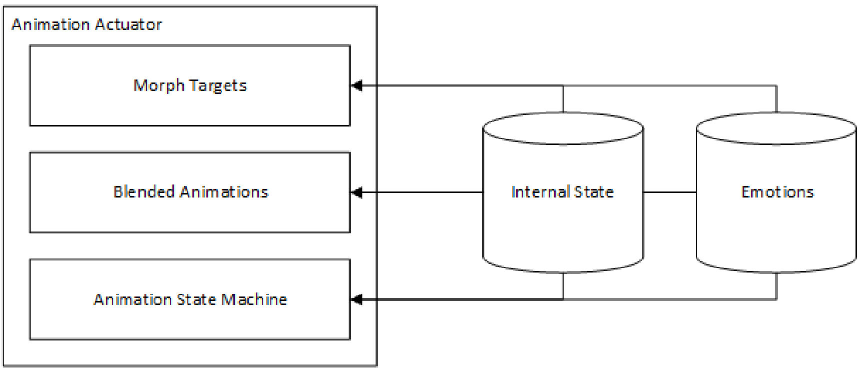


FIGURE 6.9: The Animation Actuator. It is organized hierarchically were the top layers are blended over the bottom ones. In addition, the Internal State and Emotions databases influence each of these layers.

wag). This way, it is given the illusion that Orphibs perform actions with certain moods.

To illustrate that orphibs have reproduced, as depicted in item 14 (“There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.”), we designed a *morph target* showcasing an inflated abdominal region. This is blended in proportion to the pregnancy's gestation period.

To show an orphib is growing, due to items 17 (“The creatures change the way they look with age”), 18 (“The creatures change the way they sound with age”) and 19 (“The creatures change the way they behave with age”), we devised a two types of *morph targets*. The first one, changes the creature's size. The second one, changes its facial features as well as its exposed dorsal spine's size. Both these targets are blended in proportion to the orphib's age.

Mapping the Orphibs to the Creature Believability Scale

As the game's goal is to evolve creatures so they can be evaluated, using the CBS, we had to understand what to evaluate. The Orphibs' main behavior design derived from guidelines we interpreted from the scale itself. To this extent, we developed a generative system capable of creating Orphibs with several characteristics mapped to the scale's items'. This exercise is latent in Table 6.2. It is worth stressing that these only account for one possible, proof of concept, mapping between the scale's items and gameplay mechanics as we are certain there are many others.

These mappings were converted into vectors, who were fed to the game's Egg Machine object. This, in turn would cause the object to generate different instances of the Orphib agent archetype, taking into account the parameter values from an input vector.

Overall, the idea behind a generative solution served as the early stages towards a potential automated creature design tool. This was integrated in the software prototype and after its development, it underwent a second Usability Test.

TABLE 6.2: Mappings between the CBS items and gameplay mechanics/Orphibs' agent architecture. This states the scale's items on the left column while the center and rightmost ones describe the parameter value ranges used (taking into account the scale's ranks) and their effect description, respectively. Some mappings are binary so, in those cases, only the extreme ranks are considered. These ranks are abbreviated and their meanings is as follows: NAL - Not At All; S - Slightly; SW - Somewhat; M - Moderately; TAGE - To a Great Extent.

Scale Item: Parameter	Parameter Value Range	Mapping: Effect Description
The creatures interact with surrounding environment	0% (NAL) 25% (S) 50% (SW) 75% (M) 100% (TAGE)	Reduce the speed with which the internal state is updated
The creatures have motor coordination	1 (NAL) 0.75 (S) 0.5 (SW) 0.25 (M) 0 (TAGE)	Blend ragdoll physics to the Orphibs' initial animations with a certain weight
The creatures direct their movement towards/away from targets	0 (NAL) 0.25 (S) 0.5 (SW) 0.75 (M) 1 (TAGE)	Overrides the action selection algorithm to only select directed behaviors with a random probability
The creatures are able to find objects in the environment	100% (NAL) 75% (S) 50% (SW) 75% (M) 0% (TAGE)	Reduce the Orphibs' sensors maximum range by a certain percentage
The creatures are able to expel organic waste	Remove (NAL) Add (TAGE)	Remove/Add the 'Excrete' action from the Orphibs' action database
There are signs of previous expelled organic waste, such as feces, urine, goo, etc.	Invisible (NAL) Visible (TAGE)	Turn organic waste objects as visible/invisible
The creatures' actions are appropriate to their circumstances	0 (NAL) 0.25 (S) 0.5 (SW) 0.75 (M) 1 (TAGE)	With a certain probability, the action selection algorithm is overridden by a purely random selection
The creatures are able to move autonomously	Disable (NAL) Enable (TAGE)	Enable/Disable the Orphibs' capability to move

The creatures' agility reflect their volume	5 (NAL) 4 (S) 3 (SW) 2 (M) 1 (TAGE)	With 5 sizes and 5 matching velocities, select a pair with a certain Hamming distance
The creatures react to environmental triggers	0 (NAL) 0.25 (S) 0.5 (SW) 0.75 (M) 1 (TAGE)	Reactions have a random probability to happen, when triggered
The creatures can perform tasks by coordinating with other creatures, or humans	Disable (NAL) Enable (TAGE)	Enable/Disable the Orphibs' ability to interact with the game's water fountain
The creatures audio-visually communicate with other creatures, or humans	Remove (NAL) Add (TAGE)	Remove/Add the 'Talk' action from the Orphibs' action database
The creatures engage in reproductive acts	Remove (NAL) Add (TAGE)	Remove/Add the 'Reproduce' action from the Orphibs' action database
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	Disable (NAL) Enable (TAGE)	Enable/Disable both egg objects and the Orphibs' growing mechanic
The creatures learn better responses for repeated events	0 (NAL) 0.25 (S) 0.5 (SW) 0.75 (M) 1 (TAGE)	The learning ρ coefficient is updated at a random probability
The creatures use past experience to respond to novel events	Disable (NAL) Enable (TAGE)	Enable/Disable the Orphibs' ability to hunt after learning its basics from performing the 'Play' action
The creatures change the way they look with age	0% (NAL) 25% (S) 50% (SW) 75% (M) 100% (TAGE)	Orphibs grow to a percentage of their maximum possible size
The creatures change the way they sound with age	0% (NAL) 25% (S) 50% (SW) 75% (M) 100% (TAGE)	When growing, the Orphibs' voice pitch change is reduced at a certain percentage
The creatures change the way they behave with age	Disable (NAL) Enable (TAGE)	Enable/Disable age-dependent actions

The creatures' body language anticipate their actions	Disable (NAL) Enable (TAGE)	Enable/Disable the Orphibs' animation's telegraphy
The creatures show signs of emotional reaction to objects, events and other creatures	Disable (NAL) Enable (TAGE)	Enable/Disable the Orphibs' agent architecture Emotion component
The creatures can imitate others to develop their own behaviors	Disable (NAL) Enable (TAGE)	Enable/Disable the Orphibs' ability to interact with the water fountain, after observing the player doing so
The creatures' bodies seem adapted to their habitat	Contrasts (NAL) Blends (TAGE)	The Orphibs' main skin color blend/contrast with the environment's

Videogame Usability Test

Unlike the previous test, this one was meant to focus on the game flow and assess its pacing. Additionally, we wanted to reduce any kind of user fatigue that could rise from the evaluation portion of the videogame.

Similarly to the previous test, this one's setup was also made in a lab environment where participants' actions were observed. For this test, we recruited an additional sample of 5 end users (Virzi, 1992), not yet exposed to previous prototyping or test stages. Its methodology is as follows:

1. Participants were explained the context behind the test.
2. Each participant played the prototype from 30 minutes to 1 hour. During this time, the prototype's tutorial was the only explanation regarding the game's possible interactions and mechanics. In addition, notes were taken of the ways subjects interacted with the prototype.
3. In the end, participants were asked to state any difficulties, or suggestions, they might have.

Both the observations and suggestions noted, during the test, pointed to the same issue: The prototype's pacing was too slow and it was causing user fatigue. Because participants had to evaluate 5 orphibs per batch, and several batches per session, it became evident this was impacting the players' experience. Additionally, subjects noted the game loop forced several activities before allowing evaluations. As a result, both the game loop and its design were streamlined to facilitate evaluations. The first main change was in the videogame's user interface's navigation, as illustrated in Figure 6.10.

To iterate upon the user interface, we removed the inventory altogether and, instead, bought items were directly deployed onto the level. Furthermore, the items' breaking mechanic was removed. As an alternative, a button was added to the main interface, allowing to recall every item in the world.

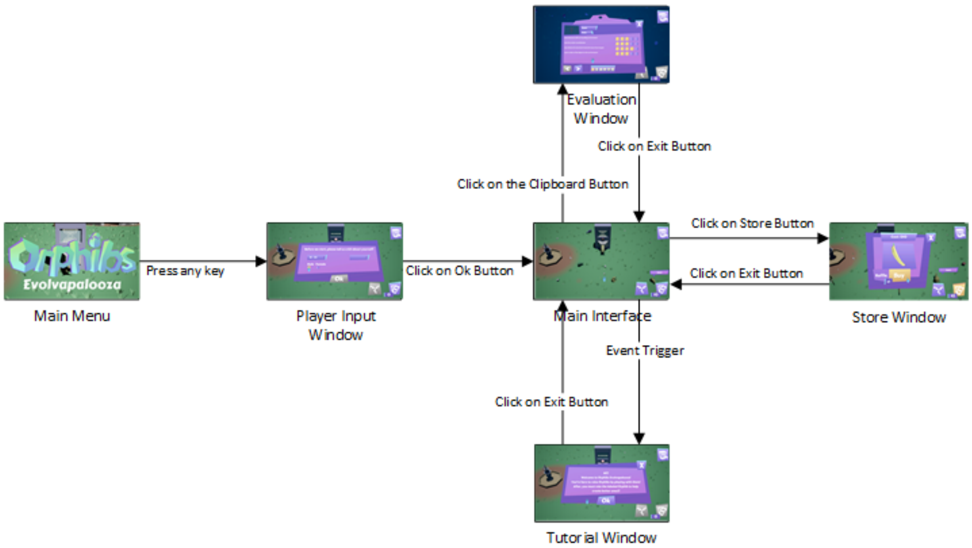


FIGURE 6.10: Orphibs Evolvapalozza Navigation Diagram.

The game flow also underwent some changes. With the goal to simplify the path between playing with orphibs and evaluating them, some of videogame's progression subloops were cut, as depicted in Figure 6.11. As can be observed, the videogame's main loop was reduced: when players generate a batch of eggs, they are awarded money. At this time, they can either use items to interact with creatures, or buy new ones, if they have enough money. Now, instead of evaluating every orphib in the batch, they only evaluate 1, indicated by a name label. orphibs can be evaluated when the players interact, at least, 7 times with them, and they appear on screen for 2 minutes. This is visually indicated whenever a player interacts with a labelled orphib. After evaluating it, players are prompted to generate a new batch, repeating the game loop. However, the game loop ends when they manage to evaluate 5 orphibs. Afterwards, the players' responses are collected.

Even with the amount of evaluated orphibs, per batch, reduced to 1, we decided to keep the original amount of 5 creatures per batch. This way, the remainder 4 orphibs were used as comparison and to allow social actions.

With the videogame's new progression flowchart and navigation model, we reiterated the prototype. To accommodate for these new changes, the aforementioned generative system, was fed vectors to create several orphib instances. Particularly, 5 different vectors were created, at coding time. These were randomly built, under the assumption that the items in the scale were orthogonal, so that each vectors' element were of different values (taking into account the items' mapping parameters' maximum domain). This allowed us to test all parameter values without extending the amount of creatures to evaluate and, consequently, increase user fatigue.

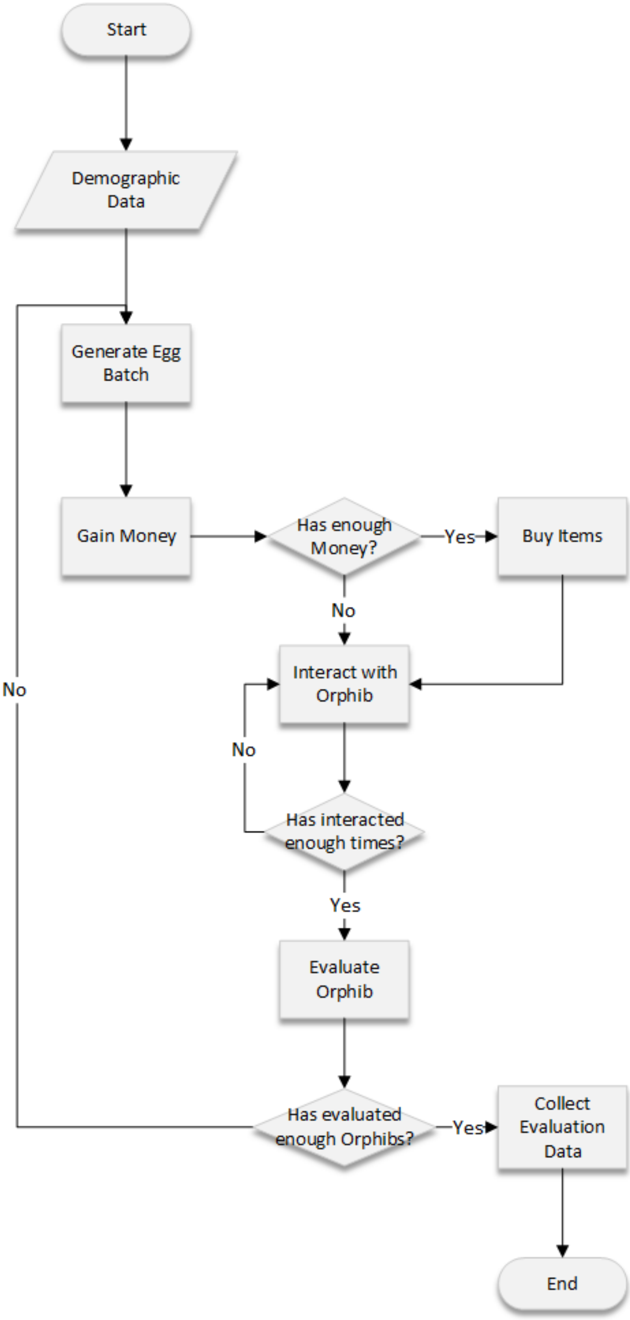


FIGURE 6.11: Orphibs Evolvapalozza revised Game's Progression Flowchart diagram.

The algorithm used in the creation of these vectors consisted in a random shuffling of each items' possible values, across the 5 vectors. However, in elements where these had only 2 possible values, the latter had to be repeated. The resulting vectors are as follows:

[1 0.5 0.5 0.25 0 1 0.5 0 0.5 1 1 1 1 0 0.25 1 0.5 0.25 0 0 0 1 0]

[0.75 0.75 1 1 1 1 0.75 1 0 0 0 0 0 1 0.75 0 0.25 0 0 1 1 0 1]

[0 1 0.75 0.75 0 0 0 0 0.25 0.25 0 0 0 0 0.5 0 0.75 0.5 1 0 0 0 0]

[0.25 0 0.25 0.5 1 0 0.25 0 1 0.5 1 1 1 1 1 1 1 0.75 0 0 0 0 0]

[0.5 0.25 0 0 0 0 1 1 0.75 0.75 0 0 0 0 0 0 0 1 1 1 1 1 1]

As a final note, each of the videogame's labelled orphib has created from one of these vectors. The remainder ones were composed of a random selection of the remaining vectors.

6.3 Deploying Orphibs Evolvapalooza

Having iterated the Orphibs Evolvapalooza prototype, informed by two usability tests, its third iteration, illustrated in Figure 6.12 was deployed to the public.

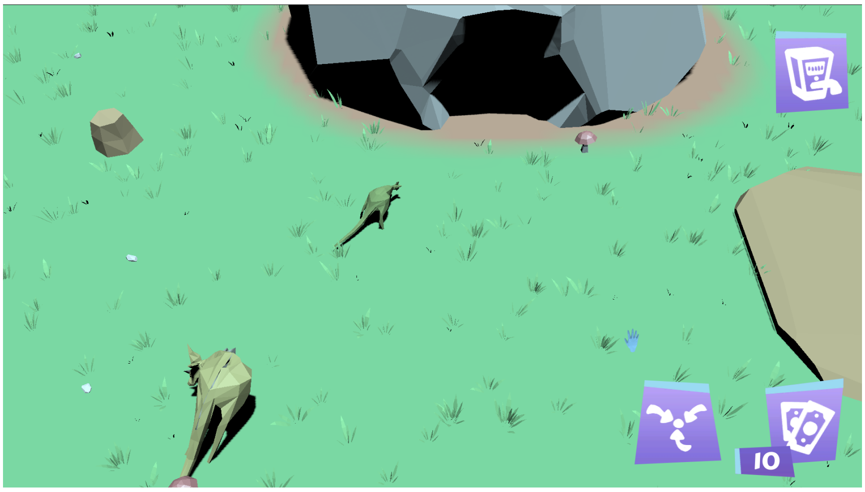


FIGURE 6.12: A screenshot of Orphibs Evolvapalooza.

During the videogame's deployment, data was collected to assess several goals. First, it served as an opportunity to perform a test replication (Hinkin, Tracey, and Enz, 1997), with a new sample. Finally, since players evaluated the same 5 orphib instances, their evaluation was further used to calculate if the players could agree on their assessment, through means of ICC.

To this extent, the experiment's methodology involved prompting participants to play the videogame to its completion. So, to sum up the game flow, the methodology is as follows:

1. After filling demographic data, participants were welcomed into the videogame, where they were explained its objectives, possible interactions and game mechanics.
2. During 5 generations participants had to:
   (a) Generate a batch of Orphibs' eggs.
   (b) Interact with, and observe, the labelled Orphib until its evaluation is unlocked.
   (c) Evaluate the Orphib using an in-game window, raking it according to each of the CBS statements.

3. After the fifth generation, the videogame informed the parcipants that the videgame had reached its completion.

6.3.1 Prototype Deployment Results

The videogame was deployed online, where 41 users participated (29 Males and 12 Females). From these 41 participants, 2 were aged between 18 and 20, 34 were aged between 21 and 29, 4 were aged in the 30 to 39 range while 1 was aged in the 40 to 49 range. Regarding media contact, 13 of these users had a weekly exposition of up to 20 hours, while 18 had 20 to 40 weekly hours, 5 had 40 to 60 weekly hours, 3 had 60 to 80 weekly hours, 1 had 80 to 100 weekly hours while another 1 had a weekly contact of over 100h. The sessions ran for an average of 32m22s, where each game loop iteration (from generating an egg batch up until an orphib is evaluated) took, on average, 5m17s. Finally, each labelled orphib, were interacted with, on average, 15 times, while the remainder ones were interacted with 5 times.

As mentioned previously, with the data collected, two analyzes were performed. Firstly, we performed a relability analysis and an EFA. This is explained in subsubsection 6.3.1. Secondly, we calculated an ICC, as depicted in subsubsection 6.3.1.

Exploratory Factor Analysis

An initial reliability analysis, using the Cronbach Alpha, showed a value of 0.94. While it is slightly below the initial 0.96 obtained in the first analysis (Referenced in Chapter 5), results still suggest it to have an excellent internal consistency (Steiner, 2003), albeit with some redundancy.

We then performed an EFA to confirm our scale's factors: PCA with a Varimax Rotation and a stop criteria of 4 fixed factors. The results are listed in Table 6.3. As can be observed, with these criteria, 4 components were extracted with, at least, a total cumulative variance above 50%, a threshold we previously established. As such, results seem to verify the scale's assumed number of dimensions.

TABLE 6.3: Extracted components from an EFA using PCA, with a Varimax Rotation with 4 fixed factors.

Component	Eigenvalue	% of Variance	Cumulative %
1	10.225	44.455	44.455
2	2.551	11.089	55.544
3	1.476	6.415	61.959
4	1.260	5.479	67.439

From looking at the item's loadings, listed in Table 6.4, it can be observed that with the same cut-threshold loading value of 0.4, every item loaded in, at least, one of the 4 extracted components, though items were arranged differently across the scale's dimensions. While this is not surprising, as the CFA hinted at a reasonable fit (Chapter 5), it is worth noting 15 items, out of 23, loaded in their expected dimensions.

As indicated, results point to the scale still being internal consistent, when used with a new sample, on a more interactive experimental setup.

TABLE 6.4: The Believability Scale second EFA. The first column depicts the original scale's items. The second column shows the loadings obtained during the EFA using a PCA, with 4 fixed factors, Varimax Rotation and a loading cut-threshold value of 0.4.

Item	Components			
	1	2	3	4
1. The creatures interact with surrounding environment	0.811			
2. The creatures have motor coordination	0.783			
3. The creatures direct their movement towards/away from targets	0.822			
4. The creatures are able to find objects in the environment	0.709			
5. The creatures are able to expel organic waste				0.821
6. There are signs of previous expelled organic waste, such as feces, urine, goo, etc.				0.790
7. The creatures' actions are appropriate to their circumstances	0.732			
8. The creatures are able to move independently	0.756			
9. The creatures' agility reflect their volume	0.695			
10. The creatures react to environmental triggers	0.735			
11. The creatures can perform tasks by coordinating with other creatures, or humans	0.485		0.526	
12. The creatures audio-visually communicate with other creatures, or humans	0.424	0.443	0.406	
13. The creatures engage in reproductive acts			0.870	
14. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.			0.855	
15. The creatures learn from past events		0.632		
16. The creatures use past experience to overcome novel situations		0.692		
17. The creatures change the way they look with age		0.736		
18. The creatures change the way they sound with age		0.786		
19. The creatures change the way they behave with age		0.823		
20. The creatures can imitate others to develop their own behaviors		0.661		
21. The creatures' body language anticipate their actions		0.556		
22. The creature shows signs of emotional reaction to objects, events and other creatures		0.558	0.500	
23. The creatures' bodies seem adapted to their habitat	0.425	0.550		

Observer Agreement Analysis

In this second analysis, the same data was used to calculate the agreement among participants through an ICC. After establishing a theoretical baseline value of 0.6, identical to our previous study, with 5 observable objects, Bujang (2017) suggested a minimum number of 30 observations to obtain significant results. Our observers' sample size was 11 units above the minimum. These results are depicted in Table 6.5.

TABLE 6.5: ICC values calculated, for each Orphib, evaluated by 41 participants.

Orphibs	ICC
1	0.351
2	0.347
3	0.047
4	0.176
5	0.313

Results show that all ICC values are below 0.40. This accounts for a poor agreement between participants (Cicchetti, 1994). Because this contradicts our initial agreement test (refer to Chapter 5), a more thorough analysis was conducted to assess the reason behind the phenomenon.

By plotting, for each analyzed Orphib, the participants' scores against each of the scale's items, using a Box Plot graph as illustrated in Figures 6.13, 6.14, 6.15, 6.16, 6.17, it becomes apparent several factors:

1. In some of the Orphibs, some items, have a wide spread, with an IQR superior to 3. This might have contributed to the low ICC.
2. Conversely, Orphib 3 has a very small dispersion across most items. While this should have contributed to a large ICC value the value listed in Table 6.5 counters this. However, upon a closer inspection, this box plot contains several outliers.
3. The items' spread changes across Orphibs.

From these observations, we speculated 4 possible reasons behind this phenomena:

1. Several participants could have gotten different experiences due to their interactions with the artifact.
2. The mappings between the generative system and the scale's items were not perceived as intended, or at all.
3. Scale's items were being misinterpreted. After the revisions the scale underwent to improve the items' readibility, we hypothesized this to be a possible reason, though not as probable as the previous two.
4. The fact there were only 5 objects interfered substantially with the evaluation results, and did not provide adequate conditions for distinguishing between creatures; Expectation effects (creature evolution) might have created a bias.

As an attempt to assess the actual reason, a follow-up interview was conducted to some of the participants.

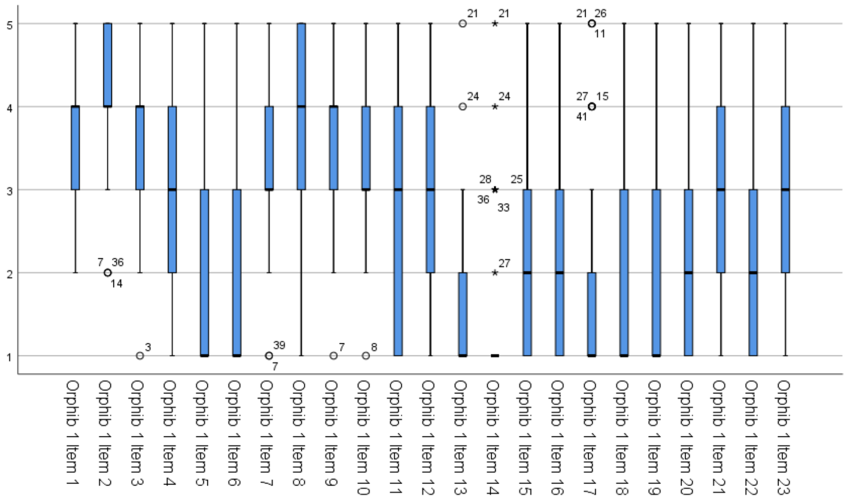


FIGURE 6.13: Orphib 1 boxplot (n = 41). Note that item 11 has an IQR of 3 and item 14 has several outliers and extreme outliers.

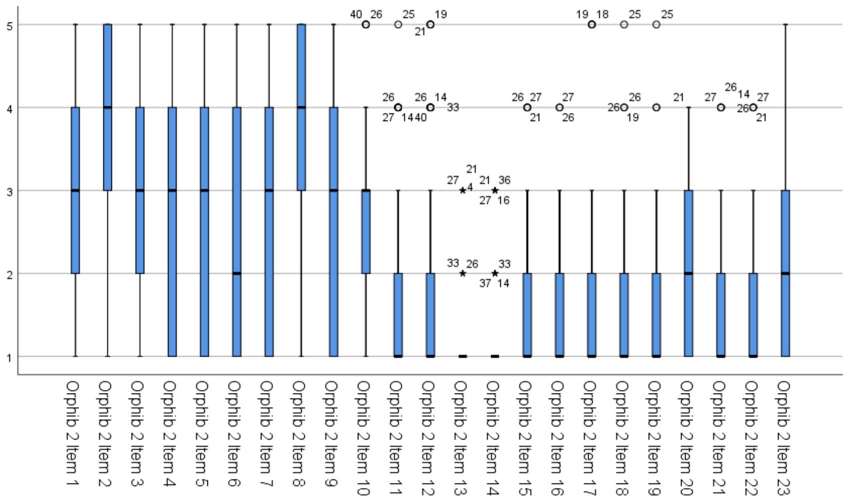


FIGURE 6.14: Orphib 2 boxplot (n = 41). Note that items 4, 5, 6, 7 and 9 have an IQR of 3.

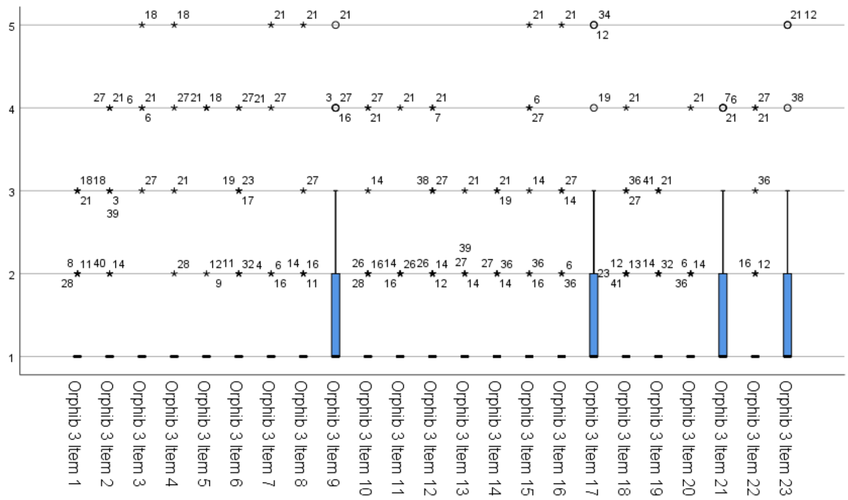


FIGURE 6.15: Orphib 3 boxplot (n = 41). Note that all items have several outliers and extreme outliers.

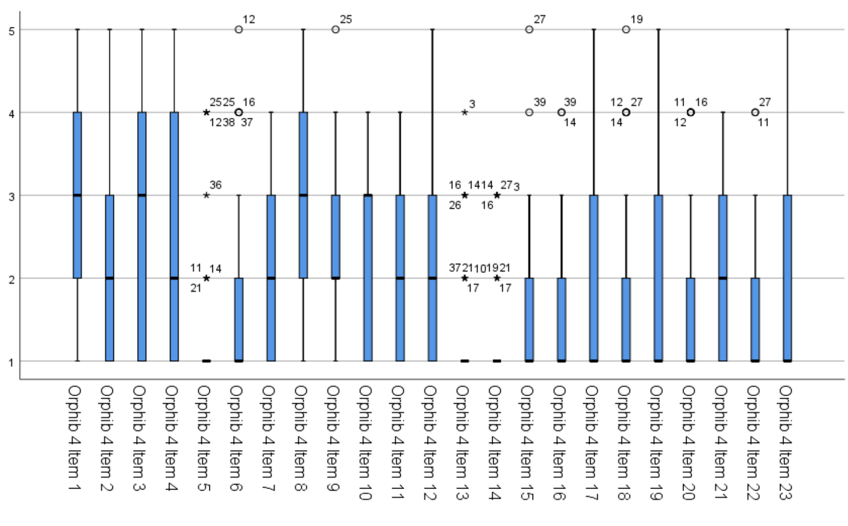


FIGURE 6.16: Orphib 4 boxplot (n = 41). Note that items 3 and 4 have an IQR of 3 and items 5, 13 and 14 have several outliers and extreme outliers.

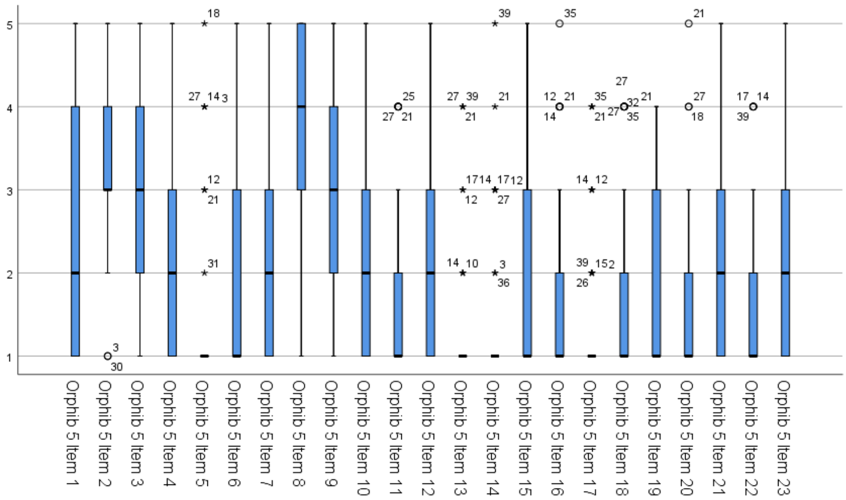


FIGURE 6.17: Orphib 5 boxplot (n = 41). Note that items 5, 13, 14 and 17 have outliers and extreme outliers.

Follow-up Interviews

The follow-up interviews focused on obtaining qualitative data that could allude towards the possible reasons behind the low ICC scores, across all Orphibs. We considered potentially problematic items, those who had an IQR of at least 3, more than 2 outliers, or had precedents in previous tests. As a result, several items passed these conditions. From these, we chose items 2, 3, 7, 10 through 13, 15 through 19, 22 and 23.

The followup interview we designed, was made with the goal to understand whether the score discrepancies came from different experiences, different perceptions of the Orphibs' instances or misinterpretations of the items. Each of its question was meant to gather some information regarding 1 item, or a group of semantically identical ones. These are listed in Table 6.6.

The interview was also preceded by a question prompting participants to describe their experience with Orphibs: Evolvapalooza in, at most, 10 lines. The interview's final form, as it was presented to the participants is detailed in Appendix A. It was administered online to 10 randomly selected participants, out of the original 41. The experience related question was analyzed through coding of specific qualifying keywords and narrative analysis whilst the remainder questions used solely narrative analysis.

When asked to detail their experience, respondents found it 'fun' (60%) and 'intuitive' (10%) but also 'a bit repetitive' (20%), 'too long' (20%), 'confusing' (30%), difficult (30%), 'a bit glitchy' (10%) and with a 'slow paced gameplay' (10%). Furthermore, one participant noted he "was met with this menu telling a lot of things (...)", he "wanted to skip all those instructions and explore away". Additionally, he initially though he was "(...) potentially making them smarter with each generation" but ultimately he "realized that" he "wasn't". The later sentiment was shared by an additional participant who stated she didn't "think they changed enough for the length of the experience". One subject referred the evaluation mechanic was interesting "since it lets the player judge things based on interactions and lets him/her repeat actions if in doubt". However, this same subject added that, given the items present in the

CBS, he potentially did not experience everything the videogame had to offer "since some events couldn't or wouldn't happen even after a long while". Similarly, another subject emphasized he "wasn't able to witness the reproduction or teach the orphibs".

TABLE 6.6: Orphibs Evolvapalooza follow-up interview's entries and their corresponding scale's items.

Question	**Items**
Please say whether you consider the Orphibs' body-movement as normal?	2. The creatures have motor coordination
Are Orphibs able to chase after toys, food, or any other thing?	3. The creatures direct their movement towards/away from targets
Please say whether you consider the Orphibs' behavior as normal?	7. The creatures' actions are appropriate to their circumstances
What did the Orphibs react to?	10. The creatures react to environmental triggers
How did Orphibs coordinate with each other?	11. The creatures can perform tasks by coordinating with other creatures, or humans
Can you please describe ways the Orphibs communicate?	12. The creatures audio-visually communicate with other creatures, or humans
Can you please describe the way Orphibs reproduce?	13. The creatures engage in reproductive acts
Did you teach Orphibs anything?	15. The creatures learn from past events
	16. The creatures use past experience to overcome novel situations
How did you distinguish between baby and adult Orphibs?	17. The creatures change the way they look with age
	18. The creatures change the way they sound with age
	19. The creatures change the way they behave with age
Can you please describe ways Orphibs express their emotions?	22. The creature shows signs of emotional reaction to objects, events and other creatures
Can you please describe what you believe should be the Orphibs' natural habitat?	23. The creatures' bodies seem adapted to their habitat

The answers collected provided some insights regarding potential justifications to explain the results' lack of agreement:

Item 2. The creatures have motor coordination - Although all participants stated, at first, the orphib's body-movement was normal, when asked to develop their answer they explained some orphibs had a 'normal' coordination while others shook their bodies violently, akin to 'ragdolls'. One respondent stated some orphibs changed their movement direction too drastically. Because all of the interviewees managed to successfully identify what constituted a 'normal body-movement' and managed to recognize possible deviations from this, we conjecture these discrepancies stem from the mapping between the item and its computational implications. Additionally, the lack of movement on one orphib might have contributed to the misperception of this item.

Item 3. The creatures direct their movement towards/away from targets - Most respondents were able to correctly identify situations where orphibs could chase after items. Particularly, 5/10 found some orphibs able to chase after either food or toys, though 2/10 mentioned orphibs were able to chase items without detailing which. Similarly, 2/10 stated only some of them could chase items while 1/10 answered they could not. As some respondents correctly identified situations where orphibs had targeting behaviors whilst others stated they did not, we speculate different gameplay experiences might have influenced the agreement discrepancies in this item.

Item 7. The creatures' actions are appropriate to their circumstances - Only 1/10 of the participants considered the orphibs' behavior as normal. From the remaining 9/10, 5/10 claimed it depended on the orphib while 4/10 stated the orphibs' behavior was not normal. When prompted to justify their answers, respondents pointed to several situations they witnessed: the orphibs' limited behaviors, occasional lack of movement, lack of ability to learn and extreme personalities (either too passive or too aggressive). From these results, we theorize the videogame might have lacked further feedback to guide players to better perceive the context behind the orphibs' actions.

Item 10. The creatures react to environmental triggers - Participants answered orphibs reacted to player interactions (7/10), petting/spanking, toys and food (5/10), the fountain (6/10) and the bird (2/10). The majority of respondents found player interactions to be the main event orphibs reacted to, though they might have dismissed it from being part of the environment. Because subjects identified correctly several of the environments' objects, we conjecture the population sample might have misinterpreted the item, most specifically, what they considered to be the environment or related interactions.

Item 11. The creatures can perform tasks by coordinating with other creatures, or humans - 7/10 of the participants assured orphibs did not coordinate with each other, 1/10 were undecided, 10% stated orphibs coordinated through conversations and the remaining 1/10 claimed one orphib went to drink water after "watching another one do it". Given there is some freedom in the gameplay, we cannot put aside different game experiences as a cause, but neither can we ignore the different interpretations the participants might have done of the orphibs' behaviors.

Item 12. The creatures audio-visually communicate with other creatures, or humans - 8/10 correctly identified sounds as their primary source. The remaining 2/10 were not sure if they communicated at all, though one participant did confess to hearing sounds. Akin to the previous situations, results suggest different experiences might have impacted the subjects' perception.

Item 13. The creatures engage in reproductive acts - 5/10 of the participants stated they did not witness such a thing, 1/10 claimed not being able to directly observe it while an additional 1/10 misinterpreted pressing the Egg generator machine as reproduction. From the answers, it became clear some participants did not experience this particular behavior.

Item 15. The creatures learn from past events - 7/10 respondents noted they were not able teach the orphibs anything, while the remaining 3/10 stated they were able to teach them to react to the fountain or to fetch toys, or just that they were able to teach orphibs 'to a degree'. However, not only most participants didn't detect the relevant mechanic, but also what they detected referred to mechanics not present. Regarding the latter, it is possible an expectation bias might have given participants the illusion they were teaching orphibs to learn things they were not programmed to (e.g. learning to 'react to the fountain'). As such, we believe this behavior was not perceivable at all.

Item 16. The creatures use past experience to overcome novel situations - Since the interview's respondents were not able to experience learning, we can assume this item's discrepancies have the same explanation.

Item 17. The creatures change the way they look with age - According to 9/10 of the respondents, size was the distinguishing factor between adults and infants. However, 4/10 of them also noted a change in behaviors when orphibs matured, 1/10 mentioned the male's orphib horn size and another 1/10 hinted at color differences. Nonetheless, 1/10 stated they were not able to distinguish between orphibs of different age. The heterogeneity of the responses leads us to propose most visual cues for aging were ignored most of the time, and it is likely this was because they were too subtle.

Item 18. The creatures change the way they sound with age - The interview's answers suggest aging might not have been perceivable, at all, by participants.

Item 19. The creatures change the way they behave with age - Similarly to the previous statement, it appears most participants did not perceive the gameplay aspects behind this item.

Item 22. The creature shows signs of emotional reaction to objects, events and other creatures - 6/10 mentioned orphibs used sounds (purrs and whines), 2/10 observed orphibs emitting floating hearts, 2/10 did not know whether orphibs expressed emotions or not and another 1/10 stated orphibs 'expressed emotions explicitly'. Because of the different responses obtained, we theorize participants had some difficulty in perceiving how orphibs were able to express emotions.

Item 23. The creatures' bodies seem adapted to their habitat - 8/10 of respondents described forest areas, some detailing the amount of resources present including more food, a river and more caves. 1/10 of them mentioned an ecosystem similar to a savanna while another 1/10 described a household environment. As respondents were able to describe biomes similar to videogame's world, we conjecture the scores' discrepancies might have been caused by a poor mapping between the item and its generative counterpart representation.

6.4 Discussion and Summary of the third DSR iteration

Taking the CBS into account, it was possible to design a videogame where its creatures were guided by the scale's items while, at the same time, the act of using the scale to evaluate creatures could be an integral part of its mechanics. This was the premise behind Orphibs Evolvapalooza.

The CBS items' were first translated into computational elements, from which the Evolvapalooza architecture was devised and then into configuration parameters. The latter was fed into a generative system capable of producing instances of Orphib archetypes. While the example we provided was a mere proof-of-concept, a more thorough dissection of the items could help define a potential search-space. With this in mind, search-based PCG (Togelius et al., 2011a) could be employed to help creature designers traverse through the design space. Additionally, the scale itself could be used to help evaluate any creature discovered from this method. In fact, it could even guide the evaluation. The experiment described in this Chapter illustrates a possible methodology used for this extent.

While the experiment's data helped verify the scale's fourth revision (the one presented in Subchapter 5.5) internal consistency, results did not support agreement between evaluators, with a large population sample. By interviewing some of the participants, we could theorize, with some supporting evidence, several factors that might explain this phenomena. Firstly, the subjects' gameplay experience differed, a common issue when evaluating aspects of the player experience using actual interactive artifacts in lieu of gameplay footage (Togelius, Yannakakis, and Karakovskiy, 2011). As a result, one of the ICC assumptions was not met (that evaluators must observe a certain subject under the same conditions) and this reflected on its values. So, the application of the scale is still subject to be influenced by evaluator experience. Nonetheless, this test was conducted previously, in Chapter 5, with more positive results. However, the previous test had a smaller sample of actual experts who were familiar with most, if not all, of the videogames. Secondly, some of the mappings between the scale's items were not correctly perceived. They were either too subtle or, the resulting implementation was not in line with the subjects' expectations for the item. Thirdly, it is possible that a small set of statements still need an additional round of language simplification.

In the end, Orphibs Evolvapalooza, demonstrated the possibility of both making a videogame whose creatures' design was guided by the scale, the initial steps towards a PCG system for generating them. In fact, the scale played an integral part in the creation phase of the artifact while, at the same time, helped develop an exercise where information could be gathered to further improve it. Moreover, Orphibs Evolvapalooza helped bring light to some potential issues to using a videogame with an integrated validation mechanics, as an experimental setup, in general, and the CBS, in particular.

Chapter 7

Conclusions

In conclusion, this research aims to provide insight on how create a system to aid designers in creating creatures: we identified an opportunity to improve the evaluation phase after deconstructing content creation as a two-phased iterative process. As a result, we provided the initial steps towards a systematic study of believability of videogame creatures. This study materialized in the Creature Believability Scale tested under several content scenarios. The usefulness of the CBS was also evaluated by using it to guide the creation of a proof-of-concept creature generator, in a gameplay context.

This research was conducted through the perspective of several research questions and its effort transpired into potential answers.

7.1 Research Answers

Referring to the question-tree present in Chapter 3, the answers collected, throughout this research, were answered recursively:

1. (b) **How do the creatures' components manifest in identifiable phenomena?**

During this thesis' initial steps, our research attempted to evaluate phenomena that could be perceived by players. This way, a creature could be framed within its videogame context. The literature hints at the notion of expressions, something we later expanded. As we established that a creature, in a videogame, is a composition of a morphology, animations and behaviors, its different components manifest themselves as an audio-visual output, over time. Thus, users can perceive what the creatures do and how they do it. From a videogame perspective, this involves sequencing animations, and sounds, bound by some form of logic, as shown in the Orphibs Evolvapalooza Agent Architecture.

1. (a) i. **Are the state-of-art criteria enough to evaluate believability?**

When considering believability assessment on videogame agents, current state-of-the-art provides two, similar, approaches. On one hand, agents are considered to simulate human beings playing a videogame, so variations of the Turing Test are used (Togelius, Yannakakis, and Karakovskiy, 2011). On the other hand, believability measures how agents fare within a narrative. While the former was beyond the scope of this research, the latter follows the line of works of believable characters (Loyall, 1997; Bates, 1994; Mateas, 1997). These works synthesize several elements a character must have (personality, emotions, self-motivation, the ability to perform parallel actions, social behaviors and behavior consistency) from an anthropomorphic perspective. Because videogames include several entities that are not anthropomorphic in nature, the CBS addresses non-anthropomorphic factors which, as results suggest, seem to contribute to evaluate a subject's perception of creature believability. We conclude the current state-of-the-art criteria are not enough to evaluate believability.

1. (a) ii. **What makes a creature feel believable?**

Building from the previous answer, given that state-of-art criteria is not sufficient to evaluate creature believability, the DSR construction process of the CBS offered several proposals to answer this research question. With this in mind, a creature is perceived as believable when:

1. It is able to react to its environment. This includes being aware of environmental cues and objects, in general, and interact with any of its surroundings, in particular.
2. It moves autonomously and with motor coordination.
3. It behaves consistently within its context.
4. Its morphology is consistent with its habitat.
5. It can mature and reproduce.
6. It can output organic objects (waste, eggs, cubs, etc.).
7. It must use its body as a means to communicate, learn or survive.

However, we were able to construct a model proposal through an inventory of items relevant to the perception of creature believability. We conclude that this proposed inventory (an item set, organized in 4 factors (Barreto, Craveirinha, and Roque, 2017)) contributes a novel and useful perspective on creature believability, beyond the state-of-the art.

1. (a) **How to evaluate believability of videogame agents?**

Psychometric scales are developed to evaluate constructs (Spector, 1992). By nature, these cannot be directly measured but are rather assessed through a set of indicators.

The literature supports the idea that believability is a construct. As such, we can evaluate it through means of a psychometric scale. The CBS provides a formal approach to assess believability and, as results suggests, its individual statements correlate with a subject's perception of this construct. However, the CBS current still needs additional research to be usable in practical situations.

1. (c) **How can the evaluation instrument guide the creation of content?**

The Orphibs Evolvapalooza case study attempts to answer this question. Its creatures' design tries to transcribe the scale's statements into editable parameters. Its implementation was a proof-of-concept, meant to illustrate how this can be made possible. A workflow can then be derived to formalize this aspect:

- During the creation phase of a certain videogame creature, the scale's items can be viewed as heuristics.
- These heuristics can be incorporated into its design and architecture.
- During the evaluation phase, playtesting and expert evaluation can be used, with the CBS as a measuring tool, to validate the previously produced artifact.

- Framed by the scale's items' scores, a second iteration can start to fine tune the creatures.
- This process is repeated until designers are satisfied with their results.

1. **How can we empower creature designers to devise (automated) creatures through an evaluation instrument, capable of measuring creature assets in a contextualizing environment?**

Our research effort started with the goal of providing insights on how to aid the design of videogame creatures. To this extent, our first approach revolved around implementing a PCG system. However, the literature either tackled each component of a creature individually, or used only a functional evaluation method that did not address how creatures were effectively perceived. Throughout this research, we argued this did not provide the knowledge necessary to output complete creatures in an integrated process and thus, while unloading content-creation from human designers (Togelius et al., 2011a), does not seem sufficient to empower content creators designing a complete creature.

There are a few ways we have shown the CBS is capable of improving upon this. By providing a set of guidelines to maximize creatures' believability, content creators can shift their mindset to reflect upon this construct and use it to guide the creation of virtual creatures.

It is also true the Orphibs Evolvapalooza case study demonstrates it is possible to create a *quasi-automated* system that can help explore a search-space defined by the scale itself. The CBS provides a means to help filter ideas from a wide selection (Csikszentmihalyi, 1997), at a level closer to the artifact's final presentation. While this is by no means a definite answer, the work provided by this research helps open the discussion which will lead to it.

7.2 Contributions

Our main contributions are new insight on the Believability concept, the construct underlying the CBS, and the CBS itself. To the extent we have been able to ascertain, the CBS can be a part of the creative process behind creatures.

Our work contributes to the Believable Agent research field, by offering a non-anthropomorphic perspective on how 'believability' can be assessed on videogame creatures.

The empirical results, listed throughout this document, suggest the scale's items contribute to the construct's perception. Thus, the CBS can be used both to evaluate existing creatures and guide the design of new ones.

Moreover, while its main contribution may not lie within the PCG area, as we initially anticipated, it does open the discussion in that direction. The Orphibs Evolvapalooza prototype provided a proof-of-concept illustrating the possibility of developing a generative system guided by the CBS.

This type of exercise can be used in conjunction with a mixed-initiative (Yannakakis, Liapis, and Alexopoulos, 2014) search-based PCG (Togelius et al., 2011a) to explore a search-space of possible creature instances and, ultimately, help designers iterate their creations.

In summary, this research effort offered a new perspective on how a creature can be designed and, even though our method still needs further work, its current form, along with the prototype created, points to several research directions new works can follow.

7.2.1 Limits and Constraints

It is relevant that, while there were some revisions made, ultimately there were not enough resources for the desirable level of validation (a subject further explained in subchapter 7.3). Additionally this scale is intended to be an ongoing research that may be subject to change, in order to become more efficient and/or adequate to other contexts (e.g. different media, also detailed in subchapter 7.3).

This said, the CBS is limited by the participants involved in its conception. It was constructed using a population sample with a distinctly Western culture that might not be valid for an eastern public.

The mapping between the generative system and the scale is subject to the interpretation in each instance. For example, the Evolvapalooza generative system was created and tested by a small sample of participants. The relation between that system and the scale caters to the opinion of the designer and may not be equally valid in other instances.

Furthermore, this research was aimed at the empowerment of designers, and is constrained, as such, to the mixed initiative paradigm (Yannakakis, Liapis, and Alexopoulos, 2014). It is not adequate for the classical PCG paradigm, where the computer acts as both the Designer and the Evaluator (Craveirinha, Barreto, and Roque, 2016).

7.2.2 Publications

During this thesis research process, several of its steps resulting in publications. Part of the literature review, from Chapter 2, resulted in the following articles:

- Barreto, Nuno and Licinio Roque (2014). "A Survey of Procedural Content Generation tools in Video Game Creature Design". In: *Proceedings of the Second Conference on Computation Communication Aesthetics and X*. Porto, Portugal.
- Barreto, Nuno, Amílcar Cardoso, and Licínio Roque (2014). "Computational Creativity in Procedural Content Generation: A State of the Art Survey". In: *Conferência de Ciências e Artes dos Videojogos*. Barcelos, Portugal.

The Orphibs II prototype, built during the exploratory phase, listed in Chapter 4, was detailed in 2 scientific papers:

- Barreto, Nuno, Luís Macedo, and Licinio Roque (2014). "MultiAgent System Architecture in Orphibs II". In: *Proceedings of the 14th International Conference on the Synthesis and Simulation of Living Systems (ALIFE 14)*. New York City, USA: MITPress, pp. 588–595.
- Barreto, Nuno et al. (2015). "Incorporation of Emotions in the Orphibs' Agent Architecture". In: *Proceedings of the 13th European Conference on Artificial Life (ECAL 2015)*. York, England: MIT Press, p. 311.

Some of the revisions made in the creation process of the CBS are depicted in an additional 2 papers:

- Barreto, Nuno, Rui Craveirinha, and Licinio Roque (2017). "Designing a Creature Believability Scale for Videogames". In: *Proceedings of the 2017 International Conference on Entertainment Computing 2017*. Tsukuba City, Japan. Springer.
- Barreto, Nuno, Rui Craveirinha, and Licinio Roque (2018). "Validating the Creature Believability Scale for Videogames". In: *Entertainment Computing – ICEC 2018*. Poznań, Poland. Springer.

7.3 Future Work

We foresee some future directions this kind of research can follow. First, additional tests should be conducted to the CBS to verify whether its fourth revision improved its goodness-of-fit (as our initial CFA suggested a reasonable fit). An additional CFA must be conducted, with a sample pool of between 300 to 500 entries. Though the issue was analyzed from different points of view, it was focused on that of believable agents. Emphasing the other, less studied, points of view (as is the case of biology), might reveal other relevant items while de-valuing some already present. So, after the CBS has been more thoroughly validated, several research questions can be considered:

1. **Can we consider creature archetypes with different expectations?** - When thinking about existing animals, we notice they vary in complexity, and we can assume they invoke different expectations (for instance, people might expect a simpler performance from an insectoid creature but not from a mammal-like one). Research can be conducted to understand the links between a creature's appearance and what subset of the CBS is sufficient to assess its believability.
2. **Can the Creature Believability Scale be used in other media?** - While the scale was created with videoludic artifacts in mind, we believe it is not its only application. The scale could be applied to other media and research can be conducted in them to determine if its items remain relevant. In fact, the scale's initial experiments showed the scale could be administered with a more passive medium, such as videos.
3. **How does the Creature Believability Scale adapt to different cultures?** - As mentioned, the scale is limited by the cultural perspective of the population sample used in its construction. It is necessary to understand how different cultures view the artifacts created using the CBS as it is.

Another interesting direction could follow improving the creature believability spectrum. Though it was an supplementary artifact, created from the CBS main research, works such as this contribute to game studies and helps give insights, from a theorical point-of-view, on which factors (the creatures' role within the game, the game's genre, etc.) translate into development effort.

The development of the Orphib Evolvapalooza videogame prototype originated additional research questions:

1. **How can we incorporate a player experience evaluation as a game mechanic?** - Even though this was attempted during the prototype's development, results showed our implementation hindered it. Nonetheless, we conjecture this might create an opportunity for future research, to understand how evaluation methods, such as questionnaires, can be incorporated as game mechanics in such a way that players are motivated to complete them.

2. **How can we translate the Creature Believability Scale's items into computational systems** - Another question initiated by the prototype, as mentioned in Chapter 6. Our proposal was only one of many solutions. Research can be conducted to understand the multiple possible computational translations of each CBS item.

3. **How can a Procedural Content Generation system make use of the Creature Believability Scale?** - The generative proposal, employed in Orphibs Evolvapalooza, was a simplified version of how we idealize a PCG system. Achieving that ideal would require defining a search space from the CBS for the PCG system to traverse. Additional research could be conducted either in this particular direction or to try to understand alternatives to this method.

Appendix A

Surveys

This appendix transcribes each survey used throughout this research, with respect to how they were presented to the subjects.

A.1 Behavior and Expression Survey

Thank you for participating in this survey. This survey is part of a PhD project centering on behavior and expression modeling in virtual creatures for use in video-games.

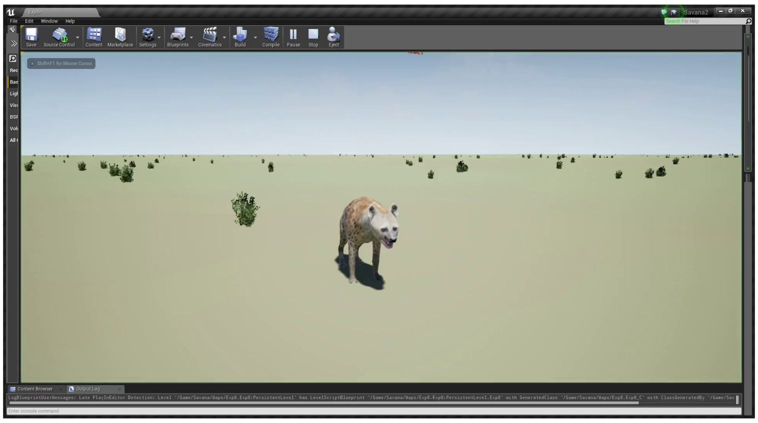
How old are you?

- 17 or less
- 18 to 30
- 31 to 59
- 60 or more

How much time do you spend playing video-games weekly?

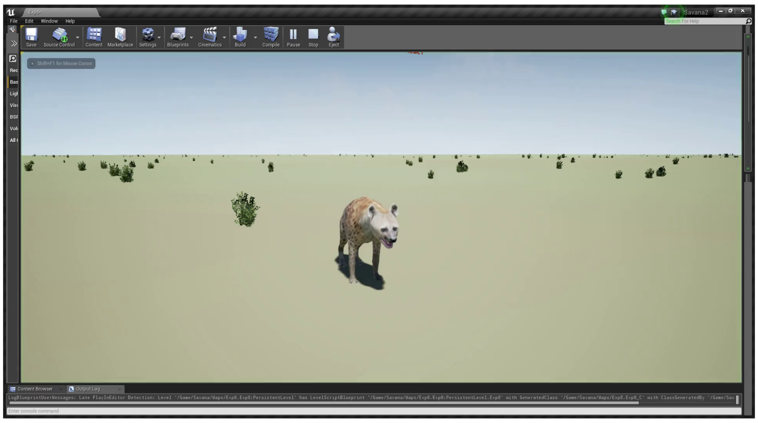
- 0 - 1 hour
- 2 - 10 hours
- 10 - 20 hours
- More than 20 hours

Before answering any further questions, please observe the video below. You may consult it at any time during the survey.



In the video, the Hyena displayed several expressions. In 5 sentences or less, describe the Hyena's behavior over time.

Before answering any further questions, please observe the video below. You may consult it at any time during the survey.



At each of the following moments in the video, what do you think the Hyena was expressing?

From 00:00 to 00:05?

1. Nervousness
2. Nothing
3. Aggression
4. Stress

From 00:06 to 00:12?

1. Nothing
2. Hunger
3. Happiness
4. Sadness

From 00:13 to 00:19?

1. Aggression
2. Tiredness
3. Submission
4. Nothing

From 00:20 to 00:26?

1. Aggression
2. Hunger
3. Happiness
4. Nothing

From 00:27 to 00:34?

1. Nothing

2. Confusion
3. Curiosity
4. Hunger

A.2 Creature Believability Questionnaire

This survey is part of a PhD research project developed at the University of Coimbra, Portugal.

For any problems with the HIT, please contact the PhD student responsible for this study: Nuno Barreto (nbarreto@dei.uc.pt).

The following survey aims to provide a believability framework for videogame creatures.

How old are you?

What is your sex?

- Male
- Female

What is your highest achieved education?

- High school
- Bachelor's degree
- Master's degree
- Professional degree
- Doctorate degree
- Other

What is your weekly contact with any kind of audiovisual media (movies, videogames, etc.)?

- Up to 20 hours/week
- 20-40 hours/week
- 40-60 hours/week
- 60-80 hours/week
- 80-100 hours/week

In the past 30 days, please name, at most, 3 videogames you have played

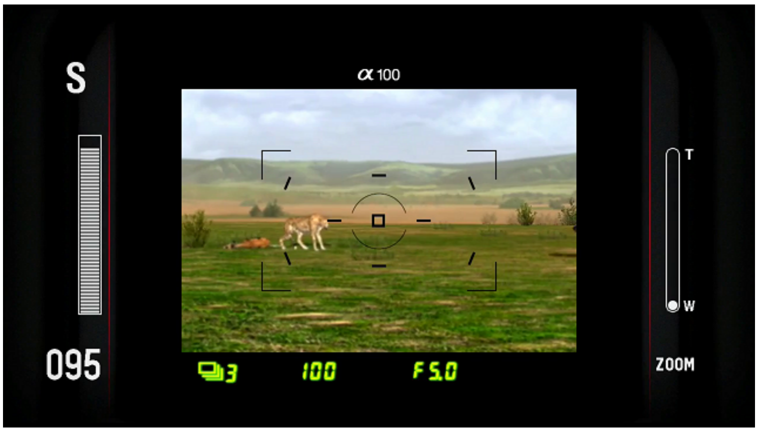
---

Instructions

During this survey you will:

1. Watch one of twenty-eight 20-30 second-long video clips, showcasing one, or several, videogame creatures, engaged in several activities.
2. Answer a small survey regarding the creatures in the clip.
3. Repeat step 1, until you viewed all 28 clips.

Responses will be rejected if they fail to answer control questions adequately and are abnormally below the expected completion time.



Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures work with other creatures for a common goal	☐	☐	☐	☐	☐
Each of the creatures' body parts have inertia	☐	☐	☐	☐	☐
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures play by themselves	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures expel material	☐	☐	☐	☐	☐
The creatures' motions are fluid	☐	☐	☐	☐	☐
The creatures work with other creatures for a common goal	☐	☐	☐	☐	☐
The creatures feel empathy towards other creatures	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures make tools	☐	☐	☐	☐	☐

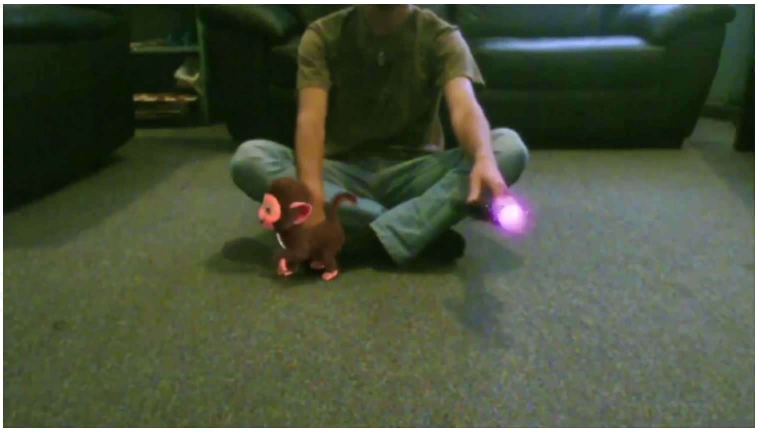
How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable



Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures play by themselves	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures have different life-stages	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures show anticipatory expressions to known stimuli	☐	☐	☐	☐	☐
The creatures express moods through its body	☐	☐	☐	☐	☐
Postures and expressions seem coherent with the creatures' behavior	☐	☐	☐	☐	☐
The creatures' behaviors differ from other creatures of the same species	☐	☐	☐	☐	☐
The creatures use objects from the environment	☐	☐	☐	☐	☐
The creatures make tools	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures' bodies seem adapted to their habitat	☐	☐	☐	☐	☐
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures have different life-stages	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures control their body	☐	☐	☐	☐	☐
The creatures communicate with other, different-species, creatures	☐	☐	☐	☐	☐
The creatures' motions are fluid	☐	☐	☐	☐	☐
The creatures absorb substances/energy from the environment to survive	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐
The creatures play by themselves	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' actions involve more than one step	☐	☐	☐	☐	☐
The creatures alternate between tasks	☐	☐	☐	☐	☐
The creatures' actions seem appropriate to their context	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' motions reflect their weight and size	☐	☐	☐	☐	☐
Each of the creatures' body parts have inertia	☐	☐	☐	☐	☐
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures alternate between tasks	☐	☐	☐	☐	☐
The creatures' behaviors differ from other creatures of the same species	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
The creatures' motions are fluid	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures absorb substances/energy from the environment to survive	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' expressions anticipate its actions	☐	☐	☐	☐	☐
The creatures' motions reflect their weight and size	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures use objects from the environment	☐	☐	☐	☐	☐
The creatures feel empathy towards other creatures	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures show positive (or negative) emotions towards objects or events	☐	☐	☐	☐	☐
The creatures use objects from the environment	☐	☐	☐	☐	☐
The creatures express moods through their body	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures change the way they behaves with age	☐	☐	☐	☐	☐
The creatures have traits particular to their sex	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behaves with age	☐	☐	☐	☐	☐
The creatures absorb substances/energy from the environment to survive	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures make tools	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' bodies and behaviors seem consistent	☐	☐	☐	☐	☐
The creatures appear to have traits particular to their sex	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures play by themselves	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures' motions reflect their weight and size	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
The creatures have traits particular to their sex	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures plays by themselves	☐	☐	☐	☐	☐
Postures and expressions seem coherent with the creatures' behaviors	☐	☐	☐	☐	☐
The creatures' actions seem appropriate to their context	☐	☐	☐	☐	☐
The creatures absorb substances	☐	☐	☐	☐	☐
The creatures' behaviors differ from other creatures of the same species	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures absorb substances or energy from the environment to keep alive	☐	☐	☐	☐	☐
The creatures play with others	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures make tools	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐
The creatures plays by themselves	☐	☐	☐	☐	☐
The creatures absorb substances	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures recognize themselves	☐	☐	☐	☐	☐
The creatures plays by themselves	☐	☐	☐	☐	☐
The creatures moves by themselves	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures work with other creatures for a common goal	☐	☐	☐	☐	☐
The creatures communicate with other, different-species, creatures	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' behaviors differ from other creatures of the same species	☐	☐	☐	☐	☐
The creatures absorb substances or energy from the environment to keep alive	☐	☐	☐	☐	☐
The creatures appear to have diverse priorities	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures' bodies and behaviors seem consistent	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures play by themselves	☐	☐	☐	☐	☐
The creatures have different life-stages	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures express moods through its body	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐
The creatures appear to have different life-stages	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures feel empathy towards other creatures	☐	☐	☐	☐	☐
The creatures make tools	☐	☐	☐	☐	☐
The creatures absorb substances/energy from the environment to survive	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects or events	☐	☐	☐	☐	☐
The creatures express moods through its body	☐	☐	☐	☐	☐
The creatures communicate with other, different-species, creatures	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures work with other creatures for a common goal	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures appear to have different life-stages	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures recognizes themselves	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures make tools	☐	☐	☐	☐	☐
The creatures communicate with other, different-species, creatures	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures direct its behaviors towards targets	☐	☐	☐	☐	☐
The creatures feel empathy towards other creatures	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures expel material	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures' behaviors differ from other creatures of the same species	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip, indicate to what extent you agree or disagree with the following statements?

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures feel empathy towards other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, different-species, creatures	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures plays by themselves	☐	☐	☐	☐	☐
The creatures absorb substances/energy from the environment to survive	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

How do you find this/these creature(s) believable?

By believable, we mean presented in such a way that it does not break suspension-of-disbelief.

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

How do you find this setting believable?

	1	2	3	4	5	6	7	8	9	10	
Believable	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Non-Believable

A.3 Creature Believability Questionnaire II

Thank you for participating in this survey.

This is part of a PhD research project, developed at the University of Coimbra, Portugal. The research's main goal is to develop a creature believability assessment tool, to help guide creature design and implementation in video games.

The survey uses gameplay footage taken from commercial games (copyrights belong to their respective owners), released over the last decade.

Video clips from that footage will be shown throughout the survey. After each one, you will be prompted to answer a few questions about what you have just seen.

The data collected from this survey will then be used to validate the assessment tool.

For any issue pertaining this survey, please contact the PhD student responsible: Nuno Barreto (nbarreto@dei.uc.pt).

How old are you?

____

What is your sex?

- Male
- Female

What is your highest achieved education?

- High school
- Bachelor's degree
- Master's degree
- Professional degree
- Doctorate degree
- Other

What is your weekly contact with any kind of audiovisual media (movies, videogames, etc.)?

- Up to 20 hours/week
- 20-40 hours/week
- 40-60 hours/week
- 60-80 hours/week
- 80-100 hours/week

In the past 30 days, please name, at most, 3 videogames you have played

____

____

____

Instructions

During this survey you will: Watch one of ten 40-55 second-long video clips, showcasing one, or several, videogame creatures, engaged in several activities. Answer a small survey regarding the creatures in the clip. Repeat step 1, until you viewed all 10 clips. Responses will be rejected if they fail to answer control questions adequately and are abnormally below the expected completion time.

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures’ expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures’ body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

____

____

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

Describe how many of each Creature you see on the video? (For ex.: one rhino, two elephants and one red alien)

By creatures, we mean entities with animal-only characteristics inspired by existing living-beings. Humans, or human-like beings, do not count towards this definition.

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Relationship with the Environment

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures interact with the environment	☐	☐	☐	☐	☐
The creatures control their body	☐	☐	☐	☐	☐
The creatures direct their behaviors towards targets	☐	☐	☐	☐	☐
The creatures locate objects in the environment	☐	☐	☐	☐	☐
The creatures expel material	☐	☐	☐	☐	☐
The creatures' actions are appropriate to their context	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Biological/Social Plausibility

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures move by themselves	☐	☐	☐	☐	☐
The creatures' motions reflect their weight/size	☐	☐	☐	☐	☐
The creatures make several simultaneous motions	☐	☐	☐	☐	☐
The creatures react to stimuli	☐	☐	☐	☐	☐
The creatures focus on stimuli	☐	☐	☐	☐	☐
The creatures coordinate with other creatures	☐	☐	☐	☐	☐
The creatures communicate with other, same-species, creatures	☐	☐	☐	☐	☐
The creatures engage in reproductive acts	☐	☐	☐	☐	☐
There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Adaptation

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' same-stimuli reactions change over time	☐	☐	☐	☐	☐
The creatures learn from past events	☐	☐	☐	☐	☐
The creatures are able to apply old behaviors to new, similar, situations	☐	☐	☐	☐	☐
The creatures change the way they look with age	☐	☐	☐	☐	☐
The creatures change the way they sound with age	☐	☐	☐	☐	☐
The creatures change the way they behave with age	☐	☐	☐	☐	☐

After watching the clip please indicate, using the provided scale, how much do you agree or disagree with the following statements?

Expression

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
The creatures' expressions anticipate their actions	☐	☐	☐	☐	☐
The creatures show positive (or negative) emotions towards objects, or events	☐	☐	☐	☐	☐
The creatures show expressions to known stimuli	☐	☐	☐	☐	☐
The creatures learn through imitation	☐	☐	☐	☐	☐
The creatures' body are adapted to their habitat	☐	☐	☐	☐	☐

A.4 Orphibs Evolvapalooza Follow-up Questionnaire

Thank you for participating in the Orphibs: Evolvapalooza experiment. With this follow up questionnaire, we hope to uncover potential issues you may have experienced with either the game, or the scale. Please, try to recall the creatures you evaluated during the experiment (Hint: They were the labelled ones) and answer as truthfully as possible to the following questions:

Please describe, in 10 lines or less, your experience playing Orphibs: Evolvapalooza

Are Orphibs able to chase after toys, food, or any other thing?

What did the Orphibs react to?

Did you teach Orphibs anything?

Can you please describe ways the Orphibs communicate?

Can you please describe ways Orphibs express their emotions?

Can you please describe the way Orphibs reproduce?

How did you distinguish between baby and adult Orphibs?

How did Orphibs coordinate with each other?

Please say whether you consider the Orphibs' behavior as normal?

Please explain why?

Please say whether you consider the Orphibs' body-movement as normal?

Please explain why?

Can you please describe what you believe should be the Orphibs' natural habitat?

Appendix B

Creature Believability Scale Construction Auxiliary Material

B.1 Creature Believability Scale v0.3 Experiment Boxplots

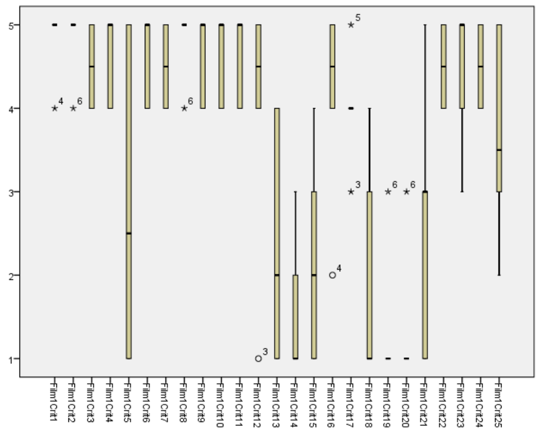


FIGURE B.1: Trico (The Last Guardian (genDESIGN and SIE Japan Studio, 2016)) boxplot (n = 6).

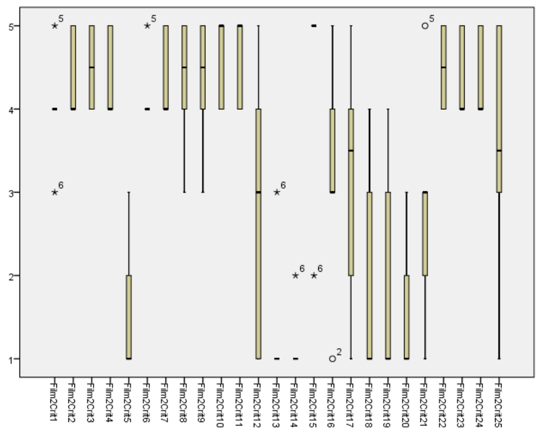


FIGURE B.2: Eyepet (EyePet (SCE London Studio and Playlogic Game Factory, 2009)) boxplot (n = 6).

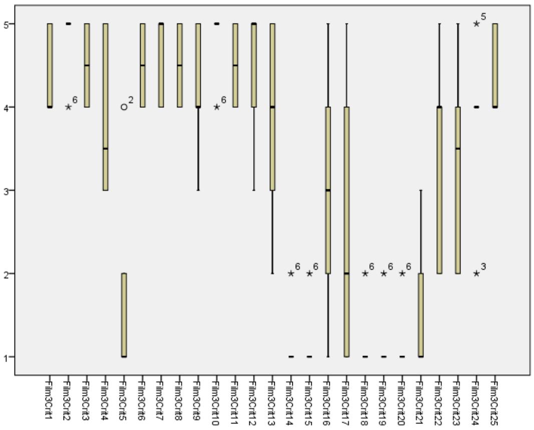


FIGURE B.3: Strider (Horizon Zero Dawn (Guerrilla Games, 2017)) boxplot (n = 6).

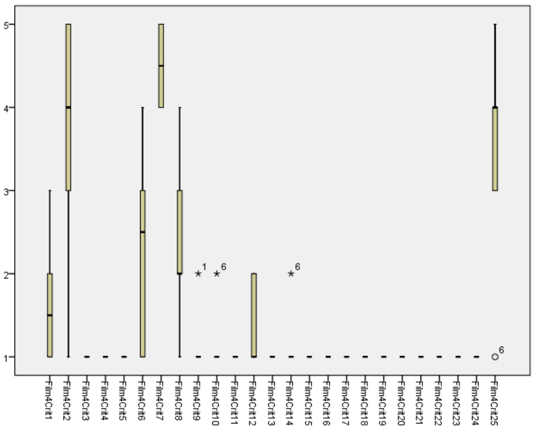


FIGURE B.4: Armadillos (from Everything (David OReilly, 2017)) boxplot (n = 6).

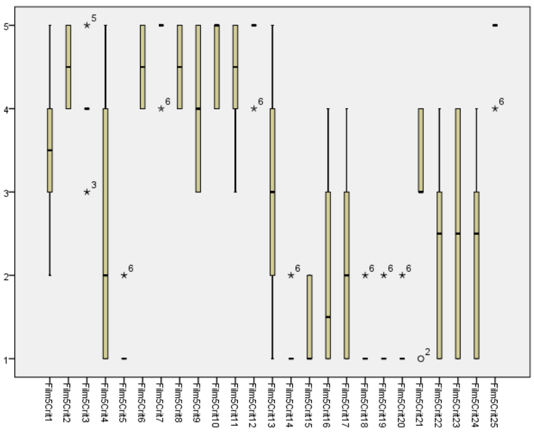


FIGURE B.5: Threadfin Butterflyfishes (Abzû (Giant Squid Studios, 2016)) boxplot (n = 6).

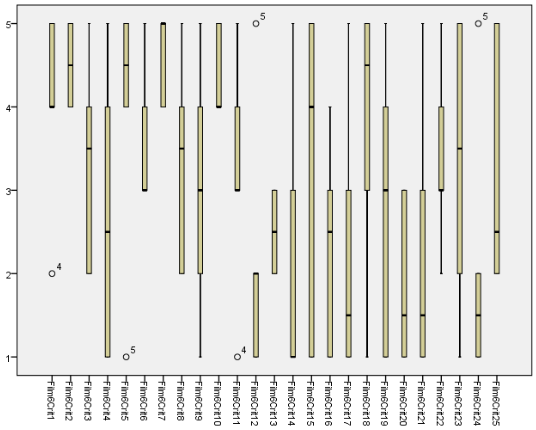


FIGURE B.6: Cifiarae Varsaku (No Man's Sky (Hello Games, 2016)) boxplot (n = 6).

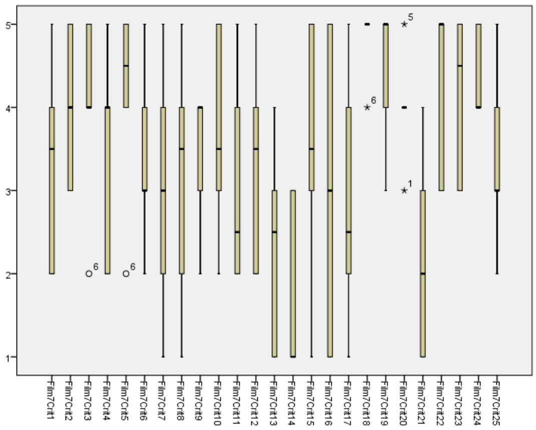


FIGURE B.7: Botamon (Digimon World: Next Order (B.B. Studio, 2017)) boxplot (n = 6).

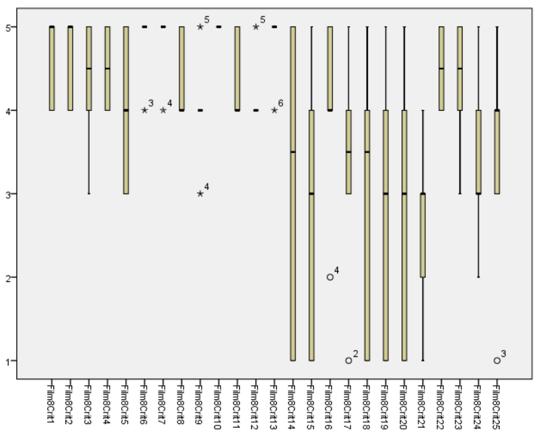


FIGURE B.8: Cat (from The Sims 3 (Maxis, 2009)) boxplot (n = 6).

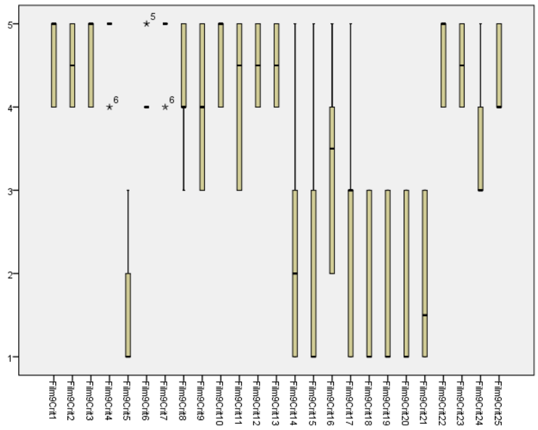


FIGURE B.9: Squirrel (SimAnimals (Electronic Arts, 2009)) boxplot (n = 6).

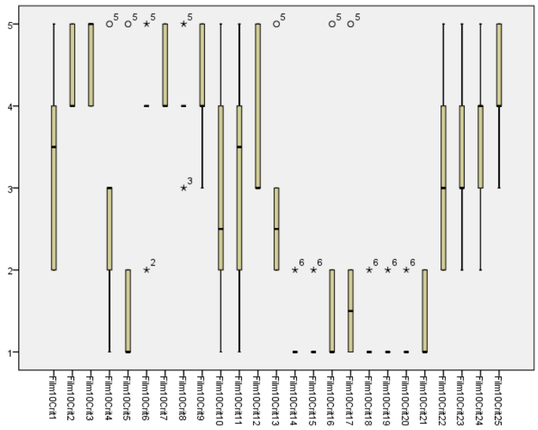


FIGURE B.10: Raptor (Turok (Propaganda Games, 2008)) boxplot (n = 6).

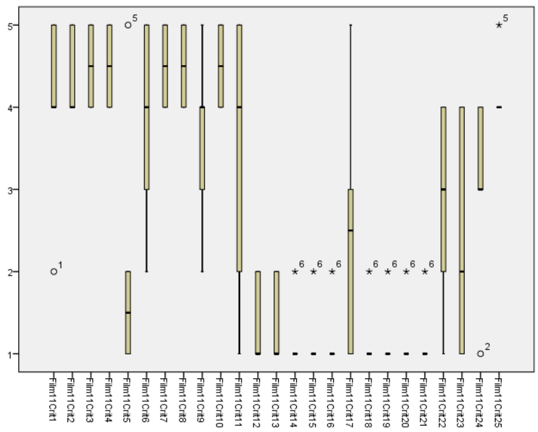


FIGURE B.11: Alligator (Mafia III (Hangar 13, 2016)) boxplot (n = 6).

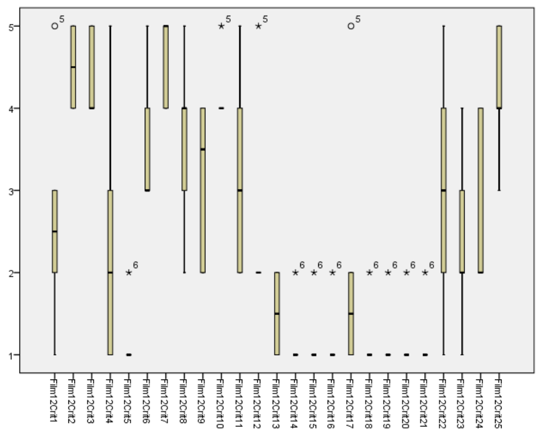


FIGURE B.12: Boar (S.T.A.L.K.E.R.: Call of Pripryat (GSC Game World, 2010)) boxplot (n = 6).

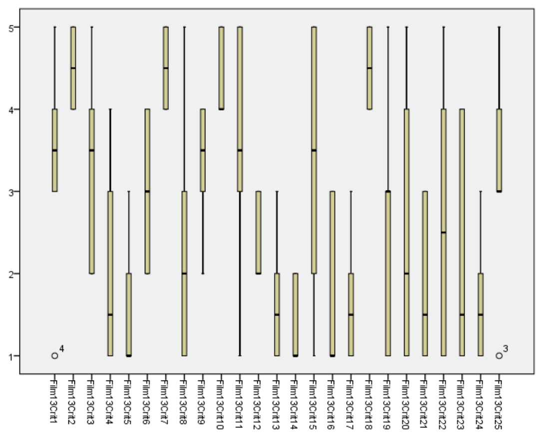


FIGURE B.13: Pig (Minecraft (Mojang, 2011)) boxplot (n = 6).

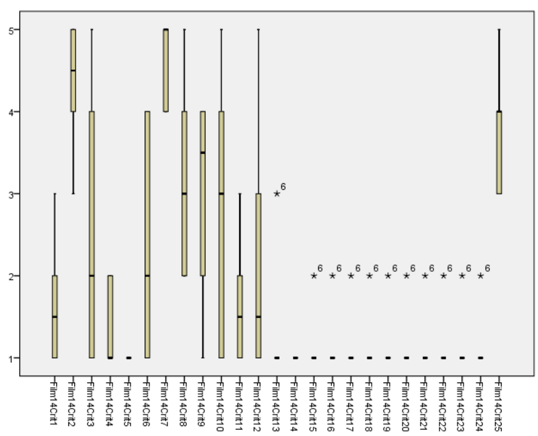


FIGURE B.14: Lyote (StarCraft II: Legacy of the Void (Blizzard Entertainment, 2015)) boxplot (n = 6).

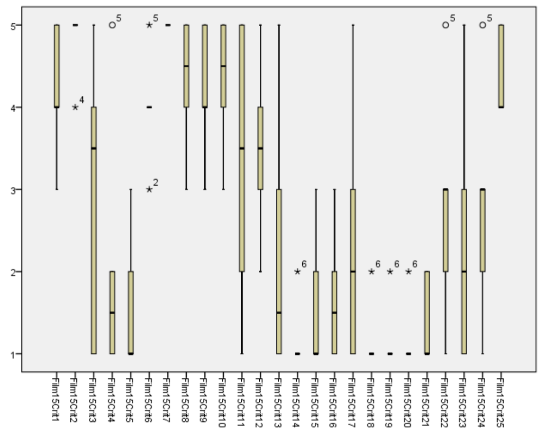


FIGURE B.15: Deer (Witcher III: The Wild Hunt (CD Projekt RED, 2015)) boxplot (n = 6).

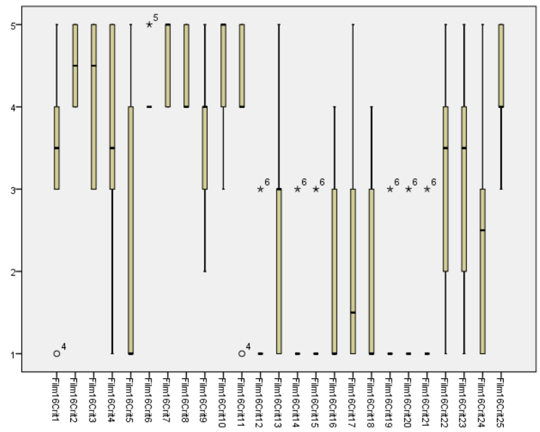


FIGURE B.16: Whitetail Deer (theHunter (Games, 2005)) boxplot (n = 6).

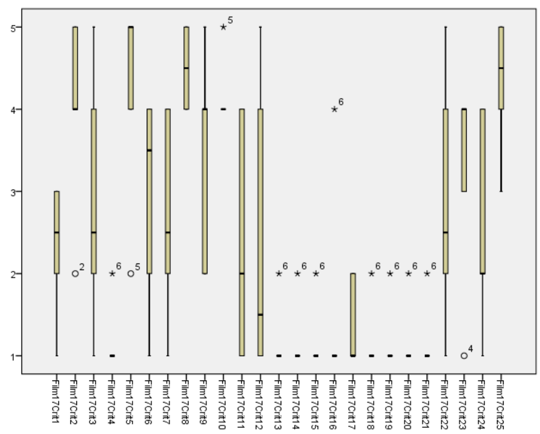


FIGURE B.17: D-Horse (Metal Gear Solid V: The Phantom Pain (Kojima Productions, 2015)) boxplot (n = 6).

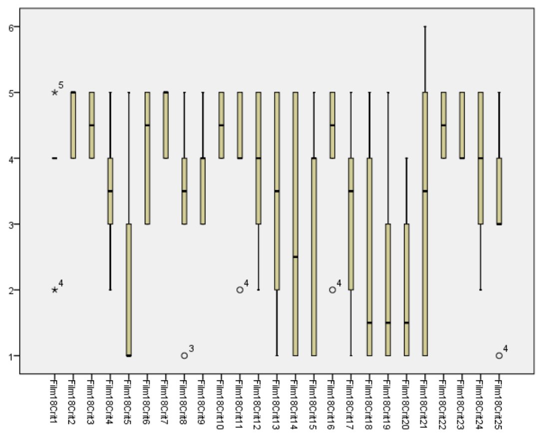


FIGURE B.18: Rasteirist (Spore (Maxis, 2008)) boxplot (n = 6).

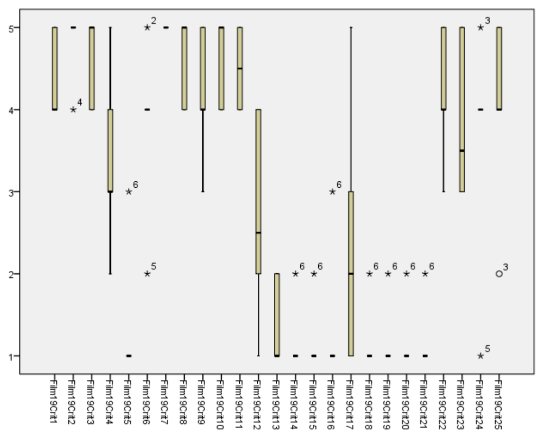


FIGURE B.19: Rhinoceros (FarCry 4 (Ubisoft Montreal, 2014)) boxplot (n = 6).

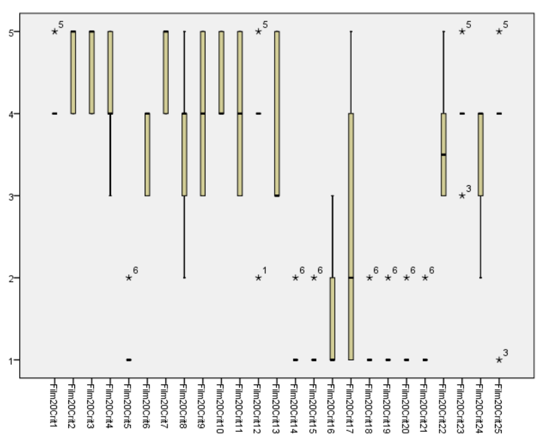


FIGURE B.20: Lizard (Rain World (Videocult, 2017)) boxplot (n = 6).

B.2 Creature Believability Scores

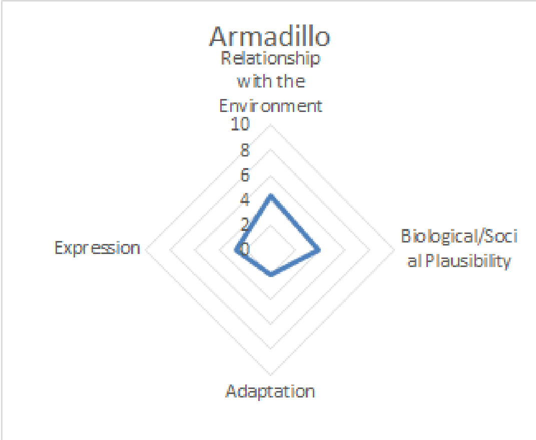


FIGURE B.21: Believability index, for the Armadillo creature, across each of the CBS dimensions.

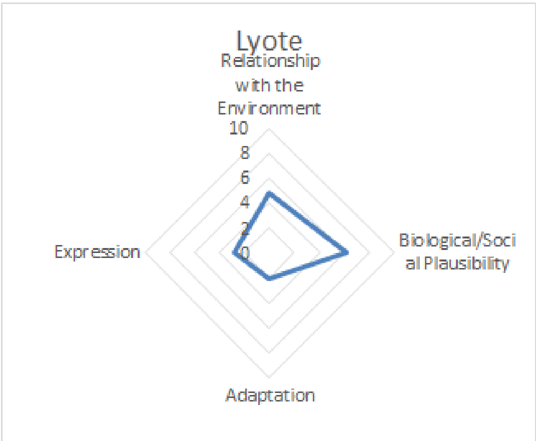


FIGURE B.22: Believability index, for the Lyote creature, across each of the CBS dimensions.

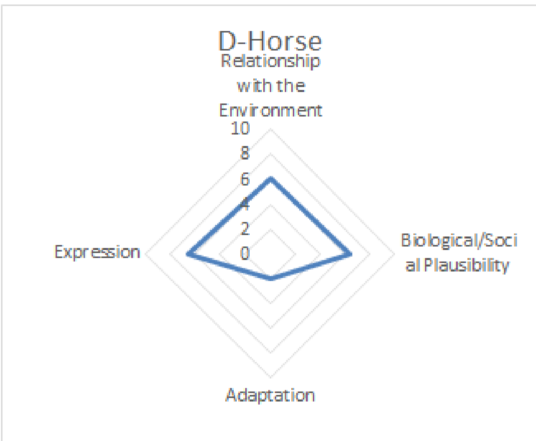


FIGURE B.23: Believability index, for the D-Horse creature, across each of the CBS dimensions.

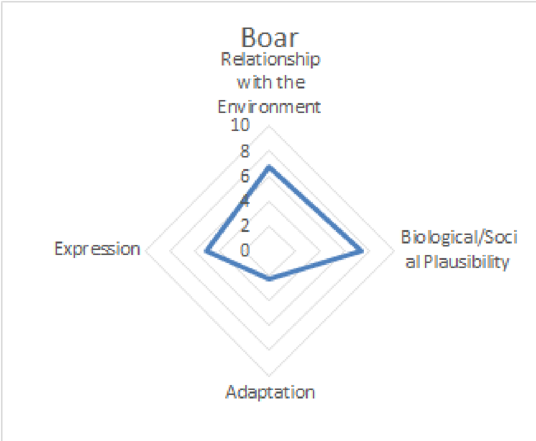


FIGURE B.24: Believability index, for the Boar creature, across each of the CBS dimensions.

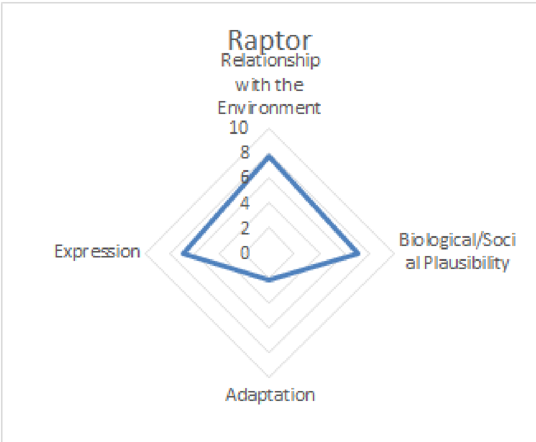


FIGURE B.25: Believability index, for the Raptor creature, across each of the CBS dimensions.

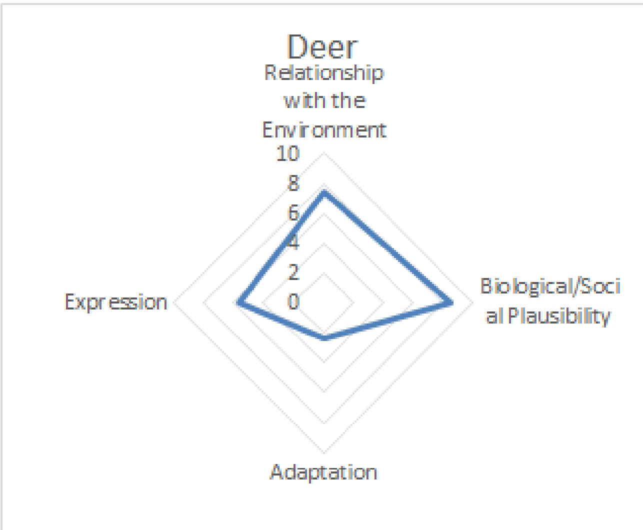


FIGURE B.26: Believability index, for the Deer creature, across each of the CBS dimensions.

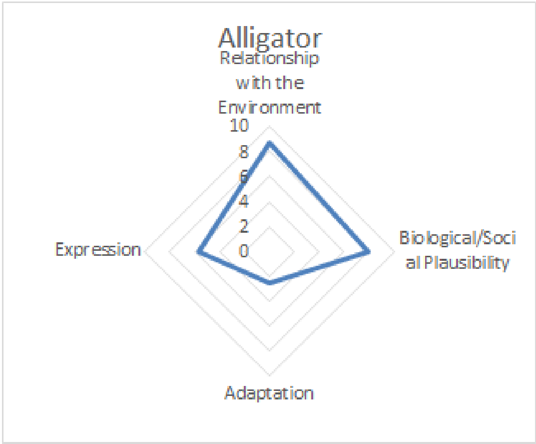


FIGURE B.27: Believability index, for the Alligator creature, across each of the CBS dimensions.

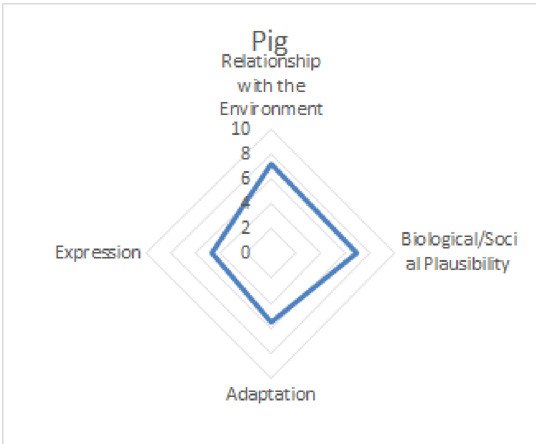


FIGURE B.28: Believability index, for the Pig creature, across each of the CBS dimensions.

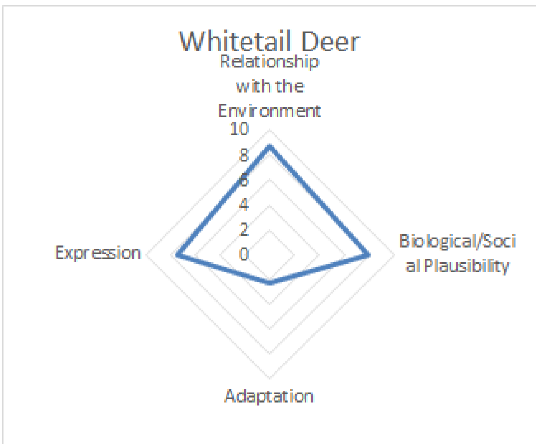


FIGURE B.29: Believability index, for the Whitetail Deer creature, across each of the CBS dimensions.

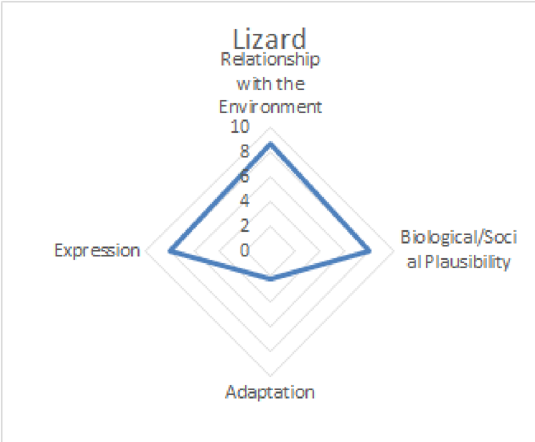


FIGURE B.30: Believability index, for the Lizard creature, across each of the CBS dimensions.

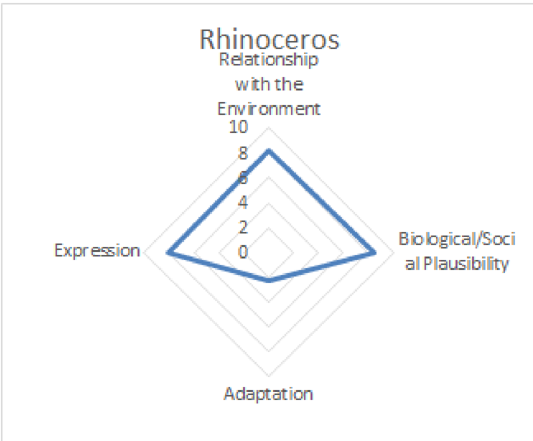


FIGURE B.31: Believability index, for the Rhinoceros creature, across each of the CBS dimensions.

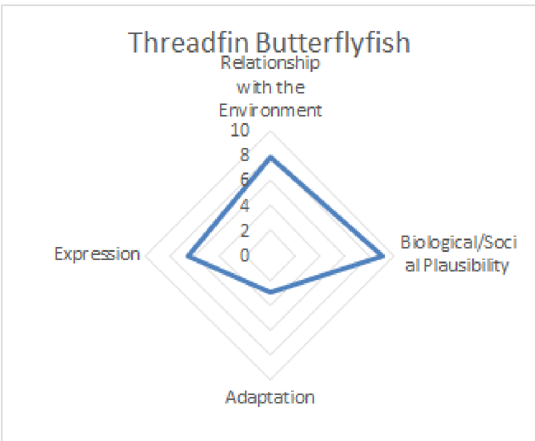


FIGURE B.32: Believability index, for the Threadfin Butterflyfish creature, across each of the CBS dimensions.

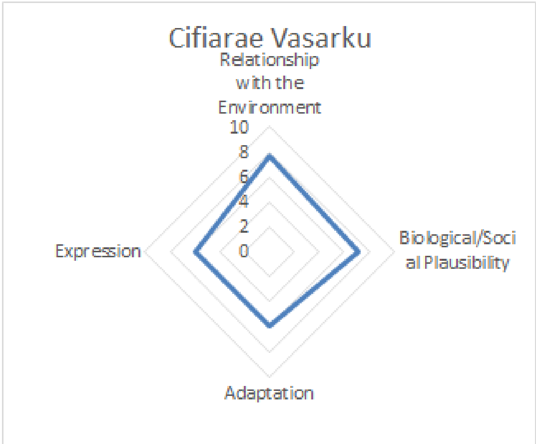


FIGURE B.33: Believability index, for the Cifiarae Vasarku creature, across each of the CBS dimensions.

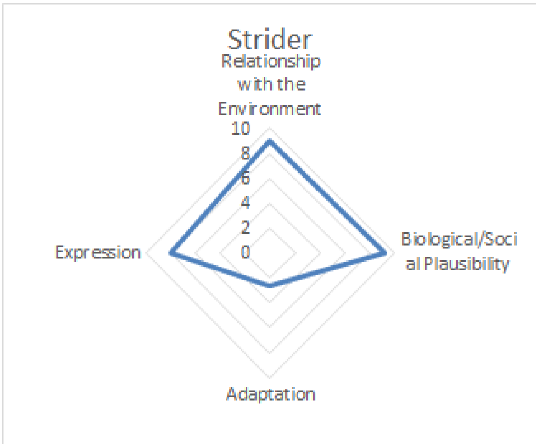


FIGURE B.34: Believability index, for the Strider creature, across each of the CBS dimensions.

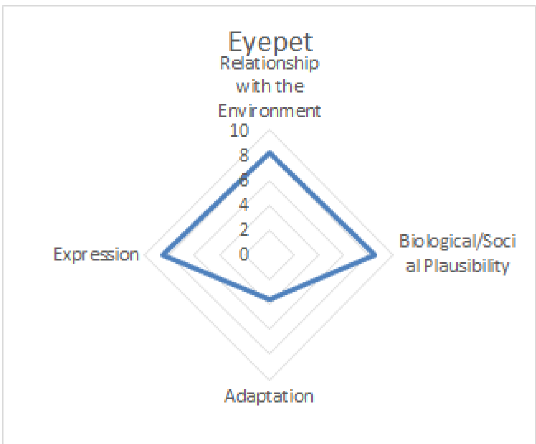


FIGURE B.35: Believability index, for the Eyepet creature, across each of the CBS dimensions.

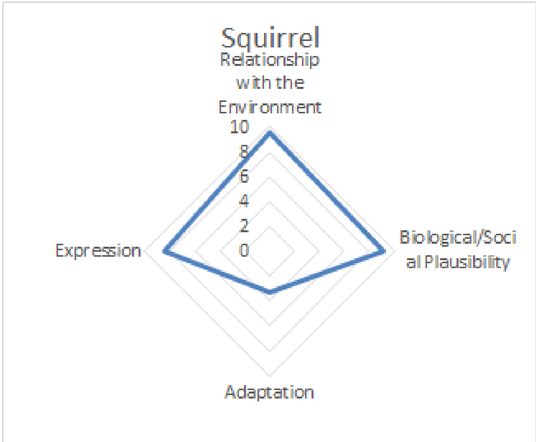


FIGURE B.36: Believability index, for the Squirrels creature, across each of the CBS dimensions.

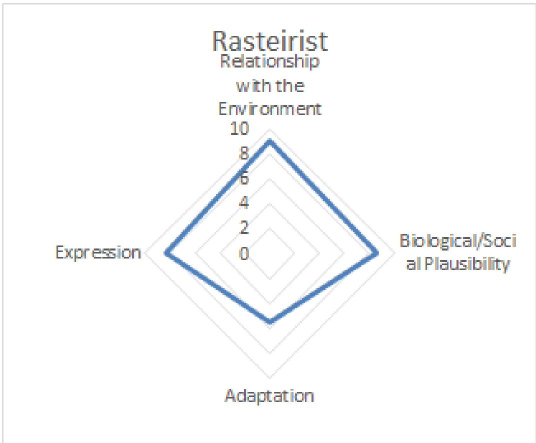


FIGURE B.37: Believability index, for the Rasteirist creature, across each of the CBS dimensions.

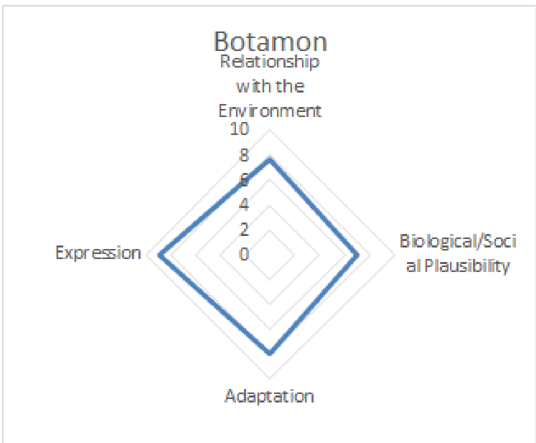


FIGURE B.38: Believability index, for the Botamon creature, across each of the CBS dimensions.

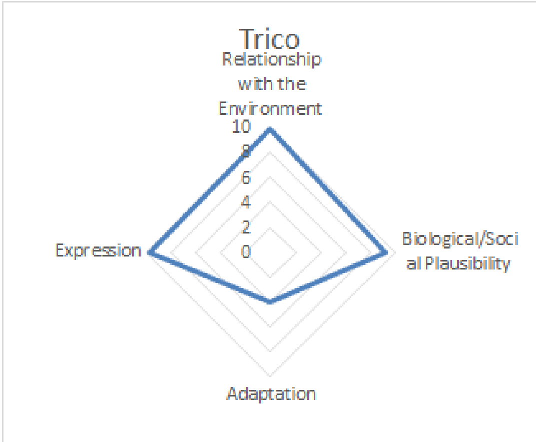


FIGURE B.39: Believability index, for the Trico creature, across each of the CBS dimensions.

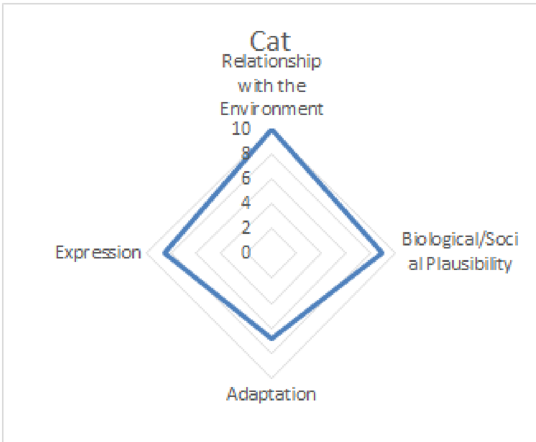


FIGURE B.40: Believability index, for the Cat creature, across each of the CBS dimensions.

Appendix C

Creature Believability Scale v0.4

TABLE C.1: The Creature Believability Scale v0.4. The likert-points (NAA, S, SW, M, TAGE) signify 'Not At All', 'Slightly', 'Somewhat', 'Moderately' and 'To a Great Extent' respectively.

		NAA	**S**	**SW**	**M**	**TAGE**
Relationship with the Environment	1. The creatures interact with surrounding environment					
	2. The creatures have motor co-ordination					
	3. The creatures direct their movement towards/away from targets					
	4. The creatures are able to find objects in the environment					
	5. The creatures are able to expel organic waste					
	6. There are signs of previous expelled organic waste, such as feces, urine, goo, etc.					
	7. The creatures' actions are appropriate to their circumstances					
Biological/Social Plausibility	8. The creatures are able to move independently					
	9. The creatures' agility reflect their volume					
	10. The creatures react to environmental triggers					
	11. The creatures can perform tasks by coordinating with other creatures, or humans					
	12. The creatures audio-visually communicate with other creatures, or humans					
	13. The creatures engage in reproductive acts					
	14. There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.					

Adaptation	15. The creatures learn from past events					
	16. The creatures use past experience to overcome novel situations					
	17. The creatures change the way they look with age					
	18. The creatures change the way they sound with age					
	19. The creatures change the way they behave with age					
	20. The creatures can imitate others to develop their own behaviors					
Expression	21. The creatures' body language anticipate their actions					
	22. The creature shows signs of emotional reaction to objects, events and other creatures					
	23. The creatures' bodies seem adapted to their habitat					

Appendix D

Orphibs Evolvapalooza: Final Game Design Document

Following the design philosophy of Petz (PF Magic, 1995), also underlain in Orphibs II, Orphib Evolvapalooza is a cartoon life-simulation videogame, where players take on the role of an animal raiser. It is inspired by the comedic interactions (e.g. by using unusual objects, such as toilet paper, to pet the animal) made available in Macho Cat: Pat the Virtual Cat (Sia Ding Shen, 2017) and the pet raising game mechanics (such as being able to pet, feed and watch the animal grow) in Creatures (Millenium Interactive, 1996), the Petz (PF Magic, 1995) franchise and Sonic Adventure's (Sonic Team USA, 2003) Chaos Garden mini-game.

Aesthetically, the videogame is meant to invoke calm, and playful, feelings, so that players feel invited to interact, and play, with the creatures.

D.1 Game Flow

While the videogame's objective is for players to evaluate creatures of 5 different generations, players are communicated that their main objective is to raise a generation of the most believable creatures, according to the CBS. In order to raise them, players can feed, pet and spank them. Furthermore, they are rewarded for interacting with the creatures. This comes in the form of points that can be used to evaluate creatures (Evolution Points), create a batch of creatures (Egg Points) and to buy raffles and win additional items (Money).

The videogame starts by collecting the players' demographic data and then greets them with a small text explaining the purpose of the videogame/experiment and what is expected of them. Then, they are given money and the actual game loop starts: players are expected to generate a batch of eggs. At this time, the generated eggs hatch and they can either use items to interact with creatures, or buy new ones, if they have enough money. Every time an item is used on a creature, the player gains points. With enough of them, one of the creatures can be evaluated and money is awarded for this action. This is meant to drive the players' attention towards exploring interactions with the creatures, so that evaluation is based on exposure to them.

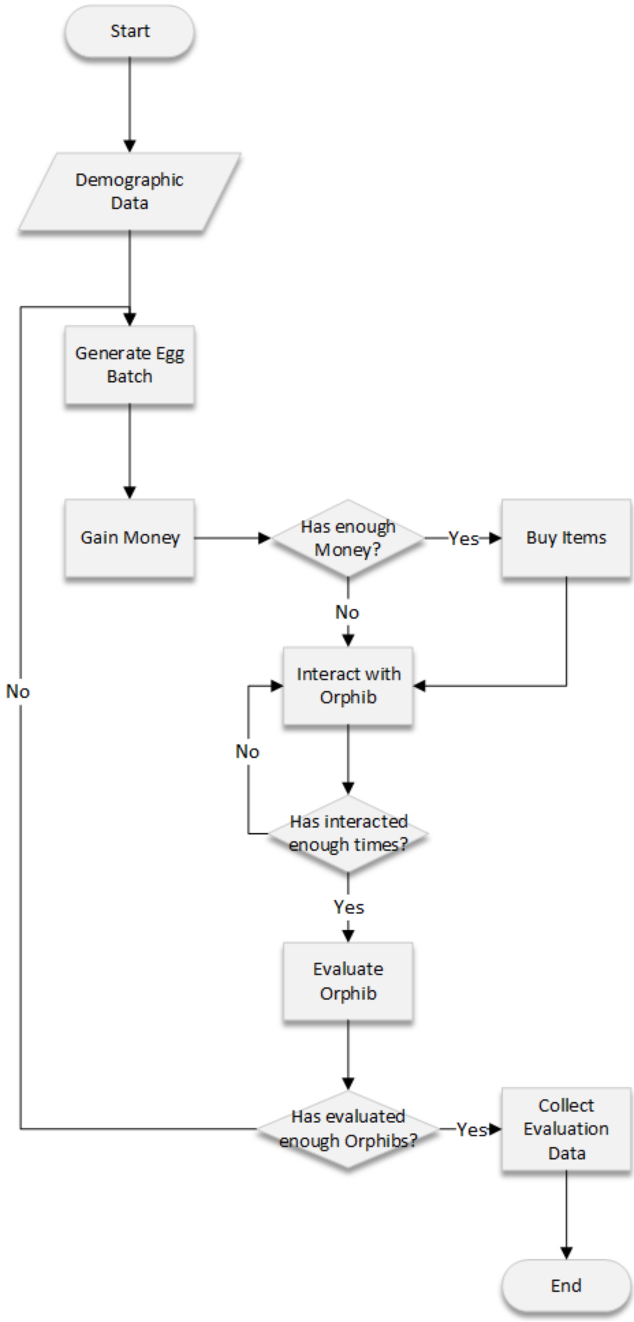


FIGURE D.1: Orphibs Evolvapalozza Game Flow diagram.

During this loop, the videogame can be in one of three different states:

- **Game State** - Most of the videogame is spent in this state. Here, players can scroll through the videogame's world where they can play with/evaluate Orphibs. In addition, they can generate a new batch of creatures replace the old ones. Transiting from this state to another one is done, with a click of a button.
- **Store State** - Transiting from the Game State, players can enter this state to buy new items. Meanwhile, the videogame is paused until players leave in the same manner they entered it: through means of a button.

- **Tutorial State** - Players can transit to this one, from the Game State, through special in-game triggers. During this state, the player is shown a tutorial/error message and the videogame is paused until the player transits back to the Game State in a manner similar to the Store State.

D.1.1 Game Mechanics

The videogame uses a top-down perspective, tilted at 45°, where players can scroll through the world. In addition, they can do four main actions:

1. They can interact with objects and creatures. This comes in the form of both neutral/friendly interactions such as grabbing items, pressing buttons or petting Orphibs and more aggressive ones (spanking Orphibs and tossing items).
2. They can buy items to play with the creature (e.g. stuffed toys, rolling pins, toilet paper, etc.). However, instead of directly choosing an item, players are awarded with a random one, from a pool of items. Unlocking tiers increases the number of available items. To buy items, players must have money.
3. They can play with creatures. When playing, players can use the items they bought, to pet creatures.
4. They can evaluate creatures. Before being able to acquire a new batch of creatures, players must evaluate every creature in a given batch. Evaluation is done using by rating the Orphib according to the CBS.

The videogame awards players money, so they can win items. They gain it every time a new egg batch is generated. This is done at a constant rate, increasing as the generations creatures increase.

A final mechanic is related to the Orphib's actual evolution. While players are told they are choosing creatures to evolve, evolution is actually deterministic and does not depend on the players' choices.

D.2 Controls

The videogame's main input is the mouse. This is meant to simplify the videogame, in order to make it accessible for a larger audience. The mouse is used to control a hand-style cursor. By moving the cursor to the world's edges, players can move around it. In addition, the mouse's main buttons are used for context sensitive actions: the left button is used for neutral/friendly actions such as grabbing items, pressing buttons or petting Orphibs while the right button allows for more aggressive actions (spanking Orphibs and tossing items).

D.3 Art Style

For the art style, we considered using what is colloquially known as 'low-poly art'. This is a visual aesthetic supposedly meant to mimic, and expand upon, the technically limited visuals achieved in the early 3D games of the late 90s, and early 00s. Since we established that believability is not necessarily related to realism, we considered this style, despite others, due to it's implementation simplicity. An example of the 'low-poly art' style is illustrated in Figure D.2.

FIGURE D.2: Screenshot of Feather (Samurai Punk, 2019), a videogame with 'low-poly art' visuals.

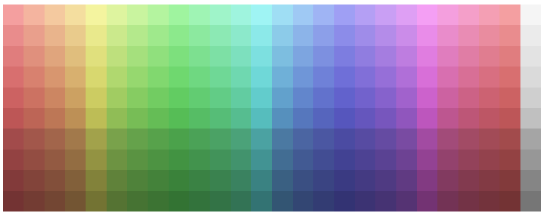


FIGURE D.3: The color palette used in Orphibs Evolvapalooza, with 260 colors. These encompass 26 hues with 10 different brightness levels.

Regarding the videogame's colors, we opted for pastel colors. On one hand, these type of colors are usually associated with the 'low-poly art' style. On the other hand, these colors are thought to invoke a "calming effect" (Burgett, 2015), something we defined for the videogame's original design intent. The videogame's color palette is illustrated in Figure D.3. It contains 26 hues, each with 10 brightness levels.

D.4 Sound Design

During the videogame's opening, there is a song playing until any input is given. This song is meant to set the videogame's mood (calm and playfulness).

The videogame's soundscape aimed to simulate a forest to avoid breaking the players' suspension of disbelief. In addition, sounds not directly connected to the creatures were kept at a comparatively lower volume to avoid distracting players.

Besides these, there are a few special events where the videogame signals the player:

- **Evaluating Creatures** - Whenever players evaluate a labeled Orphib, a sound plays to signal they can generate a new batch.
- **Being Awared Money** - Whenever players are awarded money, a sound plays.
- **Buying a Raffle** - Whenever players buy a raffle, a looping jingle plays until they confirm their action.

- **Winning an Item** - Whenever players are awarded a raffled item, a happy sound effect plays.
- **Items** - Items play sounds when they are dropped on the floor, as well as when they are used to pet a creature.

D.5 Game World

The videogame is set in a small world so players can focus on interacting with Orphibs. It is a grass field, with some rocks and logs, surrounded by trees, who act as barriers. In the middle of it, lies a water fountain. Standing near that, is an egg dispenser machine. In the world's top-right corner stands a small cave formation with a mound of dirt nearby. In addition, there is a bird creature roaming around the world.

D.5.1 Objects

The videogame's world has several man-made objects, the player can interact with.The first two are part of the world: the fountain and the egg dispenser. When the player interacts with the former, the fountain starts producing water, Orphibs can then drink while the latter creates a new batch of eggs. There are also mushrooms, the Orphibs are able to eat, and these grow in random spots along the videogame's world. These mushrooms, similarly to the aforementioned toys, can also be interacted by the player. In addition, there are several other objects, that act as toys, the Orphibs can play with. These objects are locked behind tiers. They are as follows:

- **Tier 0 (Food)** – Starts unlocked
  1. Carrot
  2. Leek
  3. Watermelon
  4. Banana
  5. Pineapple
- **Tier 1 (Toys)** – Unlocks on generation 2
  1. NotAGame BoyClone
  2. Spinning top
  3. Kendama
  4. Plushy Slinki
  5. Diabolo
- **Tier 2 (Sports)** – Unlocks on generation 3
  1. Football
  2. Rugby ball
  3. Curling disk
  4. Shuttlecock

5. Baseball bat

- **Tier 3 (Kitchen)** – Unlocks on generation 4
  1. Bowl
  2. Coffee Kettle
  3. Frying pan
  4. Bottle
  5. Rolling pin
- **Tier 4 (Bathroom)** – Unlocks on generation 5
  1. Toilet plunger
  2. Hair Brush
  3. Shampoo
  4. Toilet paper
  5. Toilet
- **Tier 5 (Gems)** – Unlocks on generation 6
  1. Diamond
  2. Ruby
  3. Emerald
  4. Sapphire
  5. Quartz Crystal

These items are illustrated in Figure D.4:

FIGURE D.4: A screenshot of the items in Orphibs Evolvapalooza.

It is worth noting that while a 'food' category is present in the previous list, these are toy lookalikes. The actual food Orphibs are able to eat are mushrooms. These grow in random spots throughout the forest and players can also interacted with them. Furthermore, there is one additional item: the clipboard. This special item can only be interacted with by the player and it allows to evaluate Orphibs.

D.5.2 Orphibs

The world is populated by Orphibs, herbivorous sexuated alien-like creatures. The player can play with 5 at a time. Their behavior is similar to that of most mammals, though they lay eggs (which only monotreme mammals do).

Guided by the CBS, Orphibs can perform the following actions:

- **Eat** - Orphibs walk up to a mushroom and eats it. This was taken from items 1 ("The creatures interact with surrounding environment") and 4 ("The creatures are able to find objects in the environment").
- **Play** - Orphibs can either fetch toys for the player or, follow him/her if he/she is grabbing a toy. This was based on items 1 ("The creatures interact with surrounding environment") and 3 ("The creatures direct their movement towards/away from targets").
- **Hunt** - Orphibs can stalk the bird present in the videogame's world and, when close enough, bite it. This was inspired by items 1 ("The creatures interact with surrounding environment") and 3 ("The creatures direct their movement towards/away from targets").
- **Excrete** - Orphibs ceremoniously expel organic materials and bury them. This was based on items 1 ("The creatures interact with surrounding environment"), 5 ("The creatures are able to expel organic waste") and 6 ("There are signs of previous expelled organic waste, such as feces, urine, goo, etc.").
- **Reproduce** - Unlike the other actions, this is not directly seen but rather, hinted at. When Orphibs want to reproduce, they move towards the cave and it closes. During this time, hearts are shown around the cave. After reproducing, one of the orphibs becomes pregnant and, eventually, lays an egg. These were based on items 13 ("The creatures engage in reproductive acts") and 14 ("There are signs of previous reproductive acts, such as eggs, cubs, pregnancy, etc.").
- **Learn** - There are three moments where Orphics can learn. First, whenever a player pets/spanks an orphib after an action, this action is reinforced/inhibited respectively. Second, when Orphibs observe the player performing some actions (interacting with the fountain), they learn it and can do it themselves. Third, Orphibs can learn to use the same actions in different contexts: after being able to fetch toys, Orphibs learn to hunt using the same basics. This was inspired by items 15 ("The creatures learn from past events"), 16 ("The creatures use past experience to overcome novel situations") and 22 ("The creatures can imitate others to develop their own behaviors").
- **Drink** - If the fountain has water on it, Orphibs can move towards it and drink its water. The reason behind this, was to account for item 10 ("The creatures react to environmental triggers").
- **Interact with the Fountain** - Orphibs can interact with the fountain, to make it pour water, akin to the player, for themselves, or other Orphibs. This was based on items 1 ("The creatures interact with surrounding environment") and

11 ("The creatures can perform tasks by coordinating with other creatures, or humans").

- **Sleep** - Orphibs move towards the mound of dirt and sleep on it. This was also to account for item 4 ("The creatures are able to find objects in the environment").
- **Talk** - Orphibs move towards other Orphibs and bark at each other. This was so that "The creatures audio-visually communicate with other creatures, or humans" (item 12).
- **Grow** - As time passes, Orphibs grow in size and change the color of their bodies, toward a greyish tint, to indicate they are maturing. Additionally, their voices' pitch lowers and they predisposition for certain actions reduces. This was based on items 17 ("The creatures change the way they look with age"), 18 ("The creatures change the way they sound with age") and 19 ("The creatures change the way they behave with age").
- **Express emotions** - Orphibs have a specific body language to indicate emotions. First, they are able to pulse a glowing light to indicate they are expressing emotions. The light's colors are warm upon positive emotions whilst cold upon negative ones. Additionally, Orphibs have also special poses for this extent. Additionally, their ears and tails act differently. These elements was based on item 21 ("The creature shows signs of emotional reaction to objects, events and other creatures").

Visually, we decided to design Orphibs as fictional creatures though, with some familiar elements. This still falls within our definition ("Zoomorphic entities, inspired by (contemporary or otherwise) living or fictional beings (...)". The main reason behind this, is to allow players to derive some meaning from the familiar aspects while, at the same time, make the creature different enough from any existing species, as to avoid any expectations associated to them. As a result, Orphibs can be seen as some form of hybrid between a kangaroo and a dinosaur, as depicted in Figure 6.3. In addition, these animal inspirations, and the orphibs' color palette, were also selected to account for item 23 ("The creatures' bodies seem adapted to their habitat"), due to the game's forest environment.

FIGURE D.5: A rendering of an Orphib.

D.6 Orphibs Evolvapalooza User Interface

There are two types of graphical user interfaces in Orphibs Evolvapalooza. The first one, used in the videogame's main screen, consists in superimposed navigation buttons. The other, used every other context, encompasses pop-up windows. The interface navigation diagram is illustrated in Figure D.6.

The videogame starts with a main menu comprising its logo. After pressing any key, it transits to the main screen. This is followed by a pop-up window, prompting players to enter demographic data. When they confirm their date, the videogame transits to the main interface, comprised of superimposed buttons over the main screen. These include two buttons on the lower right corner, and one on the upper right corner: one to recall the clipboard to the player's position, one to open the item store and one to teleport the player to the egg machine.

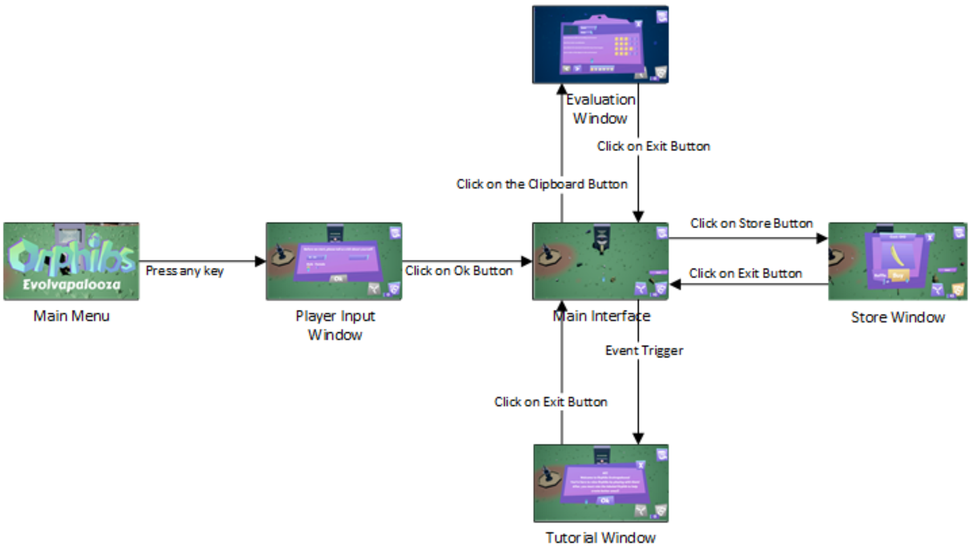


FIGURE D.6: Orphibs Evolvapalooza's navigation diagram.

When orphibs are ready to be evaluated, a clipboard image appears on top of them. Clicking on this clipboard causes a popup message to show with the scale's items and a star mapping to a 5-point Likert rank: 1 (Not at all), 2 (Slightly), 3 (Somewhat), 4 (Moderately) and 5 (To a great extent). This pop-up window disappears after pressing the exit button.

Similar to the other windows, the item store is superimposed upon the main screen. It has a circular slideshow showing a subset of all of the available items, a text with the cost of buying the next raffle, how many raffles players can buy with their current currency and a button stating 'Buy' (which is disabled if players cannot buy a raffle). Additionally, an exit button is present on the upper right corner of the menu. When pressing the buy button, the slide show begins to slowdown until it completely stops. At this point, players receive the item present on the image frame.

The final type of window players can navigate to, is the tutorial window. Unlike the other windows however, this one pops-up during special events, or after particular triggers. It is used to explain some aspects of the videogame. Every time a tutorial pops up, the videogame is paused and players must acknowledge it, with the press of a button, to resume normal play. The tutorials teach the following:

- What to do with creatures
- How to evaluate
- How to win evaluation points
- How to win money

- How to buy items
- How to get new creatures
- What to do with points
- Explain the videogame's possible interactions

Bibliography

1Games (2017). *Life Of Black Tiger*. [Online].
Adkins-Regan, E. (2005). *Hormones and Animal Social Behavior*. Princeton University Press.
A.I. Design (1980). *Rogue*. [Floppy-Disk].
Arrabales, Raúl, Agapito Ledezma, and Araceli Sanchis (2011). “ConsScale FPS: Cognitive Integration for Improved Believability in Computer Game Bots”. In: *Believable Bots: Can Computers Play Like People?* Springer, pp. 193–214.
Barreto, Nuno, Amílcar Cardoso, and Licínio Roque (2014). “Computational Creativity in Procedural Content Generation: A State of the Art Survey”. In: *Conferência de Ciências e Artes dos Videojogos*. Barcelos, Portugal.
Barreto, Nuno, Rui Craveirinha, and Licinio Roque (2017). “Designing a Creature Believability Scale for Videogames”. In: *Proceedings of the 2017 International Conference on Entertainment Computing 2017*. Tsukuba City, Japan: Springer.
– (2018). “Validating the Creature Believability Scale for Videogames”. In: *Entertainment Computing – ICEC 2018*. Poznań, Poland: Springer.
Barreto, Nuno, Luís Macedo, and Licinio Roque (2014). “MultiAgent System Architecture in Orphibs II”. In: *Proceedings of the 14th International Conference on the Synthesis and Simulation of Living Systems (ALIFE 14)*. New York City, USA: MIT Press, pp. 588–595.
Barreto, Nuno and Licinio Roque (2014). “A Survey of Procedural Content Generation tools in Video Game Creature Design”. In: *Proceedings of the Second Conference on Computation Communication Aesthetics and X*. Porto, Portugal.
Barreto, Nuno et al. (2015). “Incorporation of Emotions in the Orphibs’ Agent Architecture”. In: *Proceedings of the 13th European Conference on Artificial Life (ECAL 2015)*. York, England: MIT Press, p. 311.
Bates, Joseph (1994). “The role of emotion in believable agents”. In: *Communications of the ACM* 37 (7), pp. 122–125.
B.B. Studio (2017). *Digimon World: Next Order*. [Blu-Ray].
Bethesda Game Studios (2011). *The Elder Scrolls V: Skyrim*. [DVD-ROM].
– (2015). *Fallout 4*. [DVD-ROM].
Binary Systems (1986). *Starflight*. [Floppy-Disk].
Blizzard Entertainment (2015). *StarCraft II: Legacy of the Void*. [DVD-ROM].
Blizzard North (1996). *Diablo*. [CD-Rom].
Boden, Margaret A. (2009). “Computer Models of Creativity”. In: *AI Magazine* 30.3, pp. 23–35.
Borges, Jorge Luis and Margarita Guerrero (2002). *Book of Imaginary Beings*. Buenos Aires: Vintage Classics.
Braitenberg, Valentino (1984). *Vehicles: Experiments in Synthetic Psychology*. MIT Press.
Brown, Sorrel (2010). *Likert Scale Examples for Surveys*. URL: https://www.extension.iastate.edu/documents/anr/likertscaleexamplesforsurveys.pdf.

Bujang, Mohamad Adam (June 2017). “A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review”. In: *Archives of Orofacial Sciences* 12, pp. 1–11.

Burgett, Shoshana (2015). *Spring Color Update: The Psychology of Pastels*. URL: https://www.xrite.com/blog/spring-color-update-psychology-of-pastels.

Cambridge Dictionary (2018). Cambridge University Press.

Capcom (2008). *Monster Hunter Freedom Unite*. [UMD].

CD Projekt RED (2015). *The Witcher 3: Wild Hunt*. [DVD-ROM].

Cenydd, Llyr ap and Bill Teahan (Jan. 2013). “An embodied approach to arthropod animation”. In: *Computer Animation and Virtual Worlds* 24. DOI: 10.1002/cav.1436.

Cerezo, Eva, Alfredo Pina, and Francisco Serón (May 1999). “Motion and behaviour modelling: State of art and new trends”. In: *The Visual Computer* 15, pp. 124–146. DOI: 10.1007/s003710050167.

Cicchetti, Domenic (Dec. 1994). “Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instrument in Psychology”. In: *Psychological Assessment* 6, pp. 284–290. DOI: 10.1037/1040-3590.6.4.284.

Coleridge, Samuel Taylor (1817). *Biographia Literaria*.

Colton, Simon, Jacob Goodwin, and Tony Veale (Dec. 2018). “Full-FACE Poetry Generation”. In: *Proceedings of the Third International Conference on Computational Creativity*.

Colton, Simon, R. López de Mántaras, and O. Stock (2013). “Computational Creativity: Coming of Age”. In: *AI Magazine* 30.3, p. 11.

Cook, M, S Colton, and Alison Pease (Jan. 2012). “Aesthetic considerations for automated platformer design”. In: pp. 124–129.

Cook, Michael and Simon Colton (2012). “ANGELINA – Coevolution in Automated Game Design”. In: *Proceedings of the Third International Conference on Computational Creativity*.

Cope, David (2006). *Computer Models of Musical Creativity*. MIT Press.

Coros, Stelian, Philippe Beaudoin, and Michiel van de Panne (2009). “Robust Task-based Control Policies for Physics-based Characters”. In: *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 28.5, Article 170.

Coros, Stelian et al. (2011). “Locomotion Skills for Simulated Quadrupeds”. In: *ACM SIGGRAPH 2011 Papers*. SIGGRAPH ’11. Vancouver, British Columbia, Canada: ACM, 59:1–59:12. ISBN: 978-1-4503-0943-1. DOI: 10.1145/1964921.1964954. URL: http://doi.acm.org/10.1145/1964921.1964954.

Craveirinha, Rui (2018). “The Authorial Game Evolution Tool for Player Experience Design”. PhD thesis. Coimbra, Portugal: Universidade de Coimbra.

Craveirinha, Rui, Nuno Barreto, and Licinio Roque (2016). “Towards a Taxonomy for the Clarification of PCG Actors’ Roles”. In: *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY ’16. Austin, Texas, USA: ACM, pp. 244–253. ISBN: 978-1-4503-4456-2. DOI: 10.1145/2967934.2968086. URL: http://doi.acm.org/10.1145/2967934.2968086.

Crawford, Chris (1987). “Process Intensity”. In: *The Journal of Computer Game Design* 1.5.

Creature Labs (1996). *Creatures*. [CD-ROM].

Creswell, John W. (2013). *Research Design: Qualitative, Quantitative, & Mixed Methods Approaches*. SAGE Publications, Inc.

Csikszentmihalyi, Mihaly (1997). *Creativity : Flow and the Psychology of Discovery and Invention*. HarperPerennial.

D. Braben and I. Bell (1984). *Elite*. [Floppy-Disk].

Dahlskog, Steve and Julian Togelius (2014). "Procedural Content Generation Using Patterns as Objectives". In: *EvoApplications 2014: Applications of Evolutionary Computation*. Springer, pp. 325–336.

Dane, Patrick (2014). "'Grand Theft Auto 5' Map Compared To GTA IV, San Andreas & Real Cities". In: *GameRant*. URL: https://gamerant.com/grand-theft-auto-5-map-comparisons/.

Darwin, Charles (1889). *The Expression of the Emotions in Man and Animals*. John Murray.

David OReilly (2017). *Everything*. [Download].

Dawkins, Marian (2000). "Animal Minds and Animal Emotions". In: *American Zoologist* 40.6, pp. 883–888.

Dennett, Daniel (2009). "Intentional Systems Theory". In: *The Oxford Handbook of Philosophy of Mind*. Oxford University Press, pp. 1–22.

Dormans, J. and S. Leijnen (2013). "Combinatorial and exploratory creativity in procedural content generation". In: *Proceedings of the fourth workshop on Procedural Content Generation for Games at the Foundations of Digital Games Conference*. Chania, Greece: Society for the Advancement of the Science of Digital Games, pp. 1–4.

Duarte, Luis and Luís Carriço (Apr. 2013). "The cake can be a lie: placebos as persuasive videogame elements". In: DOI: 10.1145/2468356.2468555.

Edward, Schneider, Wang Yifan, and Yang Shanshan (2007). "Exploring the Uncanny Valley with Japanese Video Game Characters". In: *DiGRA &#3907 - Proceedings of the 2007 DiGRA International Conference: Situated Play*. The University of Tokyo. ISBN: ISSN 2342-9666. URL: http://www.digra.org/wp-content/uploads/digital-library/07312.11004.pdf.

Eiben, A. E. and J. E. Smith (2007). *Introduction to Evolutionary Computing*. Springer.

Eladhari, Mirjam P. et al. (2011). *AI-Based Game Design: Enabling New Playable Experiences*. Tech. rep.

Electronic Arts (2009). *SimAnimals*. [DVD-ROM].

Entertainment, Naked Sky (2006). *RoboBlitz*. [Download].

Epic Games (2012). *Unreal Engine 4*.

Epstein, Joshua M. and Robert L. Axtell (1996). *Growing Artificial Societies*. The MIT Press.

Estes, Richard Despard (2012). "Hyenas and Aardwolf: Family Hyaenidae". In: *The Behavior Guide to African Mammals including Hoofed Mammals, Carnivores, Primates*. University of California Press, p. 327.

Farbrausch (2004). *.kkrieger*. [Download].

Fauconnier, Gilles and Mark Turner (1998). "Conceptual integration networks". In: *Cognitive Science* 22.2, pp. 133 –187. ISSN: 0364-0213. DOI: https://doi.org/10.1016/S0364-0213(99)80038-X. URL: http://www.sciencedirect.com/science/article/pii/S036402139980038X.

Fogg, B. and H. Tseng (1999). "The Elements of Computer Credibility". In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*.

Frank, A., A. Stern, and B. Resner (1997). "Socially Intelligent Virtual Petz". In: *Socially Intelligent Agents AAAI Fall Symposium*, pp. 43–45.

Frontier Developments (2010). *Kinectimals*. [DVD-ROM].

Games, Emote (2005). *theHunter*. [Online].

Geijtenbeek, Thomas, Michiel van de Panne, and A. Frank van der Stappen (2013). "Flexible Muscle-Based Locomotion for Bipedal Creatures". In: *ACM Transactions on Graphics* 32.6.

genDESIGN and SIE Japan Studio (2016). *The Last Guardian*. [Blu-Ray].

Giant Squid Studios (2016). *Abzû*. [Blu-Ray].

Gigazine (2017). "Interview with Fumito Ueda, "The Last Guardian" Game Director and Designer". In: *Gigazine*. URL: http://en.gigazine.net/news/20170126-the-last-guardian-ueda-fumito-interview/.

Gosling, Samuel D. (2008). "Personality in Non-human Animals". In: *Social and Personality Psychology Compass* 2 (2), pp. 985–1001.

Grand, Stephen and Dave Cliff (Mar. 1998). "Creatures: Entertainment Software Agents with Artificial Life". In: *Autonomous Agents and Multi-Agent Systems* 1, pp. 39–57. DOI: 10.1023/A:1010042522104.

Grand, Stephen, Dave Cliff, and Anil Malhotra (1997). "Creatures: Artificial Life Autonomous Software Agents for Home Entertainment". In: *Proceedings of the First International Conference on Autonomous Agents*. AGENTS '97. Marina del Rey, California, USA: ACM, pp. 22–29. ISBN: 0-89791-877-0. DOI: 10.1145/267658.267663. URL: http://doi.acm.org/10.1145/267658.267663.

GSC Game World (2010). *S.T.A.L.K.E.R.: Call of Pripryat*. [DVD-ROM].

Guerrilla Games (2017). *Horizon Zero Dawn*. [Blu-Ray].

Guo, Xuekun et al. (2014). "Creature grammar for creative modeling of 3D monsters". In: *Graphical Models* 76.5. Geometric Modeling and Processing 2014, pp. 376 –389. ISSN: 1524-0703. DOI: https://doi.org/10.1016/j.gmod.2014.03.019. URL: http://www.sciencedirect.com/science/article/pii/S1524070314000265.

Hangar 13 (2016). *Mafia III*. [Blu-Ray].

Hecker, Chris et al. (2008). "Real-time motion retargeting to highly varied user-created morphologies". In: *ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2008* 27.3, Article No. 3.

Hello Games (2016). *No Man's Sky*. [Blu-Ray].

Hendrikx, Mark et al. (2013). "Procedural Content Generation for Games: A Survey". In: *ACM Transactions on Multimedia Computing, Communications, and Applications* 9.1, Article No. 1.

Hingston, P. (2011). *Believable Bots: Can Computers Play Like People?* Springer.

Hinkin, Timothy R., J. Bruce Tracey, and Cathy A. Enz (1997). "Scale Construction: Developing Reliable and Valid Measurement Instruments". In: *Journal of Hospitality & Tourism Research* 21.1, pp. 100–120. DOI: 10.1177/109634809702100108. URL: https://doi.org/10.1177/109634809702100108.

Horswill, I. D. (2009). "Lightweight Procedural Animation With Believable Physical Interactions". In: *IEEE Transactions on Computational Intelligence and AI in Games* 1.1, pp. 39–49. ISSN: 1943-068X. DOI: 10.1109/TCIAIG.2009.2019631.

Hudlicka, E. and J. Broekens (2009). "Foundations for modelling emotions in game characters: Modelling emotion effects on cognition". In: *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1–6. DOI: 10.1109/ACII.2009.5349473.

Hudlicka, Eva (Jan. 2008). "Affective computing for game design". In: *4th International North-American Conference on Intelligent Games and Simulation, Game-On 'NA 2008*, pp. 5–12.

Hudson, Jon (2013). "Creature Generation using Genetic Algorithms and Auto-Rigging". MA thesis. Bournemouth University.

Hélie, S. and R. Sun (2010). "Incubation, insight, and creative problem solving: A unified theory and a connectionist model". In: *Psychological Review* 117.3, pp. 994–1024.

Ilčík, Martin et al. (2010). "Procedural Skeletons: Kinematic Extensions to CGA-Shape Grammars". In: *Proceedings of the Spring Conference on Computer Graphics 2010*. Comenius University, Bratislava, pp. 177–184. ISBN: 978-80-223-2644-5. URL: https://www.cg.tuwien.ac.at/research/publications/2010/ilcik-2010-ps/.

Interactive Data Visualization, Inc. (2013). *SpeedTree*.

Kadhil, N. (2005). "Characteristics of living organisms". In: *Biology*. Cambridge University Press, pp. 1–10.

Khaled, Rilla, Mark J. Nelson, and Pippin Barr (2013). "Design metaphors for procedural content generation in games". In: *Proceedings of the 2013 ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 1509–1518.

Kidd, C. D. and C. Breazeal (2004). "Effect of a robot on user perceptions". In: *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*. Vol. 4, pp. 3559–3564. DOI: 10.1109/IROS.2004.1389967.

Kidd, Cory D. (2003). "Sociable Robots : The Role of Presence and Task in Human-Robot Interaction". MA thesis. Massachusetts Institute of Technology.

Kittur, Aniket, Ed H. Chi, and Bongwon Suh (2008). "Crowdsourcing User Studies with Mechanical Turk". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. Florence, Italy: ACM, pp. 453–456. ISBN: 978-1-60558-011-1.

Kojima Productions (2015). *Metal Gear Solid V: The Phantom Pain*. [Blu-Ray].

Krueger, Richard A. (2002). *Designing and Conducting Focus Group Interviews*. URL: https://www.eiu.edu/ihec/Krueger-FocusGroupInterviews.pdf.

Lee Graham (2007). *3D Virtual Creatures Evolution*. [Download].

Leutner, Franziska et al. (2017). "From Likert scales to images: Validating a novel creativity measure with image based response scales". In: *Personality and Individual Differences* 106, pp. 36 –40. ISSN: 0191-8869. DOI: https://doi.org/10.1016/j.paid.2016.10.007. URL: http://www.sciencedirect.com/science/article/pii/S0191886916310352.

Liapis, Antonios, Georgios Yannakakis, and Julian Togelius (Sept. 2012). "Adapting Models of Visual Aesthetics for Personalized Content Creation". In: *Computational Intelligence and AI in Games, IEEE Transactions on* 4. DOI: 10.1109/TCIAIG.2012.2192438.

Liapis, Antonios, Georgios N. Yannakakis, and Julian Togelius (2013a). "Enhancements to Constrained Novelty Search: Two-population Novelty Search for Generating Game Content". In: *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation*. GECCO '13. Amsterdam, The Netherlands: ACM, pp. 343–350. ISBN: 978-1-4503-1963-8. DOI: 10.1145/2463372.2463416. URL: http://doi.acm.org/10.1145/2463372.2463416.

– (2013b). "Sentient Sketchbook: Computer-Aided Game Level Authoring". In: *Proceedings of the 8th Conference on the Foundations of Digital Games*, pp. 213–220.

Lim, Mei Yii et al. (2009). "Intelligent NPCs for Educational Role Play Game". In: *Agents for Games and Simulations: Trends in Techniques, Concepts and Design*. Ed. by Barry Silverman Frank Dignum Jeff Bradshaw and Willem van Doesburg. Springer, pp. 107–118.

Lionhead Studios (2001). *Black & White*. [CD-ROM].

– (2008). *Fable II*. [DVD-ROM].

Lombard, Matthew and Theresa Bolmarcich (Sept. 1997). "At the Heart of It All: The Concept of Presence". In: *Journal of Computer-Mediated Communication* 3. DOI: 10.1111/j.1083-6101.1997.tb00072.x.

Lombard, Matthew and Theresa Ditton (1997). "At the Heart of It All: The Concept of Presence". In: *Journal of Computer-Mediated Communication* 3.2, pp. 0–39.

– (2000). "MEASURING PRESENCE: A LITERATURE-BASED APPROACH TO THE DEVELOPMENT OF A STANDARDIZED PAPER-AND-PENCIL INSTRUMENT". In:

Lombard, Matthew et al. (2000). *Measuring presence: A literature-based approach to the development of a standardized paper-and-pencil instrument*.
Lorenz, K. Z. (2010). *The Foundations of Ethology*. Springer.
Loyall, A. Bryan (1997). "Believable Agents: Building Interactive Personalities". PhD thesis. Carnegie Mellon University Pittsburgh.
Lubart, Todd (2005). "How can computers be partners in the creative process: classification and commentary on the special issue". In: *International Journal of Human-Computer Studies - Special issue: Computer support for creativity* 63.4-5, pp. 365–369.
Lucas, Simon M. et al. (2012). *Artificial and Computational Intelligence in Games*. Technical Report 5. Schloss Dagstuhl, pp. 43–70.
Machado, Fernando (2006). "Inteligência Artificial e Arte". PhD thesis. University of Coimbra.
Magerko, Brian (2007). "Measuring Dramatic Believability". In: *AAAI fall symposium on intelligent narrative technologies*.
Martin, Jess (Jan. 2005). "Algorithmic Beauty of Buildings Methods for Procedural Building Generation". In: *Honors Theses*.
Martinho, Carlos, Pedro Santos, and Rui Prada (2014). *Design e Desenvolvimento de Jogos*. Lisbon: FCA - Editora de Informática, Lda.
Mason, Winter and Siddharth Suri (2012). "Conducting behavioral research on Amazon's Mechanical Turk". In: *Behavior Research Methods* 44.1, pp. 1–23. ISSN: 1554-3528.
Mateas, Michael (July 1997). *An Oz-Centric Review of Interactive Drama and Believable Agents*. Tech. rep. School of Computer Science Carnegie Mellon University.
– (2003). "Expressive AI: Games and Artificial Intelligence". In: *Proceedings of the 2003 DiGRA International Conference: Level Up*. Utrecht, Netherlands, pp. 611–618.
Mateas, Michael and Andrew Stern (Jan. 2003). "A Behavior Language: Joint Action and Behavioral Idioms". In: DOI: 10.1007/978-3-662-08373-4_7.
– (2005). "Procedural authorship: A case-study of the interactive drama façade". In: *In Digital Arts and Culture (DAC*.
Mathias, M. et al. (2011). "Procedural 3D Building Reconstruction Using Shape Grammars and Detectors". In: *2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, pp. 304–311. DOI: 10.1109/3DIMPVT.2011.45.
Maxis (2000). *The Sims*. [CD-ROM].
– (2005). *The Sims 2*. [DVD-ROM].
– (2008). *Spore*. [DVD-ROM].
– (2009). *The Sims 3*. [DVD-ROM].
McCloud, S. (1994). *Understanding Comics: The Invisible Art*. Morrow.
McCormack, J. and M. d'Inverno (2012). "Computers and Creativity: The Road Ahead". In: *Computers and Creativity*. Ed. by J. McCormack and M. d'Inverno. Berlin, Germany: Springer, pp. 421–424.
McLeod, S. A. (2009). *Short term memory*. URL: https://www.simplypsychology.org/short-term-memory.html.
Mcmahan, Alison (Jan. 2003). "Immersion, engagement, and presence: A method for analyzing 3-D video games". In: *The Video Game Theory Reader*, pp. 67–86.
Menzel, Randolf and Julia Fischer (2011). *Animal Thinking - Contemporary Issues in Comparative Cognition*. Cambridge: MIT Press.
Merrick, Kathryn E. et al. (2013). "A shape grammar approach to computational creativity and procedural content generation in massively multiplayer online role playing games". In: *Entertainment Computing* 4.2, pp. 115 –130. DOI: https://doi.org/10.1016/j.entcom.2012.09.006. URL: http://www.sciencedirect.com/science/article/pii/S1875952112000213.

Michael Young, R et al. (Jan. 2004). “An architecture for integrating plan-based behavior generation with interactive game environments”. In: *Journal of Game Development* 1, pp. 51–70.
Millenium Interactive (1996). *Creatures*.
Millington, Ian and John Funge (2009). *Artificial Intelligence for Games*. Burlington: Morgan Kaufmann Publishers.
Mojang (2011). *Minecraft*. [DVD-Rom].
Multon, Franck et al. (1999). “Computer animation of human walking: a survey”. In: *The Journal of Visualization and Computer Animation* 10.1, pp. 39–54. DOI: 10.1002/(SICI)1099-1778(199901/03)10:1<39::AID-VIS195>3.0.CO;2-2.
Muncy, Julie (2016). “In the new game Everything, you can be, well, everything”. In: *Wired*. URL: https://www.wired.com/2016/03/everything-announcement/.
Nareyek, Alexander (2007). “Game AI Is Dead. Long Live Game AI!” In: *IEEE Intelligent Systems* 22.1, pp. 9–11.
NaturalMotion (2008). *Euphoria*.
Nava, Matt (2017). *Creating the Art of ‘ABZU’*. Video. URL: https://www.gdcvault.com/play/1024409/Creating-the-Art-of-ABZU.
Nintendo EAD (2001). *Pikmin*. [Mini Disc].
– (2005). *Nintendogs*. [Flash Card].
Ochs, Magalie, Nicolas Sabouret, and Vincent Corruble (Jan. 2010). “Simulation of the Dynamics of Nonplayer Characters’ Emotions and Social Relations in Games”. In: *Computational Intelligence and AI in Games, IEEE Transactions on* 1, pp. 281 –297. DOI: 10.1109/TCIAIG.2009.2036247.
Parenthöen, Marc, Jacques Tisseau, and Thierry Morineau (2002). “Believable Decision for Virtual Actors”. In: *2002 IEEE International Conference on Systems, Man and Cybernetics*.
Pereira, Francisco and Amílcar Cardoso (June 2000). “Dr. Divago: Searching for new ideas in a multi-domain environment”. In:
PF Magic (1995). *Petz*. [CD-Rom].
Piranha Bytes (2009). *Risen*. [DVD-ROM].
Poels, Y et al. (2012). “Are you a gamer?:A qualitative study on the parameters for categorizing casual and hardcore gamers”. In: *IADIS International Journal on WWW/Internet* 10, pp. 1–16.
Procedural Arts (2005). *Façade*. [Download].
Propaganda Games (2008). *Turok*. [Blu-Ray].
Raúl, A., A. Ledezma, and A. Sanchis (2010). “ConsScale: A Pragmatic Scale for Measuring the Level of Consciousness in Artificial Agents”. In: *Journal of Consciousness Studies* 17 (3-4), pp. 131–164.
Reeves, Byron and Clifford Nass (1998). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. The Center for the Study of Language and Information Publications.
Reisenzein, Rainer (Mar. 2009). “Emotions As Metarepresentational States of Mind: Naturalizing the Belief-desire Theory of Emotion”. In: *Cogn. Syst. Res.* 10.1, pp. 6–20. ISSN: 1389-0417. DOI: 10.1016/j.cogsys.2008.03.001. URL: http://dx.doi.org/10.1016/j.cogsys.2008.03.001.
– (Jan. 2012a). “Extending the computational belief-desire theory of emotions to fantasy emotions”. In: *Proceedings of the 11th International Conference on Cognitive Modeling, ICCM 2012*, pp. 313–314.
– (2012b). “What is an emotion in the Belief-Desire Theory of emotion?” In: *The goals of cognition: Essays in honor of Cristiano Castelfranchi*. Ed. by F. Paglieri et al. London: College Publications, pp. 181–211.

Reynaud, Q., J.-Y. Donnart, and V. Corruble (2014). “Evaluating the believability of virtual agents with anticipatory abilities”. In: *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*.

Rhino Studios (2008). *Afrika*. [Blu-Ray].

Ribeiro, Paulo et al. (2003). “A Model For Creativity In Creature Generation”. In: *Proceedings of the Fourth International Conference on Intelligent Games and Simulation*, pp. 175–179.

Riedl, Mark O. and R. Michael Young (2005). “Lecture Notes in Computer Science”. In: ed. by Themis Panayiotopoulos et al. London, UK, UK: Springer. Chap. An Objective Character Believability Evaluation Procedure for Multi-agent Story Generation Systems, pp. 278–291. ISBN: 3-540-28738-8. DOI: doi:10.1007/11550617_24. URL: http://dx.doi.org/10.1007/11550617_24.

Rockstar North (2013). *Grand Theft Auto V*. [Blu-Ray].

Rockstar Studios (2018). *Red Dead Redemption 2*. [Blu-Ray].

Roettl, Johanna and Ralf Terlutter (July 2018). “The same video game in 2D, 3D or virtual reality – How does technology impact game evaluation and brand placements?” In: *PLOS ONE* 13.7, pp. 1–24. DOI: 10.1371/journal.pone.0200724. URL: https://doi.org/10.1371/journal.pone.0200724.

Roque, Licinio (2010). “Early Game Design Rehearsal with Paper Prototyping”. In: *IX Simpósio Brasileiro de Games e Entretenimento Digital*.

Rosenkind, M., G. Winstanley, and A. Blake (2008). “Creating Believable Characters using a Synthetic Psychology Approach”. In: *Proceedings of the CGAMES’2008, 13th international Conference on Computer Games: AI, Animation, Mobile, Educational and Serious*.

Rosenkind, Micah, Graham Winstanley, and Andrew Blake (2012). “Adapting Bottom-up, Emergent Behaviour for Character-Based AI in Games”. In: *Research and Development in Intelligent Systems XXIX*. Ed. by Max Bramer and Miltos Petridis. Springer, pp. 333–346.

Samurai Punk (2019). *Feather*. [Download].

Santos, S. and L. Roque (2010). “Ensaio de Reescrita de Comportamentos em Videojogos com Base no Ajuste e Computação de Modelos de Petri Net”. In: *IX Simpósio Brasileiro de Games e Entretenimento Digital*. Florianópolis, Brasil.

SCE London Studio and Playlogic Game Factory (2009). *EyePet*. [Blu-Ray].

Shaker, Noor, Julian Togelius, and Mark J. Nelson (2016). *Procedural Content Generation in Games: A Textbook and an Overview of Current Research*. Springer.

Shapiro, A., F. Pighin, and P. Faloutsos (2003). “Hybrid control for interactive character animation”. In: *11th Pacific Conference onComputer Graphics and Applications, 2003. Proceedings.* Pp. 455–461. DOI: 10.1109/PCCGA.2003.1238294.

Shapiro, Ari et al. (2007). “A Dynamic Controller Toolkit”. In: *Proceedings of the 2007 ACM SIGGRAPH Symposium on Video Games*. Sandbox ’07. San Diego, California: ACM, pp. 15–20. ISBN: 978-1-59593-749-0. DOI: 10.1145/1274940.1274945. URL: http://doi.acm.org/10.1145/1274940.1274945.

Shlyakhter, I. et al. (2001). “Reconstructing 3D tree models from instrumented photographs”. In: *IEEE Computer Graphics and Applications* 21.3, pp. 53–61. ISSN: 0272-1716. DOI: 10.1109/38.920627.

Sia Ding Shen (2017). *Macho Cat : Pat the Virtual Cat*. [Download].

Side Effects Software (2015). *Houdini*.

Simmons, Maryann, Jane Wilhelms, and Allen Van Gelder (2002). “Model-based Reconstruction for Creature Animation”. In: *Proceedings of the 2002 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. SCA ’02. San Antonio, Texas: ACM, pp. 139–146. ISBN: 1-58113-573-4. DOI: 10.1145/545261.545284. URL: http://doi.acm.org/10.1145/545261.545284.

Sims, Karl (1994). "Evolving 3D Morphology and Behavior by Competition". In: *Artificial Life* 1.4, pp. 353–372. DOI: 10.1162/artl.1994.1.4.353. eprint: https://doi.org/10.1162/artl.1994.1.4.353. URL: https://doi.org/10.1162/artl.1994.1.4.353.

Skrba, Ljiljana et al. (Sept. 2009). "Animating Quadrupeds: Methods and Applications". In: *Computer Graphics Forum* 28. DOI: 10.1111/j.1467-8659.2008.01312.x.

Smelik, R. M. et al. (2009). "A Survey of Procedural Methods for Terrain Modelling". In: *Proceedings of the CASA workshop on 3D advanced media in gaming and simulation (3AMIGAS)*.

Smith, Gillian, Jim Whitehead, and Michael Mateas (2010). "Tanagra: a mixed-initiative level design tool". In: *FDG '10: Proceedings of the Fifth International Conference on the Foundations of Digital Games*. New York, NY, USA: ACM, pp. 209–216.

– (Oct. 2011). "Tanagra: Reactive Planning and Constraint Solving for Mixed-Initiative Level Design". In: *Computational Intelligence and AI in Games, IEEE Transactions on* 3, pp. 201 –215. DOI: 10.1109/TCIAIG.2011.2159716.

Sohrabi-Shiraz, Ariane (2018). "Red Dead 2 map size comparison: How big is Red Dead Redemption 2 map compared to GTA 5?" In: *DailyStar*. URL: https://www.dailystar.co.uk/tech/guides/740674/Red-Dead-2-map-size-comparison-Red-Dead-Redemption-2-map-GTA-5.

Sonic Team USA (2002). *Sonic Adventure 2: Battle*. [Mini Disc].

– (2003). *Sonic Adventure DX: Director's Cut*. [Mini Disc].

Spector, Paul E. (1992). *Summated Rating Scale Construction : An Introduction*. Sage Publications, Inc.

Square Enix 1st Production Department (2010). *Final Fantasy XIII*. [Blu-Ray].

Steiner, David L. (2003). "Starting at the beginning: an introduction to coefficient alpha and internal consistency". In: *Journal of Personality Assessment* 80.1, pp. 99–103.

Sumner, Robert W. et al. (2005). "Mesh-based Inverse Kinematics". In: *ACM SIGGRAPH 2005 Papers*. SIGGRAPH '05. Los Angeles, California: ACM, pp. 488–495. DOI: 10.1145/1186822.1073218. URL: http://doi.acm.org/10.1145/1186822.1073218.

Takagid, Hideyuki (2001). "Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation". In: *Proceedings of the IEEE* 89.9, pp. 1275–1296.

Tarn Adams (2006). *Slaves to Armok: God of Blood Chapter II: Dwarf Fortress*.

Thomas, Frank and Ollie Johnston (1997). *The Illusion of Life: Disney Animation*. Hyperion.

Tobler, R. F., S. Maierhofer, and A. Wilkie (2002). "A multiresolution mesh generation approach for procedural definition of complex geometry". In: *Proceedings SMI. Shape Modeling International 2002*, pp. 35–271. DOI: 10.1109/SMI.2002.1003526.

Togelius, J., G. N. Yannakakis, and S. Karakovskiy (2011). "Assessing Believability". In: *Believable Bots: Can Computers Play Like People?* Ed. by Julian Togelius et al. Springer, pp. 215–230.

Togelius, Julian, Tróndur Justinussen, and Anders Hartzen (2012). "Compositional procedural content generation". In: *Proceedings of the The third workshop on Procedural Content Generation in Games*, p. 16.

Togelius, Julian et al. (2011a). "Search-Based Procedural Content Generation: A Taxonomy and Survey". In: *IEEE Transactions on Computational Intelligence and AI in Games* 3.3, pp. 172–186.

Togelius, Julian et al. (2011b). "What is procedural content generation? Mario on the borderline". English (US). In: *2nd International Workshop on Procedural Content Generation in Games, PCGames 2011 - Co-located with the 6th International Conference on the Foundations of Digital Games*. ISBN: 9781450308724. DOI: 10.1145/2000919.2000922.

Togelius, Julian et al. (2013). "Procedural Content Generation: Goals, Challenges and Actionable Steps". In: *Artificial and Computational Intelligence in Games*. Ed. by Simon M. Lucas et al. Vol. 6. Dagstuhl Follow-Ups. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, pp. 61–75. ISBN: 978-3-939897-62-0. DOI: 10.4230/DFU.Vol6.12191.61. URL: http://drops.dagstuhl.de/opus/volltexte/2013/4336.

Tolkien, J. R. R. (1947). "Essays Presented to Charles Williams". In: Oxford University Press. Chap. On Fairy-Stories.

Treanor, Mike et al. (2015). "AI-Based Game Design Patterns". In: *Proceedings of Foundations of Digital Games*. Pacific Grove, USA.

Tu, Xiaoyuan and Demetri Terzopoulos (Apr. 2000). "Artificial Fishes: Physics, Locomotion, Perception, Behavior". In: *Computer Graphics* 28. DOI: 10.1145/192161.192170.

Ubisoft Montreal (2014). *Far Cry 4*. [DVD-ROM].

Umarov, I. and M. Mozgovoy (2012). "Perspectives, Believable and Effective AI Agents in Virtual Worlds: Current State and Future". In: *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)* 4 (2), pp. 23–50.

Unity Technologies (2005). *Unity*.

Upper One Games (2014). *Never Alone*. [Digital].

Vaishnavi, V., W. Kuechler, and S. Petter (2004). *Design Science Research in Information Systems*. URL: http://www.desrist.org/design-research-in-information-systems/. Last updated December 20, 2017.

Videocult (2017). *Rain World*. [Download].

Virzi, Robert A. (1992). "Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough?" In: *Human Factors* 34.4, pp. 457–468. DOI: 10.1177/001872089203400407.

Wampler, Kevin and Zoran Popovic (July 2009). "Optimal Gait and Form for Animal Locomotion". In: *ACM Trans. Graph.* 28. DOI: 10.1145/1576246.1531366.

Warpefelt, H. (2016). "The Non-Player Character - Exploring the believability of NPC presentation and behavior". PhD thesis. Stockholm, Sweden: Stockholm University, Faculty of Social Sciences, Department of Computer and Systems Sciences.

Warpefelt, Henrik, Magnus Johansson, and Harko Verhagen (2013). "Analyzing the believability of game character behavior using the Game Agent Matrix". In: *Proceedings of DiGRA 2013: DeFragging Game Studies*. Atlanta, USA.

Whitlatch, T. (2015a). *Principles of Creature Design: Creating Imaginary Animals*. Design Studio Press.

– (2015b). *Science of Creature Design: Understanding Animal Anatomy*. Design Studio Press.

Wiemeyer, Josef et al. (Aug. 2016). "Player Experience". In: pp. 243–271. ISBN: 978-3-319-40611-4. DOI: 10.1007/978-3-319-40612-1_9.

Wiggins, Geraint A. (2006). "Searching for Computational Creativity". In: *New Generation Computing* 24 (3), pp. 209–222.

Winegarner, Tyler (2013). *The Witcher 3: Wild Hunt is 35 times the size of the Witcher 2*. http://www.gamespot.com/videos/the-witcher-3-wild-hunt-is-35-times-the-size-of-th/2300-6409854/. Accessed: 01-2016.

Wobbrock, Jacob O. and Julie A. Kientz (2016). "Research contributions in human-computer interaction". In: *Magazine interactions* 23.3, pp. 38–44.

Yannakakis, Geogios N. (2012). "Game AI Revisited". In: *Proceedings of the 9th Conference on Computing Frontiers*. CF '12. Cagliari, Italy: ACM, pp. 285–292. ISBN: 978-1-4503-1215-8. DOI: 10.1145/2212908.2212954. URL: http://doi.acm.org/10.1145/2212908.2212954.

Yannakakis, Georgios N., Antonios Liapis, and Constantine Alexopoulos (2014). "Mixed-Initiative Co-Creativity". In: *Proceedings of the 9th Conference on the Foundations of Digital Games*.

Zhao, Jianmin and Norman I. Badler (Oct. 1994). "Inverse Kinematics Positioning Using Nonlinear Programming for Highly Articulated Figures". In: *ACM Trans. Graph.* 13.4, pp. 313–336. ISSN: 0730-0301. DOI: 10.1145/195826.195827. URL: http://doi.acm.org/10.1145/195826.195827.

Zubek, Robert (2010). "Needs-Based AI". In: *Game Programming Gems 8*. Ed. by Adam Lake. Course Technology PTR, pp. 302–311.