

UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Annotating Player Experience in Telltale Games The Wolf Among Us

Permalink

<https://escholarship.org/uc/item/69n1d6tp>

Author

Murray, John Thomas

Publication Date

2018

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

**ANNOTATING PLAYER EXPERIENCE IN TELLTALE GAMES
THE WOLF AMONG US**

A thesis submitted in partial satisfaction
of the requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER SCIENCE

by

John T. Murray

June 2018

The Thesis of John T. Murray
is approved:

Professor Noah Wardrip-Fruin, Chair

Professor Michael Mateas

Professor Katherine Isbister

Tyrus Miller
Vice Provost and Dean of Graduate Studies

Copyright © by
John T. Murray
2018

Table of Contents

List of Figures	v
List of Tables	vi
Abstract	vii
Acknowledgments	viii
1 Introduction	1
1.1 Background	3
1.2 Summary and Overview	5
2 Measuring Affect	8
2.1 Facial Action Units	8
2.2 Measuring Body Responses	10
2.3 Heart Rate	11
2.4 Skin Conductivity	11
3 Study Design	13
3.1 The Wolf Among Us	13
3.2 Sensual Evaluation Instruments	16
3.3 Protocol	18
3.4 Expected Results	21
4 Narratologically-inspired Annotation Schema	23
4.1 Beats	24
4.2 Choices	27
4.3 Story Events	28
4.4 Lenses	29
5 Results	32
5.1 Normalization	33

5.2	Participants	36
5.3	Emotional Signals	37
5.4	Story Elements	37
5.4.1	Character Affinity Analysis	37
5.4.2	Emotional Choice Analysis	39
5.4.3	Beat Analysis	41
5.4.4	Value analysis	45
5.4.5	Story Analysis	47
5.5	Summary	49
6	Conclusions and Future Work	50
A	Beat Charts	53
	Bibliography	56

List of Figures

3.1	SEI Models	16
3.2	Video Recording Setup. The top left is the gameplay video along with the ATF display, the top right is an overhead view of the usage of the SEI, and the bottom left and bottom right are the other views of the space and game screen.	17
3.3	Beat span annotations in Premiere Pro	20
5.1	Normalized ratio of affect events to time	34
5.2	Normalized mapping of players to characters according to affect measures	38
5.3	Sum Methods per Time Window	41
5.4	Sum of Each Method Per Beat	42
5.5	SEI Usage Per Beat	44
5.6	Value-oriented sorting. Each player's relative affective response to each value is shown.	46
5.7	Player-sorted Value Beat Emotional Signals	47
5.8	Player affect by event	48
A.1	Skin Conductivity Peaks Per Beat	54
A.2	Heart Rate Peaks Per Beat	54
A.3	Valence (Facial Expressions) Events Per Beat	55

List of Tables

2.1	Facial Action Units Recognized by OpenFace and Affdex SDK . .	9
2.2	Categorical Emotion Values provided by Affdex SDK	10
4.1	Dramatic Beats and Cast in Chapters 1, 2 and 3	25
4.2	Story Events	28
4.3	Story Values	30
4.4	Cast Involvement in First Three Chapters	30
5.1	Sample Database Records	35
5.2	Sample Database Records (Continued)	35
5.3	Choices with emotion proximity. ChId represents Choice ID, Ch__[Button] represents the button used to select a particular choice label. . . .	40
5.4	Unanimously Agreed Choice Prompts. DID representes Decision ID, ChID represents Choice ID, while Ch__[Button] represents but- ton associated with choice option	40

Abstract

Annotating Player Experience in Telltale Games *The Wolf Among Us*

by

John T. Murray

Modern adventure games combine storytelling with gameplay using conventions familiar to film and television audiences. These games are different from others in that the experience depends on narrative content, and so is challenging to evaluate for improving content. The subgenre pioneered by Telltale Games uses dramatic choices to shape player experience. I developed and applied an annotation schema to a dataset of 6 player traversals of the first episode of *The Wolf Among Us* by Telltale Games. The schema supports comparison of player responses consisting of a combination of psychophysiological measurements and self-report events. These events are segmented and classified based on concepts drawn from Robert McKee, including the beat and story values, as well as traditional narrative elements such as character and plot. The results provide evidence that content annotations can facilitate analysis of players responses in longer traversals in spite of different levels of expressivity.

Acknowledgments

I am grateful for the mentorship, guidance and support of my advisors, Noah Wardrip-Fruin and Michael Mateas. I will continue to aspire to the example they have set, both in their patience and in their high expectations, as well as their passion for the field. I am particularly grateful for Noah Wardrip-Fruin's insightful direction through the minefield of research as well as his support over the years. I thank my committee member Katherine Isbister, whose work has served as an inspiration and whose help has been invaluable in the process of understanding player emotions. I am grateful for the other members of the faculty whom I have worked with and learned from over the years at UCSC, including Jim Whitehead, Arnav Jhala, Marilyn Walker, James Davis and many others.

I had the amazing opportunity to be part of something truly special made possible by the hard work and dedication of Michael Mateas and Noah Wardrip-Fruin: the Expressive Intelligence Studio. While a member there, I met both lifelong friendship as well as collaborators who I hope to continue knowing for years to come, including Jacob Garbe, Aaron Reed, Kate Compton, Johnathan Pagnutti, Stacey Mason, James Ryan, Peter Mawhorter, Anne Sullivan, Josh McCoy, Jo Mazeika, Gillian Smith, Joe Osborn, Ben Samuel, Chris Lewis, Adam Smith, Chaim Gingold, April Grow, Eric Kaltman, Mike Treanor, James Skorupski and many others.

I have also had the privilege of collaborating with a number of scholars and artists who have shaped my work and taught me much about what it's for, including Anastasia Salter, Raquel Robinson, Mark Marino, and Joellyn Rock. I am grateful for the support of other members of the electronic literature community, including my former advisor Matthew Kirschenbaum, Caitlin Fisher, Dene Grigar, Mark Sample, Rob Wittig, Leonardo Flores, Stuart Moulthrop and Jason Nelson.

Finally, I thank my partner, my family, and my close friends for their support.
Without them, this would not have been possible.

Chapter 1

Introduction

Stories are designed to elicit emotional responses and to engage their audience. Modern adventure games, particularly those defined by game developer Telltale Games, have evolved the traditional puzzle-focused storygame closer to the types of experiences found in television and film, both in content and in presentation [1]. These games take advantage of licensed content to tell stories by using a combination of interesting choices and participatory action sequences (known commonly as quick-time events (QTEs)). The experience that players of these games have is inseparable from the story itself and represent an excellent example of combining stories and games. Unlike many games, the player experience in this genre derives from the narrative content portrayed, within which player actions are contextualized. This dependency on narrative content motivates the methodology of this thesis, which combines content analysis with recordings of player experience data. The data represent player choices and emotions, as detected and reported by various means. This thesis presents a methodology to annotate player traversals for analysis and a case study of its application using player responses to *The Wolf Among Us* (TWAU) by Telltale Games [2]. I developed three distinct lenses to approach storygame content for understanding player responses: story values,

characters, and choices. I use the content encoding schema to annotate videos with distinct segments of play that correspond to narrative beats and choices in order to contextualize player responses using physiological data [3], and I use the sensual evaluation instrument (SEI) to record instances of self-reported affect [4]. My collaborator, Raquel Robinson, completed much of the annotations for the physiological and facial affect measures used in this thesis, while I provided annotations for the choices and narrative content. For further details about the study procedure, as well as a comparison of the different measures collected and the tool used, All The Feels (ATF), see Robinson’s master’s thesis [5].

The content of interactive narrative games can be viewed as two interconnected layers: the ludic layer, where the player takes actions with the goal of achieving some desired outcome, and the narrative layer, where the content itself contains a structure and is designed with the goal of eliciting a response from the player through sequence and presentation. The genre of the modern adventure game intertwines the two. The genre, which I label cinematic choice-based adventure games based on cinematic conventions and the emphasis on the act of choosing rather than on the branching narrative often attributed to interactive storygames, was established by Telltale Games through its various series that shared many features in common. In the genre, each choice is dependent on the narrative and not some underlying simulation. The choices themselves can play many roles in causing an effect in a player. Mawhorter’s theory of choice poetics provides one such approach to understanding choices and their effects [6]. In it, he breaks down choices into components for analysis to identify rhetorical patterns and the nature of common types of choices, such as dilemmas and their realization as choices produced by a generator he developed. Not every decision within the game has the same weight in the story, and not every emotional moment can be

traced to a player decision, so one way of approaching the narrative layer is to measure and chart the timing and content of the narrative itself. The narrative schema described in this thesis is a result of this strategy.

This thesis is organized as follows. In this chapter, I review related work in player experience analysis. In the next chapter, Chapter 2, I describe the specific psychophysiological measures collected in the study. The study design is presented in Chapter 3, along with the features that comprise the different analytical lenses. The schema itself is described in Chapter 4, along with its inspirations. Chapter 5 presents the analysis process and results. I pay particular attention to relating the various metrics that result from the visualizations back to the underlying content and how such an analysis is facilitated by the annotation schema. Finally, I discuss the implications and motivate future work in Chapter 6. The work in this thesis provides a framework for future studies to combine content analysis, modeling, and empirical studies in order to improve and understand storygames.

1.1 Background

Player experience evaluation has evolved out of the realization that the experience of entertainment software is distinct from that of other types of software [7]. The current state-of-the-art methods include qualitative instruments, such as surveys and questionnaires [8], think-aloud [9], and interviews. Quantitative methods include video analysis of playthroughs [10], physiological measurements [11], and telemetry data [12]. My collaborator and I build on these techniques in a study design we conducted of players playing *TWAU* by using a mixed-method approach that borrows aspects of each to bring to bear on a particular genre of storygames.

Stories and games have been combined in both research contexts and commer-

cial games. One key example of interactive drama, a genre with real-time interaction between players and characters, is *Façade*, which uses real-time sequencing of dramatic beats [13]. *Façade* presented a challenge for evaluation because of the experimental nature of using parser-based input, as well as the novelty of experiencing dynamic content sequencing for players [14, 15]. Other experiments in interactive narrative generation have resulted in work that focuses on understanding particular aspects of the narrative experience, such as surprise [16, 17], suspense [18], and emotions [19]. These works use a methodology that requires the games to be crafted specifically for the experiment to provide the appropriate experimental control. On the qualitative side of analysis, El-Nasr et al. analyzed *Façade*, one of the closest studies to the present work [20]. In the analysis, the authors address player behavior, emotion, and interpretation by using a variety of methods. Another user study that focused on narrative was conducted by Mallon and Webb [21], which also used a reader-response approach.

Recent work in computational narratology has drawn from methods used in computational linguistics of modeling and annotating specific artifacts [22, 23, 24] and sought to examine them as corpora [25, 26] from which patterns can be detected and theories can be developed. These inspired the design of the study and provided a roadmap for future uses of the dataset that resulted from the study.

Gameplay telemetry is a valuable source of empirical data and is a cornerstone of industry assessment of game success [12]. Narrative games have also been the subject of telemetry assessment [27]. Other approaches focus on evaluating the properties of the narrative itself [28]. A promising strand of research is the use of direct measures of brain activity and physiological measurements to evaluate player engagement with narratives [29]. These studies suggest that heart rate

variability may be a more relevant measure than heart rate peaks are.

Several systems were designed to aid in the analysis of player experience data. Two of these are Microsoft Tracking Real-time User Experience (TRUE) [30], and Data Cracker [31], a tool for online user experience evaluation built alongside the game *Deadspace 2* and with the input of the development team. Microsoft TRUE combined data analytics with specific measurement protocols built in, such as surveying users during gameplay. Microsoft TRUE’s subjective analysis was accomplished by pausing the game every three minutes and prompting the players to answer a survey question. The system was deployed in a large-scale case study in a released game, providing evidence that the technique could provide valuable insights into aggregate datasets. Both Microsoft TRUE and Data Cracker focus on spatial and logical components that are key to the experience of many genres, whereas the present work focuses on the story content found in storygames. The present study also looks for smaller sets of players, and the method of annotating player response data and actions can be used with already-released games without instrumenting the system itself, making it useful in research contexts outside of a particular game development context.

1.2 Summary and Overview

This thesis describes a player experience study that involved collecting player traversal videos from seven participants who played *TWAU* by Telltale Games. The study design and the collection of the data were conducted collaboratively, as well as some of the initial annotations. In this chapter, I describe an annotation schema based on narrative features suggested by story craft and narratology, and I use it to annotate portions of the dataset. The study was designed to obtain a dataset that could be used for a variety of purposes. The collection and initial

results were promising but revealed several shortcomings in the available tools and the initial annotation schema.

The approach provided evidence supporting the schema’s usefulness in detecting content preferences among the players, and it further provided a means of understanding the overall emotional flow of the work by aggregating the different measures. By manually inspecting the videos, including reviewing the think-aloud and watching the gameplay itself, I confirmed that the features indicated were significant. A vast number of potential interactions exist among the content, context, and player physiology, and these could cause variations in the observed signals.

However, the process of annotating the data was labor intensive. Once tabulated, the annotations were challenging to associate back to specific features in the original footage, as the process used to calculate the results (Google Sheets) was not connected with a video player that could directly move to the relevant locations. The use of aggregate data prevented a more fine-grained analysis of associating content which may have given rise to player responses, and the database representation presented a challenge in tabulating player choices and associating content variations directly to player choices. The data sensors did not always have continuous coverage. This was not highlighted in the visualization and, as a result, may have affected some of the distributions of annotated events. I address these issues by looking at aggregate measures (in particular, the sums) across all participants for those that have gaps, although future work will need to account for periods during which either the participant is not actively playing or gaps exist in the data because of sensor issues.

My collaborator and I (hereafter, we) were interested not only in studying the relationship between narrative contexts and player responses but also in comparing the evaluation methods used [32]. This thesis focuses directly on the applicabil-

ity of a narratologically inspired annotation schema to reveal the relationships between response patterns and content segments.

Chapter 2

Measuring Affect

Affective computing has provided methods to assess emotions through various sensors and processing of facial videos. In this chapter, I review several key measurements used in the player study and that are present in the dataset, as well as the tools and methods associated with them.

2.1 Facial Action Units

One of the primary ways by which computing systems can understand and classify facial expressions is through an existing taxonomy known as the facial action coding system (FACS). A Swedish anatomist, Hjortsjö [33], developed the original taxonomy, which Ekman and Friesen adapted and published in 1978 [34]. The system was updated in 2002 and has served as a gold standard for research on automated facial recognition and emotion classification [35]. The reasoning behind using visual images and movies to classify emotions is that humans are capable of recognizing emotions based on pictures alone and that these expressions are the result of combinations of common facial muscles. These combinations are called action units (AUs).

Two systems are available to analyze player facial expressions, and both are incorporated into the dataset. The Affdex SDK is a toolkit that incorporates trained models that have been created using open datasets of images of faces with various features labeled, such as eye position and facial feature positions. The second library, OpenFace by Baltrusaitis et al. [36], provides a set of algorithms for training a classifier, as well as a set of pre-trained models based on openly available facial expression datasets. Both systems can identify a subset of AUs, described in Table 2.1. One of the advantages of the Affdex SDK is that it provides a ready set of classifications for the traditional categories of emotions, and can be readily incorporated into other software through a library.

Table 2.1: Facial Action Units Recognized by OpenFace and Affdex SDK

AU	Description	OpenFace	Affdex
1	Inner brow raiser	X	X
2	Outer brow raiser	X	X
4	Brow lower	X	X
5	Upper lid raiser	X	
6	Cheek raiser	X	
7	Lid tightener	X	
9	Nose wrinkler	X	X
10	Upper lip raiser	X	X
12	Lip corner puller	X	X
14	Dimpler	X	
15	Lip corner Depressor	X	X
17	Chin raiser	X	X
18	Lip pucker		X
20	Lip press	X	X
23	Lip tightener	X	
25	Mouth open	X	X
26	Jaw drop	X	
28	Lip suck	X	X
43	Eyes Closed		X
45	Blink	X	
	Smirk		X

Raquel Robinson et al. incorporated the openly available Affdex SDK system

Table 2.2: Categorical Emotion Values provided by Affdex SDK

Emotion	Action Units
Joy	6+12
Sadness	1+4+15
Surprise	1+2+5+26
Fear	1+2+4+5+7+20+26
Anger	4+5+7+23
Disgust	9+15+16
Contempt	12+14(one side)

[37] into All The Feels (ATF) used during the study [3]. The set of action units used to classify each emotion is shown in Table 2.2, from the iMotions website [38], which uses Affdex to process emotional recognition. The classification method is based on EMFACS, the emotional facial action coding system developed by Friesen and Ekman [39].

2.2 Measuring Body Responses

The results over the past decade clearly show that the relationship between mental processes and physically measurable responses is not straightforward. Nacke describes the relationships between mental processes and bodily responses as being either one-one, one-many, many-one, many-many, or null [40]. These classifications characterize how one system may influence or be indicative of either a single or many mental or physiological processes, such as those described in the following sections. Nacke further recommends starting the process of selecting measures with electro-dermal activity (skin conductivity [SC]), as it has a direct relationship with arousal and can represent responses in small time windows.

2.3 Heart Rate

Some physiological response measures relate to the central nervous system. Many hospital dramas rely on omnipresent monitors of vital signs to set up the scene of a patient near death. Beyond simply indicating physiological arousal, heart rate (HR) can be a valuable indicator of mental workloads. Heart rate variance, a measure of the variability of time between heart beats, can indicate increased mental workloads and stress [41]. Heart rate can also be measured through changes in skin tone, which may indicate a means of measuring the heart rate of players who were not previously attached to sensors, such as streamers [42]. Heart rate measures, however, can be affected by other activities, such as think-aloud [40].

2.4 Skin Conductivity

Skin conductance level, also shortened as skin conductivity or SC, is a physiological measure of electric resistance across the skin associated with the polygraph (lie detector) test. It can be measured simply and affordably with two electrodes, often attached to the fingers. The idea is that certain microscopic variations in SC result from the responses of the body's sweat produced by the *eccrine* glands as a result of changes in neurotransmitter levels associated with psychological arousal [43]. These variations can be measured for certain stimuli, in which case it is called galvanic skin response, or can be measured over time, in which case it is called electro-dermal activity (EDA). In this thesis, these are generally referred to as SC, though the usage is more as an indication of EDA. The delay of a response to a stimulus is approximately five seconds. The measure tends to drift over time as sweat accumulates, which suggests that a shorter time window average might

be of use [40, 44, 45] and [11]. Examples of events that may cause detectable changes include the nervousness one feels when on a first date or when telling a lie with high stakes. In using EDA to obtain features, [46] find local minima as points of interest, as a decrease in the SC is considered a measure of an increase in sudation, “which usually occurs when one is experiencing emotions such as stress or surprise.” Its use in game design evaluation has been primarily in high temporal resolution detection of events. An example of the use of GSR in an application is an automated indexing system for an audio stream [47]. For a comprehensive summary of the brain and its relationship to research in play, see [48].

Chapter 3

Study Design

To understand how player responses were related to the content, my collaborator and I organized a study. This study consisted of seven uncompensated graduate students whose age ranged from 20 to 35. Each participant was instructed to play through the first episode of *TWAWU*, wearing an Empatica E4 wristband sensor, and to speak any thoughts aloud. We collected the data through the output ATF, described further in Section 2.1, which uses the sensor data from the wristband to visually present the state of a streamer in the form of bars and emoticons. We also encouraged the participants to use the SEI [49, 4], designed for nonverbal self-report, shown in Figure 3.1. Because of some technical issues, only six player traversals are included in this thesis (noted as P2 to P7). This chapter starts with a summary of the game’s plot, followed by a description of the protocol and the expected results.

3.1 The Wolf Among Us

TWAWU is a cinematic adventure game released by Telltale Games in 2013 [2]. It was released first for PCs, followed by other platforms. It is based on the

world of the *Fables* comic series by Bill Willingham, published by Vertigo. The story, however, is original to the game. The game has received widespread critical acclaim, in part because of its compelling story. It is an exemplar of the genre established by Telltale Games, one which I call cinematic choice-based adventure games. Games in this genre have features that have attracted criticism, as the impact of player choice on the selection of content is not always apparent and, upon replay, can be seen as less important than it seems at the time. This consistent use of content, however, makes the genre interesting for annotation-based approaches, as the similar content can provide an anchor for what is otherwise a dynamic experience. The role that decisions and content play in shaping the player experience can further be analyzed by comparing player traversals, or playthroughs.

The first episode introduces the player character, Bigby, as the sheriff of Fabletown. Fabletown is a community of fairy-tale ex-patriots who have settled in a small part of New York. Fables, as they call themselves, have long lifespans and have brought some of the magic and wealth along with them from their homelands. Some fables are non-human and are ordered to have in their possession a magic talisman, called a glamour, that disguises them as humans. This causes friction between both humans and non-humans, as well as exacerbates class tensions. The story begins with Bigby investigating a disturbance at a tenement house. Mr. Toad, the owner, is caught out of glamour and in violation of community rules. Upstairs, the disturbance is a result of a fable, the Woodsman, assaulting another fable who refuses to identify herself. After a fight between Bigby and the Woodsman, Bigby invites the woman back to his place for questioning. At his apartment, inside the Woodlands, a luxury apartment complex for the more well-off fables, Bigby encounters Colin, one of the three little pigs that he antagonized in the old world. Colin remarks on his violent tendencies and the fact that Bigby

must make amends while persuading Bigby to share his drink with him. Bigby gets a short nap before being woken by the assistant mayor, Snow (White). She takes him to the front of the apartment, where someone has left the head of the fable whom Bigby previously rescued. Before Bigby returns to the office to begin the investigation of the fable's identity, Grendel, a non-human fable waiting in line for help at the Mayor's office, accosts Bigby about not being very helpful. Bigby brushes him off and goes into the office, regardless of the player choice during that scene. Inside, the deputy mayor Ichabod Crane is looking for someone to blame and is verbally attacking Snow. After Bigby steps in, he leaves with an odd note about a massage appointment. Bufkin, a flying monkey, then reveals himself and offers to assist in the investigation. Bigby and Snow use the books in the office to find the fable's identity as Faith, or Donkeyskin. They also discover she had a husband, Prince Lawrence, who might know more, and they use another character, the Magic Mirror, to find out that he is in trouble. Before they leave, Mr. Toad calls again, asking for help because there is an apparent intruder. The story branches at this point based on whether the player decides to investigate Mr. Toad or Prince Lawrence first.

The remainder of the story follows the investigation, culminating in a scene in the Trip Trap Bar, which involves a confrontation with Grendel and a choice between apprehending the Woodsman or Tweedle Dee, another character that Bigby encountered during his investigation and who was probing Faith's death. The episode ends with the player discovering Snow's head on the doorstep in the same fashion as Faith's, although this time, the non-fable police are involved.

This thesis focuses on the first three chapters in the game, which culminates with the decision to go to Mr. Toad's apartment or Prince Lawrence's apartment. The rest of this chapter describes the study setup and protocol.

3.2 Sensual Evaluation Instruments

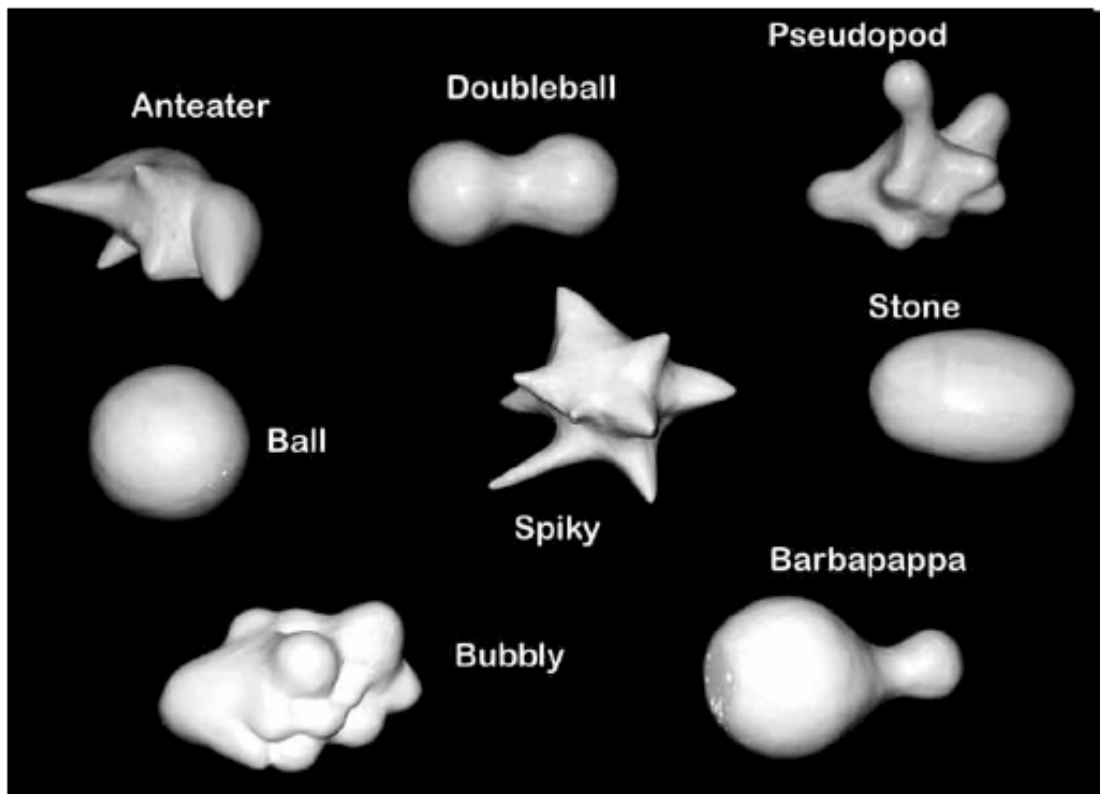


Figure 3.1: SEI Models

We decided to collect a variety of measures of player emotions through traditional measures, such as think-aloud and surveys, as well as lesser-used methods, such as the SEI. The SEI functions as a non-verbal, cross-cultural self-report method for augmenting physiological and qualitative measures by allowing participants to express their feelings using a set of sculptures developed by Isbister et al. See [49] for an example of its use in a narrative game context. The SEI consists of eight sculptures that are designed to be touched, held, or gestured with (see Figure 3.1 for the shapes and names of each). Players interact with these sculptures to indicate some internal feeling was taking place, and these moments serve both as a self-report of their experience, as well as anchors for discussion

during a post-game interview.



Figure 3.2: Video Recording Setup. The top left is the gameplay video along with the ATF display, the top right is an overhead view of the usage of the SEI, and the bottom left and bottom right are the other views of the space and game screen.

We also hypothesized that the SEI would provide insights into the more subtle emotional experiences that were prevalent in this type of narrative game experience, indicating points of interest that may not be otherwise evident from the biometric signals alone. For this analysis, I treated the SEI as an indicator of the presence of some player affect, without differentiating between each instrument. As I will demonstrate, this is useful in identifying locations of emotion through frequency measures, although multiple uses of the SEI within a short period may indicate ambiguity of player feelings rather than increased intensity. For this chapter, any SEI usage is considered an affect event (which, together with other events, can be combined into an affect signal) and is undifferentiated by the particular instrument used. Attending to patterns that relate to specific instruments or relating instruments through their usage patterns is a topic of future work.

3.3 Protocol

Each session was conducted as follows. The players first took a pre-game questionnaire, with questions covering familiarity with the *Fables* comic series and whether they have played other titles by Telltale Games. As a prerequisite to participate in the study, we required that the participants should not have played the game before so that their reactions would be initial traversals of the game content. However, we learned afterward that one participant was familiar with the comic series on which it was based, coloring this participant’s knowledge of the final scene. We then calibrated the SEI by showing the users a series of 10 photos from the International Affective Picture System [50] and having participants indicate with the SEI which association came to their mind for that particular instrument. This was used for familiarizing the users with the process of reporting affect using the instruments instead of recording specific reactions to particular stimuli. We then instructed the participants to play the entire first episode while expressing aloud any thoughts that emerged, as well as encouraged them to use the SEI as much as possible. We found that some early participants (before P2) were engrossed in the game and forgot to gesture with the SEI, so the usage varied. After noting this, we set a timer to go off periodically and serve as a reminder, although this was not as much of an issue with the other participants. We left the room for the duration of the gameplay session, only returning in case of technical difficulty.

The facial recognition software, Affdex, from ATF was also running and is displayed along with SC and HR readings in the lower right-hand corner of the video. At the end of the session, we conducted a retrospective think-aloud with the players about how they felt at certain peak moments of the game, how they felt about various characters, and why they used specific SEI objects during the game.

They also took questions from the Immersive Evaluation Questionnaire [51]) and combined them with questions relating to emotion. These provided broad strokes of the players' retrospective assessment of their experience. However, these are not included in this thesis, as I instead focus on moment-to-moment measurements and annotations.

Once we finished collecting data, I processed the gameplay screen capture and the external cameras' videos into a single video and adjusted the timing to synchronize them by using the sounds and movements shown in the videos. I used Adobe Premiere Pro to create a single video for each participant containing the gameplay video (with ATF data as a picture-in-a-picture), the top-down view, and left and right room cameras (see Figure 3.2). We both used the marker feature of Adobe Premiere Pro to hand-annotate the features described in the next section because of the sheer amount of footage, totaling over 13 hours. These annotations were exported as a comma-separated value file and tabulated with Google Sheets using a sheet for each data table.

The annotations consisted of features, including beats, choices, and affect events. My collaborator identified peaks in HR and SC through visual inspection. Choice prompts are locations in the game where a set of choices associated with controller buttons is displayed, often only for a brief time. I created a text annotation consisting of the selected choice, the choice texts, and marked each point when a player either pressed a button (indicated by a white highlight of the corresponding option) or when the timer expired.

At this point, I used the location of the annotations of choices as a basis for a set of time windows of eight seconds before or after each choice. These were used to select the presence of the affect annotations using a formula. Next, I charted various statistical measures of the choices, paying attention to the distributions

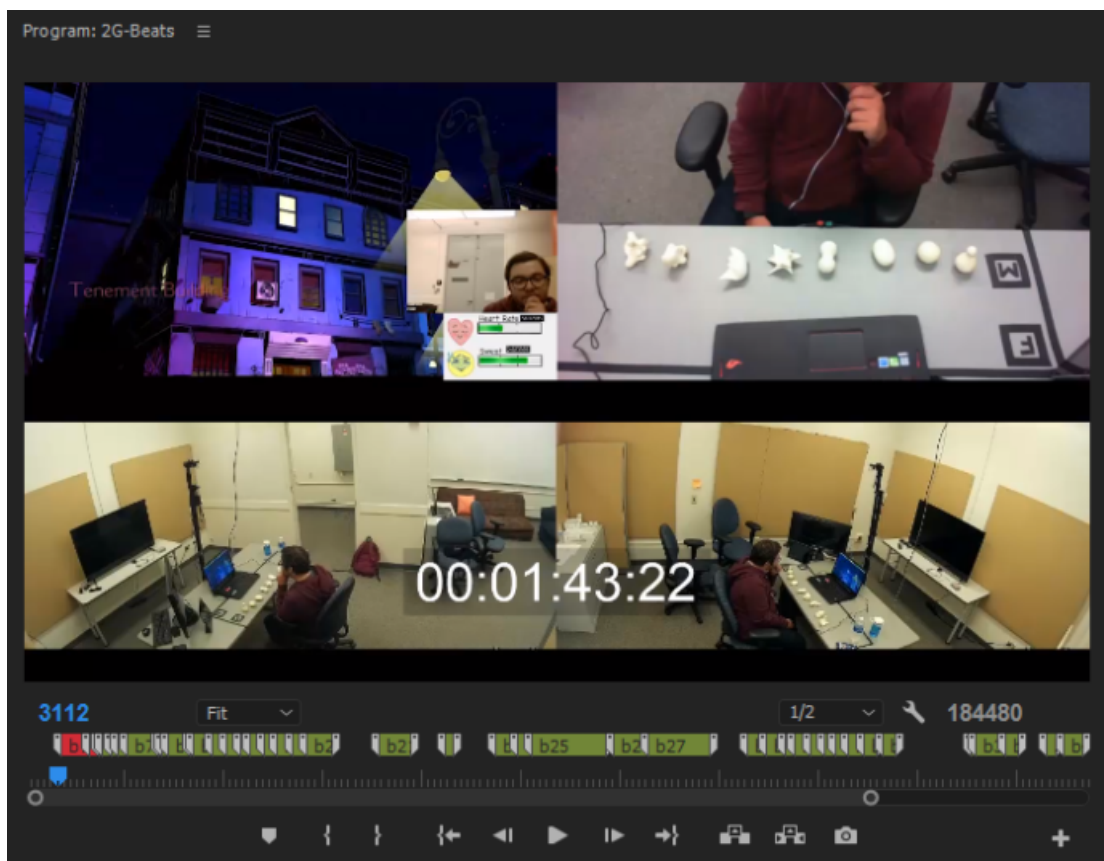


Figure 3.3: Beat span annotations in Premiere Pro

of the choices themselves. The results of these are detailed in the Section 5.4.2.

Using pairs of videos in a side-by-side configuration, I identified and annotated segments of each traversal as containing specific dramatic beats, described further in Section 4.1. The annotation was completed by identifying the particular shot or frame associated with the start and end of the beat and matching it. The next section presents and discusses the initial features used to annotate the narrative elements of the playthroughs.

The multi-channel nature of the data necessitated that the videos be synchronized to attribute response measures to particular elicitation events or to determine that they were a part of a non-game related activity, such as interruptions. These data can be divided into player response data and game content data. The narrative elements selected for analysis include the dramatic beat, story values, and character. Refer to Section 3.1 for a summary of the plot and story.

3.4 Expected Results

TWAW has many emotionally charged events in which players might react, as well as ambiguous outcomes for many of the choices provided. The game takes advantage of the fact that the characters have pre-existing relationships with the player character and his role both as a sheriff and as the eponymous Big Bad Wolf.

I anticipated that the content segments would contain the most instances of affect (excluding events that require carefully timed player input, also called QTEs, which, through their nature, evoke stress and an increased heart rate). I also predicted that there would be common patterns among all players in their emotional involvement over the sequencing of beats, as the story is relatively fixed in its composition of events.

One of the more surprising results was the amount of variation in the expressiveness of each of the players, which necessitated the normalization process to account for the different quantities of expressive events while facilitating comparison of like content. Some of the variations in SEI usage were more prominent than physiological responses were. While some events consistently caused players to report feelings, many of the spikes in usage observed occurred during different segments. The approach used in this thesis allows the consideration of these variations with more than a pair of player videos by aligning the data to the content.

In the next chapter, I will describe the annotation schema, followed by the results and the discussion.

Chapter 4

Narratologically-inspired Annotation Schema

TWAU has three types of gameplay. The first is choosing between narrative options. The second is quicktime events (QTE), which are button-pressing sequences that require fast reaction times, usually corresponding to a violent or risky event and an on-screen prompt. The third is free-roam exploration coupled with hotspots, most akin to classical adventure games, where the player can move the player character about a region of interest and engage with elements of the environment. I decided to base my analysis of the narrative content on the story beat and focus primarily on the segments including the narrative choice type of gameplay. The story beat is also used as an organizing “content piece” in the interactive drama *Façade* [52]. The beat functions as a means of segmenting character interactions using behaviors and values. Although McKee’s book is not grounded in a scholarly tradition, his work has inspired a number of approaches to narrative and writing in the interactive digital narrative community. I also selected a set of key story events that I believed to have the potential for player response and which I believed to be directly related to the primary plot. These

provided spans of time that I used to focus comparisons of each player-decided narrative traversal and that served to align the player data. I also annotated interruptions and excluded those periods from the analysis, although the periods of interruption may have an effect that isn't accounted for in terms of the duration of the interruption and the return of a participant's arousal levels to a neutral state. The following sections describe each annotation feature in detail.

4.1 Beats

McKee defines the “beat” as a unit of action [53] that involves one character initiating a behavior and another character or environment reacting. The classic example is one character using a strategy to gain something that they desire while another character either deflects or concedes. In *TWAU*, one tense scene involves a character (Ichabod Crane, deputy mayor) savaging the player character with blame (beat #31). This beat ends when the first character turns their attention to another subject and ceases expressing blame. Each beat is further annotated with a story value, which is a label that describes the overall values at stake in the story. It can be considered an empathetic label, as it is primarily based on the reception of the story material by an audience, rather than purely the concerns of the characters involved, and so is distinct from systems which look purely from an agent's perspective. The story value I assigned to this particular beat as a binary-opposition is ego/community, as Crane is attempting to save face at the expense of the community by laying blame on others. During that beat, the receivers of his blame react: in the case of the beat, this reaction is in part determined by player choice. Bigby can avoid blame by staying quiet, assume blame from Snow, or try to redirect the conversation.

I analyzed the first three chapters of *TWAU* for the location of individual

Table 4.1: Dramatic Beats and Cast in Chapters 1, 2 and 3

ID	Description	Cast
1	Toad discovered not being in Glamour	Bigby;Toad
2	Toad asks Bigby to deal with the Woodsman	Bigby;Toad
3	Toad yells at TJ to get back in his room	Toad;TJ
4	Woodsman asks Faith who he is, beating on her.	Woodsman;Faith
5	Faith spits blood on Woodsman	Woodsman;Bigby;Faith
6	Bigby asks Woodsman why he did it	Woodsman;Bigby
7	Woodsman attacks Bigby, who defeats him.	Woodsman;Bigby
8	Bigby asks Faith to leave	Woodsman;Bigby;Faith
9	Woodsman threatens Faith and refuses to pay	Faith;Bigby;Woodsman
10	Woodsman insults Faith	Bigby;Faith;Woodsman
11	Bigby destroys Toad's car.	Toad;Woodsman;Bigby
12	Toad laments car and Bigby's help	Toad;Bigby
13	Woodsman attacks Bigby	Woodsman;Bigby
14	Faith loots Woodsman, Bigby's transforming	Bigby;Faith;Woodsman
15	Faith kicks Woodsman then pushes ax in	Bigby;Faith;Woodsman
16	Bigby gently questions Faith (Cigarette)	Bigby;Faith
17	Bigby Asks about Faith's employer	Bigby;Faith
18	Faith stops Bigby from pursuing Woodsman	Bigby;Faith
19	Bigby asks about obligation; Faith demures	Bigby;Faith
20	Faith touches Bigby's wounds; plans meeting.	Bigby;Faith
21	Beauty asks Bigby to cover for her	Bigby;Beauty
22	Beast asks about Beauty's whereabouts	Bigby;Beast
23	Bigby asks Colin to get out of chair	Colin;Bigby
24	Bigby provides Colin a smoke	Colin;Bigby
25	Colin entreats Bigby to give him a drink	Colin;Bigby
26	Snow retrieves Bigby	Snow;Bigby
27	Snow asks Bigby how he knows Faith	Snow;Bigby;Faith
28	Snow wants to know Bigby's early impressions	Snow;Bigby;Faith
29	Snow wants to tell Crane	Snow;Bigby;Crane
30	Grendel is upset that Bigby isn't more helpful	Bigby;Grendel
31	Crane is attempting to assign blame	Crane;Snow;Bigby
32	Crane blames Bigby and Snow; asks for leads	Crane;Bigby;Snow
33	Asks for massage and wine before storming out	Crane;Snow;Bigby
34	Snow stops Bigby from provoking Crane	Snow;Bigby
35	Bufkin says hi	Bufkin;Bigby
36	Snow asks for help in researching identity	Snow;Bigby;Bufkin
37	Mirror accuses Bigby of being Cruel	Mirror;Bigby
38	Bigby can't read it, but Bufkin offers to help	Bufkin;Snow;Bigby
39	Bufkin reads story of Faith	Bufkin;Snow;Bigby;Faith
40	Declare Husband person of interest	Bigby;Snow;Lwrnc.
41	Finds out Husband is in trouble	Bigby;Mirror;Snow;Lwrnc
42	Toad calls about intruder	Bigby;Snow;Toad

beats. I identified a total of 42 beats, listed in Table 4.1. Other annotators may select different segments for beats, but these beats served the desired purpose of segmenting content and aligning player data on characters and story values. The criteria used for designating a dramatic exchange as a beat was whether a single behavior was being pursued by a character and was directed toward another character. A new beat begins when either the agent of the action or the behavior itself changes.

Each beat was further tagged with the characters involved, including the character taking action and the responding character, as well as the story value(s) at stake. I found, as expected, there were recurring themes of truth/lies, justice/injustice, and community/ego. The idea of a story value pair was introduced by McKee. His idea was that during each unit of action, some movement between a positive and negative valence of a story value would occur, such as from negative to positive, or from positive to double positive. The story value of justice, for instance, is invoked initially when Bigby visits Toad’s apartment, as Toad is getting justice for the disturbance upstairs. The value swings directly to negative, however, when he realizes that he is breaking the currently unjust rules of Fabletown by not being in glamour. This variation from one side of value to another is often imprecise and dependent on context, but it generally recognizes the emotional nature and thematic unity of actions in a scene. The dynamics caused by value shifts keep the reader’s interest, and the cause of the change is often the conflict itself.

In this initial annotation, each story value pair is used to describe the main values that are at stake in a particular beat. This classification approach doesn’t capture finer details of the how values manifest in a scene. For instance, the portrayal of values in a scene is often characterized as moving from one “charge”

to another, either from a positive to a negative. The annotation schema avoids this level of annotation and instead focuses on which binary pair is associated with a particular beat.

The annotation process was applied for sequences involving dramatic choices. Some of the quicktime segments were integrated into a dramatic beat, such as the first fight sequence (Bigby and the Woodsman fighting), which was consolidated into a single beat with the Woodsman as the agent initiating the action (starting the fight). There are segments of content that are dependent on players selecting hotspot while exploring areas. Sometimes, upon selecting an item or option, a character will interact with the player character based on the object selected, as is the case in the business office scene and the final scene with Grendel. These beats are significant in that they are not always required to progress, which means they may require a different strategy than the simple annotation schema described in this thesis. One possibility is a graph-based approach, which accounts for the dependent relationship between beats and various dependencies.

4.2 Choices

Dramatically presented choices are the central game mechanic in this kind of storygame and consist of two or more options shown to the player for selection. In timed prompts, the default option is chosen if the timer expires. It is often silence but can also be some form of inaction. For instance, when Bigby is presented a choice of whether to give Faith money after noting she would be short, the timer expiring indicates that Bigby chose not to give her any.

I developed a simple annotation schema that records the location and content of “choice prompts” in the game as integers representing the frame location from the beginning of the file. I assigned each choice prompt an ID (noted as “ChID”

or “ChoiceID” in the database records) based on having the same choice text (the labels displayed to the player as part of a prompt). In non-dramatic segments, I encode these with a separate annotation type, indicating that they are menu options, as they play a different role than dramatic, timed choices. They are usually part of some process to obtain exposition (plot points delivered outside of a dramatic context), such as interactions with the Magic Mirror that revealed Prince Lawrence’s being in trouble.

4.3 Story Events

Table 4.2: Story Events

ID	Event Description
1	Bigby Catches Toad Out of Glamour
2	The Woodsman Hits Faith
3	Faith Rescues Bigby
4	Bigby Responds to Beast
5	Bigby (doesn’t) Give Colin a Drink
6	Faith’s Head is Discovered
7	Faith’s Identity is Revealed
8	Prince Lawrence is Shown in Mirror
9	Toad Calls for Help
10	Prince Lawrence is Informed/Discovered
11	Bigby (Fails to) Catch Toad in a Lie
12	Snow White Confides in Bigby
13	Grendel Insults Snow
14	Grendel Fights Bigby
15	Bigby (Doesn’t) Remove(s) Grendel’s Arm
16	Snow’s Head is Revealed

Another concept which McKee uses but which has a more general meaning in narratology is the story event, also known as a plot point, where an irreversible change that propels the story forward occurs. The story event can include information revealed, significant actions taken, and revelations about other characters

or situations. Some of these events invite player participation in their outcome, but the presence and content of the events are mostly fixed. I selected the events listed in Table 4.2 that describe the plot of the first episode. Each propels Bigby forward in his story. These usually are the key point of a scene, and so beats usually precede a story event or build up to one. For instance, the first scene between Bigby and Toad reveals the nature of glamours as a means for non-human residents to pass as humans and that it is against the rules to be out without having one. This sets up one of several reasons behind the class resentment that leads to Grendel provoking and attacking Bigby in the climax.

4.4 Lenses

Combining objective and subjective measures into a single analysis requires elements that relate them to one another. I developed a set of low-level features that enable parts of the player experience to be analyzed. Interpretive lenses can be used to guide the application of these feature sets. I applied these lenses to analyze the first three chapters of the game, “Disturbance,” “Woodlands,” and “Mirror Mirror.” The fourth chapter provides two variants based on player choice at the end of the third chapter, which presents a challenge for linearly analyzing beats. The content is largely the same between playthroughs for the first three chapters. The following sections go into more detail into how these features will be used to reveal the relationship between players and the story, as well as defining the objectives for the various charts and calculations that I employ.

Story Values

The first perspective looks at a story as a set of values expressed through the discourse. According to McKee, scenes “turn” on a value change. Each value is expressed as a binary opposition, where each beat revolves around the dominance

Table 4.3: Story Values

Value	Associated Beats
Community/Ego	4,5,7,9,11,13,15,31,33,34
Survival/Bravery	8,14,19,
Truth/Lies	6,21,22,26,27,29,36,
Justice/Injustice	1,2,10,12,16,17,25,30,41,42
Duty/Hedonism	32
Other/None	28,40

of one or the other. The principle value-pairs in the first three chapters of *The Wolf Among Us* are listed in Table 4.3. These represent important motivations for both the characters in taking actions and the audience in caring about them and are present in virtually every dramatically portrayed scene that includes conflict. A player’s affinity for particular types of content can be approximated by comparing their response during beats associated with story value pairs.

Character

Table 4.4: Cast Involvement in First Three Chapters

Cast Name	Beats Present
Bigby	All but 3
Toad	1,2,3,11,12,42
TJ	3
Woodsman	4,5,6,7,8,9,10,11,13,14,15
Faith	4,5,8,9,10,14,15,16,17,18,19,20
Beauty	21
Beast	22
Colin	23,24,25
Snow	26,27,28,29,31,32,33,34,36,38,39,40,41,42
Crane	29,31,32,33
Grendel	30
Bufkin	35,
Mirror	37,41
Lawrence	40,41

Some characters may be more compelling or resonate with players more than others. Toad’s plight as a disadvantaged member of the community is overshadowed

owed by his acrimonious attitude toward Bigby, leading to the suspicion that Bigby (and the player) expresses in some scenes. Grendel's valid grievances are hidden behind a wall of rage and disrespect. Faith and Snow are both sympathetic and conflicted. I investigated ways to estimate individual players' feelings during scenes in which characters are involved (if not physically present). In this phase, I did not include either the valence or the intensity beyond measuring the frequency of occurrences, although other measures of emotion may shed additional light.

Choices

There are two ways I analyzed the choices and the selections players made. First, I analyzed the decisions players made for agreement (e.g., whether many or all players made the same decision when presented with the same choice prompt). This provided a rough way of determining if there were perceptions of different outcomes and also was a way of finding segments of gameplay where player experience diverged due to either preference or the experience itself.

These three perspectives represented different layers of narrative engagement:

1. **Story value:** specific labels associated with the stakes for which a character is fighting and that make players relate to said characters.
2. **Characters affinity:** how some characters may be more or less engaging to the player than others.
3. **Decisions and interactivity:** where the recorded player actions may or may not line up with other players, including choices and decisions.

In the next section, I review the results of these perspectives.

Chapter 5

Results

In this chapter, I describe and discuss the results of applying the annotation schema to the first three chapters and present a preliminary analysis.

The annotations, which we recorded through analysis of individual videos, were imported through the CSV files into a Google Sheets document. Each table corresponds to a sheet. The calculations were performed by a combination of table formulas and references, often using selections based on values stored in temporary tables that recorded preliminary data. For instance, one such table summed the number of events in a given category along columns for each of a set of segments based on the start and stop point for a given player’s traversal.

The analysis process starts by assessing the relationships of each player using the lenses to filter the player affect events into bins based on either time or content segmentation, which are then displayed as charts. These charts fall into two categories: segment-oriented charts and content-oriented charts. The segment-oriented charts provide a means to interpret player responses associated with specific segments, in this case beats, across multiple response measures. The content-oriented charts take a particular content annotation, such as values or characters, and represent evidence for a player’s affinity for engaging or responding during

content segments that feature that element (such as a particular character, value pair) over others. Although a direct causal relationship cannot be established from these measures alone, they provide a means of assessing the large number of annotations with respect to different aspects of the narrative, and serve to focus attention on areas of interest in the larger space of the content recorded. These charts also facilitate comparisons of larger trends among players over the entirety of the dataset.

5.1 Normalization

Each player has a different amount of expressiveness for a given type of affect event. This makes it difficult to compare players on an even footing, so I developed a formula to represent a “normalized” value that was scaled according to the rate found in the dataset. Since the normalization is novel, I will use two examples to describe the process and the rationale. The first uses sample data, whereas the second uses actual data. The first is shown in Figure 5.1, where there are two players. The diamonds represent an expressive event, whereas the rectangles indicate segments of interest. For Player A (PA), three events occur within the segments of interest (labeled 5 min and 5 min), whereas Player B (PB) exhibits 9 events during the same labeled segment. The difference is both in the amount of time each set of segments occupies as well as the total number of events. For PA, this calculation results in a ratio of $1/3$, whereas for PB this is $1/2$. This indicates that PB had a higher concentration of events in the period ($1/2$) to PA ($1/3$). The comparison of these ratios indicates relative concentration of events in the corresponding segments of time, even if the segments differ in length.

For the second example, I will use actual data. Player 2 (P2) triggered 60 facial expressions from ATF (labeled as valence events) during beats in which Faith was

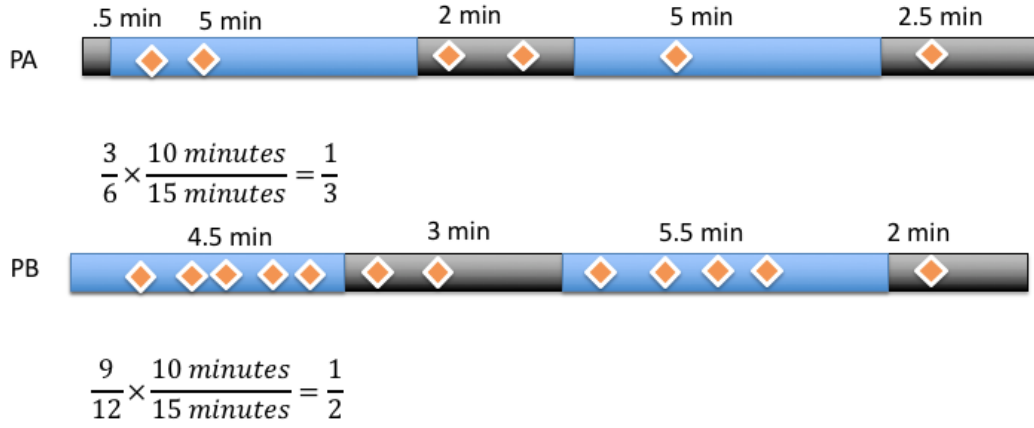


Figure 5.1: Normalized ratio of affect events to time

present. These beats occupied 23,831 frames, or 13 minutes and 14 seconds. The total time measured, however, was 96,134 frames (the total number of frames from the start of the video until the end of beat 42). Beats in which Faith was involved occupied 24.79% of the time measured. During the total period, P2 had 177 expressions recorded. This means that P2 presented 34% of their expressions during a beat where Faith was involved, whereas the beats only occupied 24.79% of game time. Multiplying these two percentages produces the dimensionless scaled ratio of the number of events that occur during the time windows to the total number of events, scaled by the proportion of time the event was measured by the total amount of time considered. The time considered in the denominator excludes interruptions for players that have them, as it only looks at time in which the player is engaged in the game.

The formula for the ratio is:

$$\frac{events(Faith)}{totalEvents} \times \frac{time(Faith)}{totalTime}$$

The ratios can then be compared to one another as they represent relative preferences compared to the total number of expressions per unit of time, rather than absolute counts that may be skewed by either the total number of events or

the total amount of time that the given content occupied.

An example of a short scene with an important impact on player response is the introduction of the antagonist, Grendel, which may fall by the wayside without this step, as the beat occupies a relatively small amount of time compared to the total time measured.

Table 5.1: Sample Database Records

ID	PID	Type	ID	ID1	ID2	In
1-2	3	1	1-2	1	2	5857
1-3	3	1	1-3	1	3	6372
1-4	3	1	1-4	1	4	6369

Table 5.2: Sample Database Records (Continued)

ID	Ch	A	B	X	Y
1-2	0	I don't make the rules	Not my problem	...	Get it fixed.
1-2	0	Do it yourself.	I'm heading up.	...	Why's he so pissed?
1-3	2	So what have I walked into?	[Head Upstairs]	...	What do you want me to do?

The quantitative data consists of a database that relates each of the inter-player content types to individual playthrough video timelines. Annotations are identified by type, with types including the content types such as the chapter and beat types, as well as affect reports (think-aloud, SEI usage, facial expressions, and HR/SC peaks). Each annotation represents this type as an integer (1-12) and an ID if present. These are separated into parts by the hyphen for choices, with the first part indicating the scene and the second indicating the choice sequence within the scene. For choices, it was important to correlate choices with the same set of options even if a choice was present in some traversals and not others, and so these are assigned the same id by identifying shared choice options. This enabled identifying shared choices between traversals. Each row contains associated data, such as the text of the choice-prompt and the decision the player made (See

Fig. 5.1). Several columns were only used by some of the types, such as the `out` column, which was useful for beats but was not used for choices. Choices were associated with the time at which the player made a selection or when the choice timer expired. Secondary tables collect the independent items, including the complete set of beats, choices, characters, and values. These are used to select related events according to the time windows for each narrative content type in a relational-database fashion.

For this initial study, I focused on the first three chapters in the first episode of *TWAU*, which amounts to about the first hour of gameplay and is summarized in section 3.1. The three chapters provide a classic introduction to a mystery, with the player character being called to the scene of a disturbance only to meet and become attached to the victim. Bigby, the sheriff of Fabletown, is confronted with his past as a violent character and is affected by the violent death of the just-introduced character named Faith. The third chapter describes the community context of the leadership of Fabletown and the core relationships that define Bigby, particular those with Snow and Colin.

5.2 Participants

Due to technical difficulties, for this analysis, I excluded P1’s data, as the physiological measures and camera data were not sufficiently complete for this analysis during the first part of the session. The remaining six participants are discussed in the results below. Of the six remaining participants, five previously played a Telltale Games title (four played *The Walking Dead Game*, one played *Sam and Max*, and one played numerous other titles). There were three men (P2, P4, and P5) and three women (P3, P6, and P7).

5.3 Emotional Signals

For this thesis, I use the term **emotional signal** to refer to a time series of measurable events that indicate the presence of affect. As discussed before, these include heart rate, skin conductivity (sweat), facial expression recognition by ATF (at the peak of an expression detection as indicated by an emoticon), player self-reports by way of one of the eight SEI, and think-aloud. These signals occur at different times due to the nature of interactive digital narratives, and so one of the challenges that I address is that of aligning player’s emotional signals to each other. While I was mostly concerned with narrative content and raw affect, it would be worthwhile to collect additional measures for insight into motivations and personality in the future. For instance, players had varying attitudes toward flirtation and romance elements hinted at during the first chapter, and this may have elicited other responses throughout the game.

5.4 Story Elements

The following sections use the lenses described in Section 4.4. These represent the transformation of objective observations into narrative-focused metrics for understanding both the players’ relationship to the content and the variations in player experience.

5.4.1 Character Affinity Analysis

The goal of a character-oriented analysis is to differentiate between players based on which characters engaged them the most. Since each sensory measure is a measure of engagement, the combination, when associated with characters, can be used to differentiate player attitudes towards characters. Using the annotations

described in Section 4, I calculate players’ affinity for each character using the normalization procedure from 5.1.

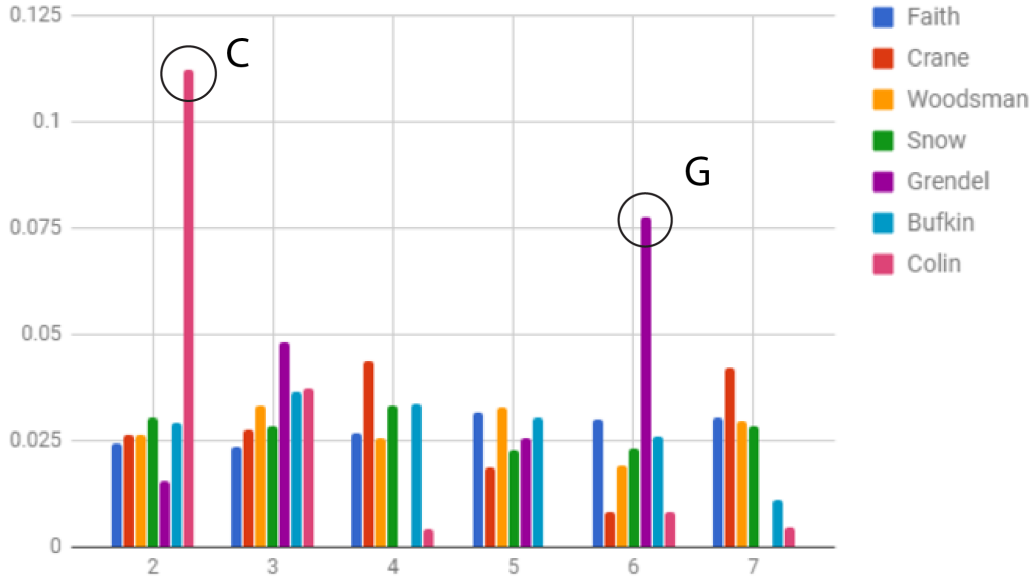


Figure 5.2: Normalized mapping of players to characters according to affect measures

The most prominent feature is P2’s response to Colin (indicated as annotation C). During the think aloud, the player reacted strongly to the character and especially to the story on which he was based (“The Three Little Pigs”). In Figure 5.2, the two character-oriented reactions that stand out are P6’s response during scenes where Grendel is present (indicated by Annotation G) and P7 and P4’s responses to scenes involving Crane (the red bar that is higher than the rest). Future work will involve charting these affective responses through player decisions. For instance, during the climax scene, P6 and P3 decide to dismember Grendel’s arm, which may have been indicated by previous responses and decisions. The unique feeling that a player has toward each character in a narrative defines their reactions and their enjoyment. This measure of player-to-character relationships may be preliminary and more associative than causal, but the concept of creat-

ing a fingerprint of how a player feels about a character indicates that storygame researchers and even the systems themselves may be able to focus more on measuring and adapting to players with appropriate modeling. Such assessment would benefit from a finer-grained approach to assessing responses, possibly at the level of individual lines of dialogue or character actions.

5.4.2 Emotional Choice Analysis

I annotated choices for the entire episode, noting the player’s choice as well as the options presented. Some choice sets had different options based on previous decisions, in which case they were assigned a further variant as a lowercase letter in the id. In Premiere Pro, I labeled in each player video every choice prompt occurrence with a semicolon-delimited list of elements including the ID, the choice the player made, and the choice options, in a predetermined order. I analyzed the choices in three ways: first, through determining the location of choices within their respective beats, and second through analyzing their proximity to emotional events, and finally for agreement.

For the first three chapters, there were 50 unique choice sets, including variants. Of these, 10 were shown only to some players, while the other 40 were shown to all players in each traversal. 47 of these choices occurred during a beat. 34 of these choices occurred within the middle of a beat, defined by the timestamp falling between the middle 60% of the beat based on time. 6 occurred during the final 20% of the beat (2-5a, 7-2, 7-7, 12-1 and 12-2), and 8 took place during the first 20% of the beat (2-4, 2-5b, 2-5d, 2-7, 5-1, 7-3, 7-4, 8-2).

In this analysis, I use the choices from the entire episode. In Table 5.3, I show the top six choices based on their proximity to emotional annotations. The number of players that chose the same response is listed along with the most

Table 5.3: Choices with emotion proximity. ChId represents Choice ID, Ch_[Button] represents the button used to select a particular choice label.

ChId	Count	Ch_X	Ch_Y	Ch_A	Ch_B
2-2	44	This is your last warning	[threaten him]	...	You're drunk
3-1	35	Sorry about the car.	How's your insurance?	...	Get off the street
7-5	35	My job	Don't need advice	...	Not my fault
2-7	33	HEY!	Will you excuse me a moment	...	[Throw him out]
19-16	31	Keep it	[Open it]	...	It belongs to Lawrence.
12-1	30	I'm fine	Fuck off	...	I'm not great.

popular choice. These reveal some problems with the straightforward approach of mapping affect annotations onto content: the scene where the top two options occur are artificially boosted due to the combat sequence that follows both. While using a time window may help, a type-based approach to dealing with player affect may be more effective than a single blanket time window.

Table 5.4: Unanimously Agreed Choice Prompts. DID represents Decision ID, ChID represents Choice ID, while Ch_[Button] represents button associated with choice option

ChID	DID	Ch_X	Ch_Y	Ch_A	Ch_B
13-1	2	Glamour	What do you want?	...	The car.
7-1	2	Yeah. Get out	C'mon, I'm tired.	...	There's only the one.
7-7	1	[Give Colin a Drink]	[Take Drink]		
19-9	4	What toy could have made this mark?	He left the toy for that long?	...	The broken lamp was here.
19-16	2	Keep it	[Open it]	...	It belongs to Lawrence.

For the final method, 5.4 tabulates a set of choices in which all 7 participants selected the same option. All players chose to give the character Colin their drink,

and all characters responded with the most polite response for Choice ID 13-1 and 7-1. The Decision ID represents the most popular choice as an index of the column of the selected choice. These agreed choices provide some insight into the overall nature of choices where the answer is obvious, as in the case where all participants decided to be nicer to Toad (rather than bring up the two previous incidents) as well as giving Colin the drink he desired. More interesting is that none of the players decided to follow the high road and not read the final note from Faith to her husband, choice 19-16, and instead opted to read it. Correlating player agreement and responses is another area worth investigating further.

5.4.3 Beat Analysis

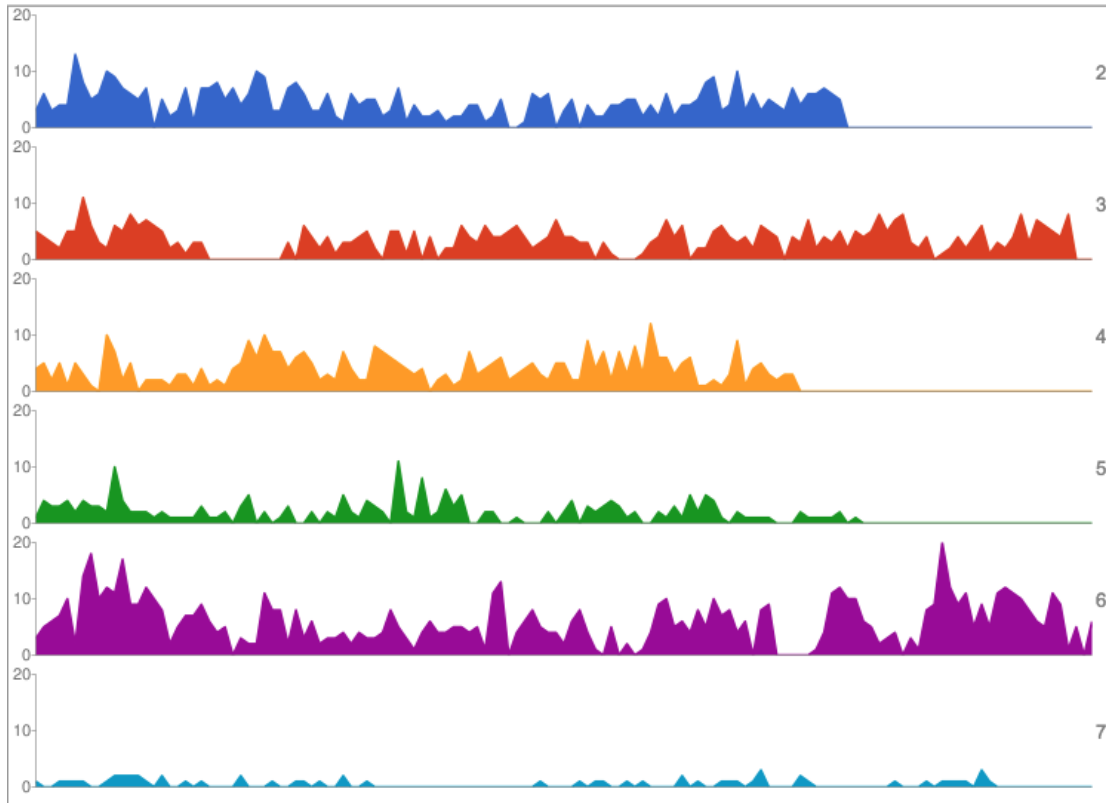


Figure 5.3: Sum Methods per Time Window

This section describes the benefits of using beats as an organizing content element and considers the trade-offs between self-report (such as SEI usage) and psychophysiological measurements.

Each player spends a different amount of time in each content segment due to the nature of gameplay. One method of analyzing player affective signals is to view each segment of time uniformly, as a sort of binning approach. This is shown in Figure 5.3, where any expressive event during each period of 60 seconds is summed in a time series. The chart shows both the variations in timing as well as the variations in expressiveness between the players. Where there are no data, players experienced a technical difficulty during the testing session, as indicated by the flat line on P3's graph. The variance in the timing made it clear that another approach to collating the data for analysis was required.

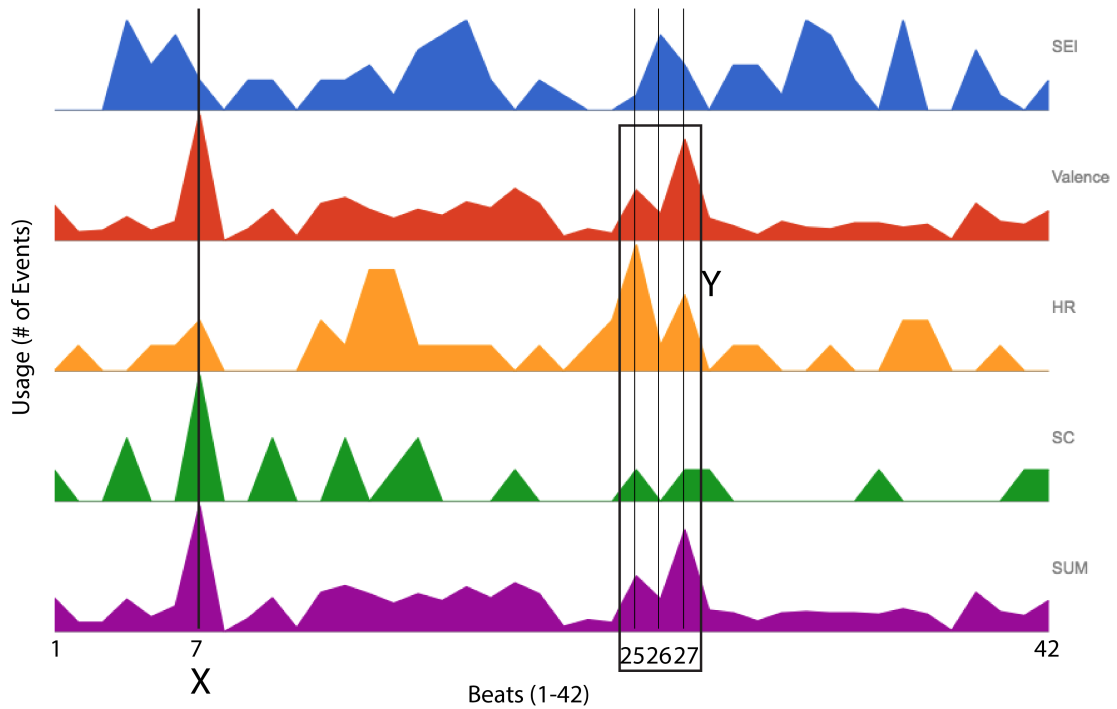


Figure 5.4: Sum of Each Method Per Beat

The next two analyses are based on segmenting the first three chapters into 42

beats, as described previously. Each beat was used to select events that occurred between the start and end timecodes. Each expressive event category is summed across all players and shown in Figure 5.4. The time window for annotations was shifted 8 seconds later from the start and end of the content annotations to account for delays in response to the content, although further fine tuning of this shift may be worthwhile. The x-axis represents an ordered sequence of beats from 1-42, whereas the y-axis represents the number of affect annotations in the time window for each beat. The full set of charts for each measure can be found in Appendix A. In this chart, beat 7 (noted as X) indicates when the Woodsman attacks Bigby in the game's first quicktime sequence. The encounter requires quick reflexes and exhibits a spike across the various physiological measures that reflects the tension involved. Beats 25-27, indicated as annotation Y, starts with the beat where Bigby argues with Colin, followed by the one where Snow retrieves Bigby while looking distressed, followed by the revelation that Faith was killed. These also showed clear physiological responses up to the revelation from players across all measures, with a dip in response measures during the suspenseful beat where Snow retrieves Bigby (beat 26). Discovering Faith's death is a key moment in the game and was carefully set up by the previous chapter, and this moment elicited sympathy from the players and foreboding at her fate.

The next chart breaks out each player's individual SEI usage 5.5. It revealed a spike (annotation A) for P3 when Snow retrieves Bigby. Unlike the physiological measures, a higher number of SEI usages usually indicates an increase in ambiguity, as the player is selecting multiple objects to indicate their feelings rather than using the same instrument multiple times as an indicator of intensity. P3 used three different SEIs in that beat, indicating a shift in feeling in one of the more suspenseful moments. P4 and P7 have spikes in their respective signals for SEI

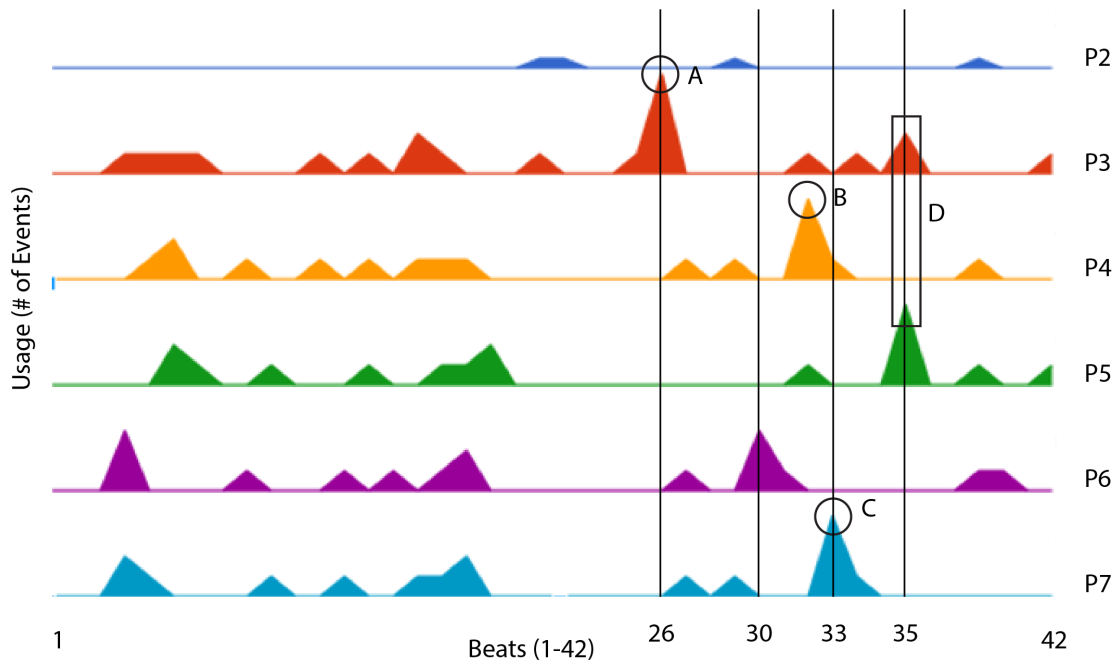


Figure 5.5: SEI Usage Per Beat

usage at beats 32 and 33 (annotation B and annotation C), during a tense scene with Mayor Crane. Another peak that occurs in the SEI signal is that of beat 35, where P5 and P3 both select multiple SEIs during a beat where the assistant flying monkey, Bufkin, is introduced. This scene follows the intense scene with Crane and, like the introduction of Colin and Toad, drives home the fantastical nature of the community and its inhabitants.

Future versions of this analysis should further discriminate between the different instruments used, although each player may end up assigning different meanings to different instruments. Additionally, the duration and nature of the SEI usage may be significant, although here we have only noted the location of the initial use. The nature of usage can further be understood using other statistical measures. I used each in a summative signal to achieve an overview of the space of potential affective events across multiple players, although breaking each sig-

nal into individual players assisted in locating the potential causes of the various features described. A more finely grained view of content can further distinguish between instruments used in response to one part of a beat or another.

I also used the timing information from the beats to analyze the position of the choice prompts in relation to their containing beats. Of the 48 choice prompts annotated that occurred within a story beat, six occurred in the first 20% of a containing beat, and eight occurred in the final 20% of a containing beat. The remaining 34 took place somewhere in between. This makes sense, as the characters (including the player character) have more of a chance to frame the decision, whereas an initial choice-prompt would put a lot of pressure on the narrative designers to shape the rest of the content of the beat based on the initial player choice.

5.4.4 Value analysis

The final analysis involves examining how players reacted to the main themes in the story. These are embedded throughout, and so simple ordered sequences of beats may not reveal a preference toward one value over another. I came up with the following list of binary-oppositions to describe the underlying force in each beat: Justice/Injustice, Ego/Community, Truth/Lies, Nice/Bad, and Duty/Hedonism. I could not classify two beats. These beats did not play a dramatic role in the sense of an active conflict involving a story value. One involves a character asking Bigby about his early guesses as to who the culprit might be, which functions to further solidify Snow's relationship and to pique the player's interest. The second is likewise an opportunity for the player to engage in speculation regarding the husband of the victim.

I labeled each beat with the story values and applied the same normalization

as was described in Section 5.1. These are shown in Figure 5.6.

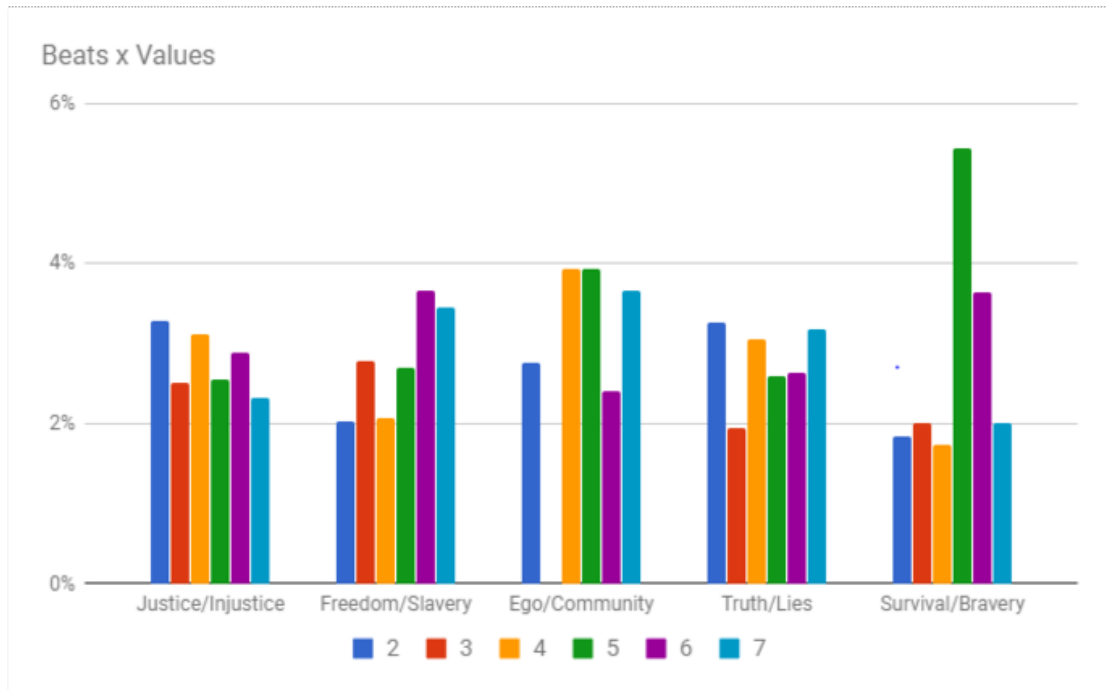


Figure 5.6: Value-oriented sorting. Each player’s relative affective response to each value is shown.

It is worthwhile to note that P5 stands out from the rest, with a strong reaction to beats 8, 14, and 19, all of which concerned survival or bravery. These beats involve Faith’s pursuit of her self-interest in the face of the Woodsman and the world at large.

In Figure 5.7, the distribution of story values is not normalized, and so the absolute relationship between the prominent themes of Truth/Lies and Justice/Injustice are contrasted with the less common beats that focus on the story values of Freedom/Slavery and Survival/Bravery. These story values are further developed later in the episode and the series.

The analysis yielded patterns that were not immediately apparent from watching the videos themselves, especially given the total length of the gameplay videos. It was difficult to remember each of the player’s responses to particular scenes,

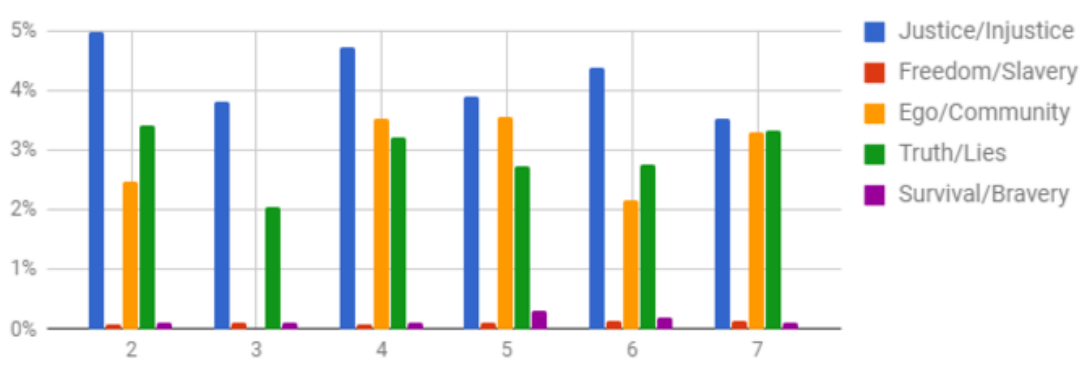


Figure 5.7: Player-sorted Value Beat Emotional Signals

and analyzing the response data required a closer inspection of the relevant scenes relating to abnormal spikes, such as the response to Colin. The other challenge was the difficulty of ensuring annotations corresponded across traversals and retained the relevant structural features. While I was interested in how themes were related to player responses to content, I did not create a label in the annotation schema for how choice texts or outcomes were related to the story values. This is left for future work.

5.4.5 Story Analysis

I recorded in each video key events and selected a time window of 25 seconds after to record the number of affective events correlated with each key event. The results are shown in a stacked bar chart in 5.8. This chart, unlike the others, is *not* normalized according to the player’s event frequency and therefore shows the contrast in the expressiveness of P6 relative to P7.

Player Affective Engagement by Story Event

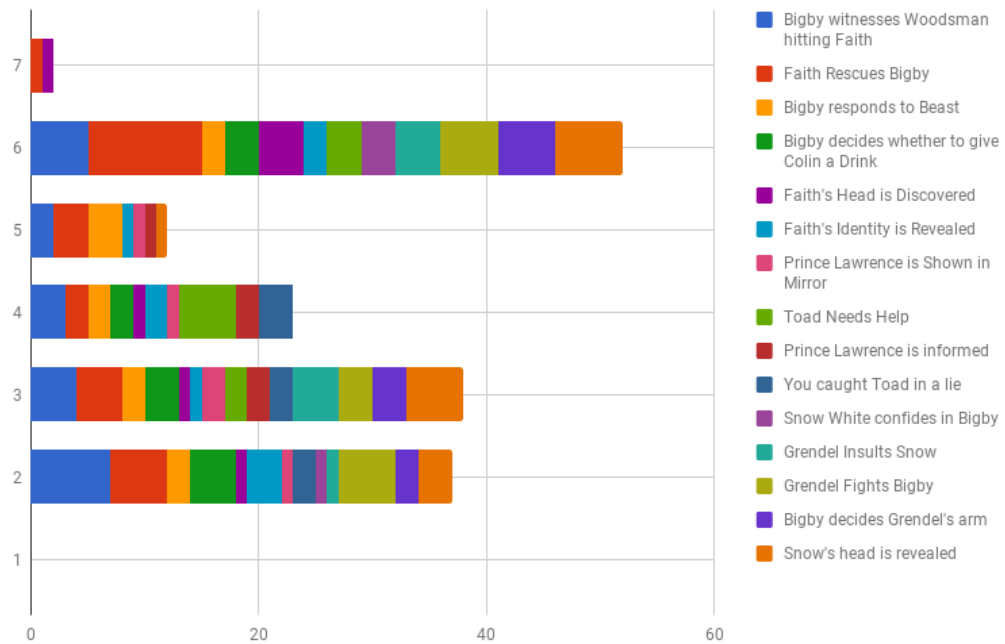


Figure 5.8: Player affect by event

5.5 Summary

The results presented in this chapter provide evidence that using various metrics in conjunction with a content-oriented schema can help guide not only assessment of traditional responses associated with games but also assist in interpreting the preferences that players exhibit with respect to certain classes of narrative content. While the granularity at this point is fairly coarse, this methodology does provide a means of comparing multiple players at once in a way that would be difficult to do using traditional analysis of player traversal videos, as it highlights both points of interest (through the points where SEI differed) as well as common patterns among all players (where the narrative was designed to elicit similar responses).

Chapter 6

Conclusions and Future Work

This thesis contributed evidence that combining measurable aspects of player response can highlight properties of the larger content structure, as well as indicate individual player preferences. The method of using content annotations to sort and compare player traversals can be seen as a promising direction to combine qualitative and quantitative measures of longer storygame experience records. The process of normalization described can be used to identify preferences for content despite varying distributions of time and events between players as well as player expressivity.

The study revealed several challenges for future work in relating content and player experience, and these were as follows:

1. The relationships between measures of player affect and content were implicit and difficult to disentangle with the current set of labels.
2. The summative method of analyzing the affective signal was coarse: it did not identify particular choices that were related to story values or their valence (positive or negative charges), nor did it allow for discriminating positively and negatively valenced emotional responses.

3. Questions that depended on patterns of content and player response were difficult to pose in the database representation.
4. The role that story value played in the mixed reception of content was observed but not in fine-enough granularity to identify potential causes.

Long-form narrative works are challenging to describe using a simple schema. The annotation schema described in this thesis provides one example of a set of criteria for breaking down content into segments and assigning labels; however, the exact labels and beat locations may be interpreted differently by another annotator, which could result in different segments and story values being selected. We are far from achieving a solution for the automated recognition of beats within the context of a modern adventure game. My annotations demonstrate evidence that even a small number of player traversals can be used to obtain insights when annotated.

The analysis inspired several new research questions, which provide a basis for future work: Were players consistent in their decisions concerning story values? Did the variations in content play a role in their level of engagement? Were specific content design patterns present that caused the effects observed? To what extent can such an analysis be automated, and what are the ethical implications of using such a method? Beginning to address these questions would require a fully annotated dataset and an annotation schema that can associate appropriate features of the content with points of interest in the traversals.

This thesis operationalized a set of labels and concepts to classify content and associate it with an affective signal derived from annotated player responses. Another set of binary-opposition pairs may have also suited, and, thus, certain aspects of this analysis must be necessarily qualitative and interpretative. Likewise, the exact specification of what individual practitioners and researchers consider

to be a beat in drama and narrative design is fluid and ill-defined and could use additional formalization in this context.

Storygames provide a rich genre to study for a better understanding of player responses to complex yet structured content. The subgenre of cinematic choice-based adventure games has enough regularity of content to enable the analysis of different player traversals, even in the presence of different decisions, in a comparative fashion. As the ability to extract more and more details from video data becomes possible, this type of analysis will become more useful but only if appropriate tools are developed to ease some of the challenges identified.

Appendix A

Beat Charts

Each component of the affective signal aggregate is broken out here. Each chart shows an array of players (P2-P7) and the number of events noted during each beat (1-42) along the x axis. For each, the peaks were noted through visual inspection of the values or indicators in the source video footage, which included indications from ATF for heartrate, skin conductivity, as well as facial expressions detected by the Affdex SDK.

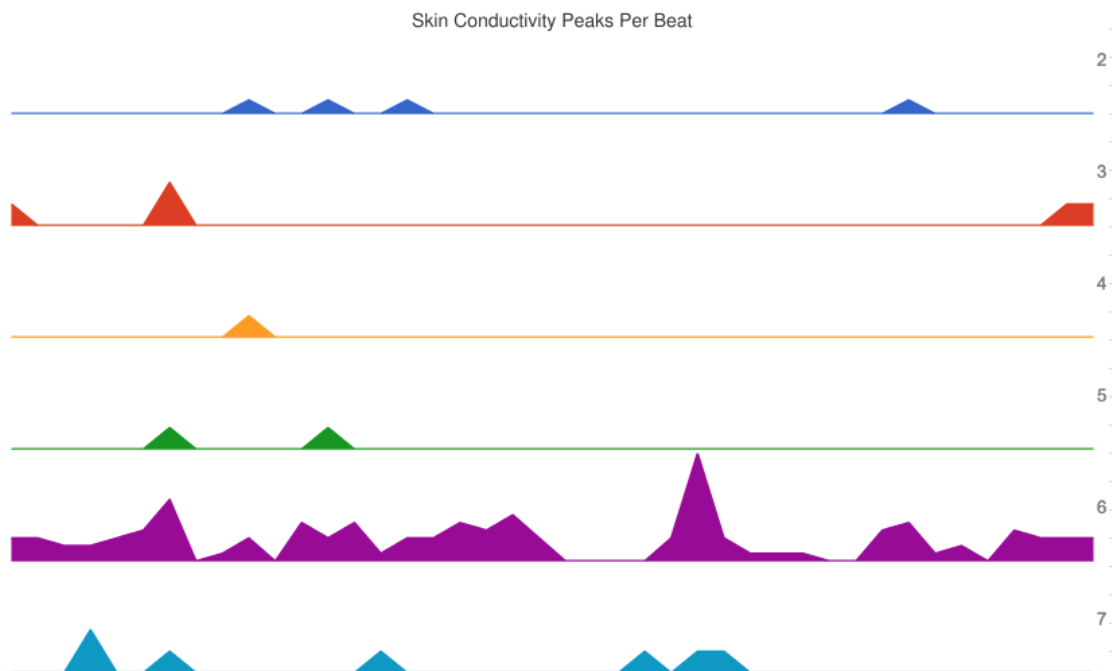


Figure A.1: Skin Conductivity Peaks Per Beat



Figure A.2: Heart Rate Peaks Per Beat

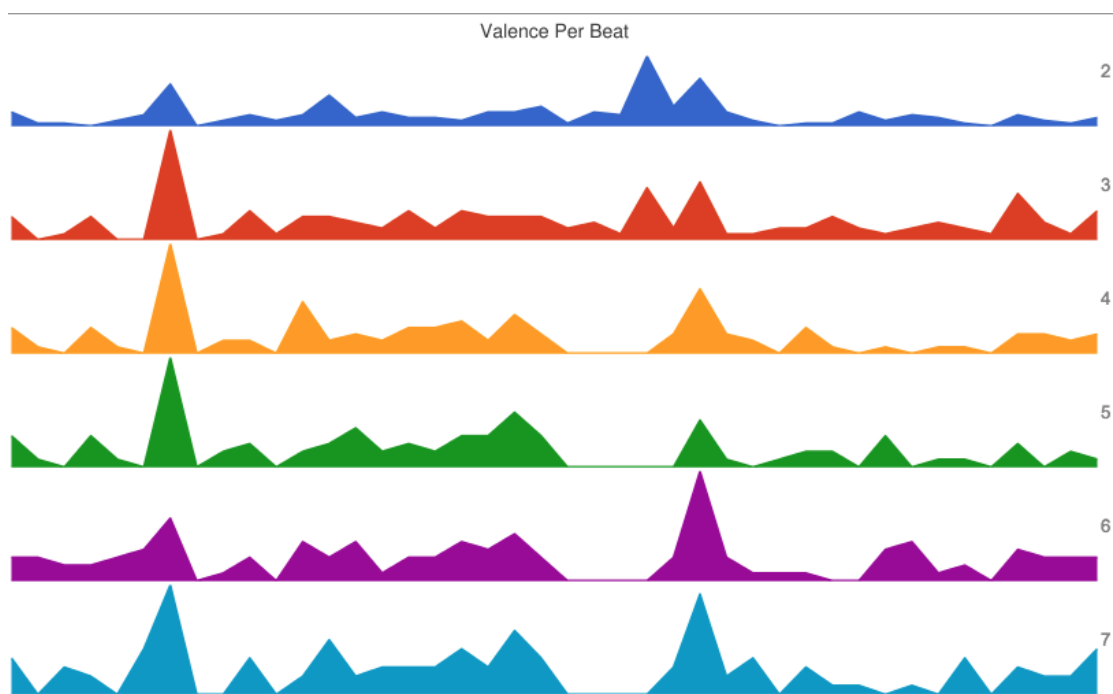


Figure A.3: Valence (Facial Expressions) Events Per Beat

Bibliography

- [1] L. Parker, “Telltale and the value of episodic content.” <https://www.gamespot.com/articles/telltale-and-the-value-of-episodic-content/1100-6405527/>, Mar. 2013. Accessed: 2017-9-14.
- [2] Telltale Games, “The wolf among us,” 2013.
- [3] R. Robinson, K. Isbister, and Z. Rubin, “All the feels: Introducing biometric data to online gameplay streams,” in *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, CHI PLAY Companion ’16, (New York, NY, USA), pp. 261–267, ACM, 2016.
- [4] K. Isbister, K. Höök, M. Sharp, and L. Jarmo, “The sensual evaluation instrument: Developing an affective evaluation tool,” in *CHI 2006 Proceedings*, pp. 1163–1172, 2006.
- [5] R. B. Robinson, *All the Feels: A Twitch Overlay that Displays Streamers’ Biometrics to Spectators*. PhD thesis, UC Santa Cruz, 2018.
- [6] P. A. Mawhorter, *Artificial Intelligence as a Tool for Understanding Narrative Choices*. PhD thesis, University of California, Santa Cruz, 2016.
- [7] J. L. G. Sánchez, F. L. G. Vela, F. M. Simarro, and N. Padilla-Zea, “Playability: analysing user experience in video games,” *Behaviour & information technology*, vol. 31, pp. 1033–1054, Oct. 2012.
- [8] H. Qin, P.-L. Patrick Rau, and G. Salvendy, “Measuring player immersion in the computer game narrative,” *International Journal of Human-Computer Interaction*, vol. 25, pp. 107–133, Feb. 2009.
- [9] C. T. Tan, T. W. Leong, and S. Shen, “Combining think-aloud and physiological data to understand video game experiences,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 381–390, ACM, Apr. 2014.

- [10] R. Marczak, J. van Vught, L. E. Nacke, and G. Schott, “Feedback-Based gameplay metrics: Measuring player experience via automatic visual analysis,” *Proceedings of The 8th Australasian Conference on Interactive Entertainment Playing the System - IE ’12*, pp. 1–10, 2012.
- [11] R. L. Mandryk, K. M. Inkpen, and T. W. Calvert, “Using psychophysiological techniques to measure user experience with entertainment technologies,” *Behaviour & information technology*, vol. 25, no. 2, pp. 141–158, 2006.
- [12] A. Drachen, “Behavioral telemetry in games user research,” in *Game User Experience Evaluation*, Human–Computer Interaction Series, pp. 135–165, Springer, Cham, 2015.
- [13] M. Mateas and A. Stern, “Integrating plot, character and natural language processing in the interactive drama facade,” in *Proceedings of the 1st International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE-03)*, 2003.
- [14] M. Mehta, S. Dow, M. Mateas, and B. MacIntyre, “Evaluating a conversation-centered interactive drama,” *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems - AAMAS ’07*, p. 1, 2007.
- [15] M. Seif El-Nasr, D. Milam, and T. Maygoli, “Experiencing interactive narrative: A qualitative analysis of facade,” *Entertainment computing*, vol. 4, no. 1, pp. 39–52, 2013.
- [16] B.-C. Bae and R. M. Young, “A computational model of narrative generation for surprise arousal,” *IEEE Transactions on Computational Intelligence in AI and Games*, vol. 6, pp. 131–143, June 2014.
- [17] B.-C. Bae and R. M. Young, “A use of flashback and foreshadowing for surprise arousal in narrative using a Plan-Based approach,” in *ICIDS ’08 Proceedings of the 1st Joint International Conference on Interactive Digital Storytelling: Interactive Storytelling* (U. Spierling and N. Szilas, eds.), vol. 5334 of *Lecture Notes in Computer Science*, (Berlin, Heidelberg), pp. 156–167, Springer Berlin Heidelberg, Nov. 2008.
- [18] Y. G. Cheong, *A Computational Model of Narrative Generation for Suspense*. PhD thesis, Raleigh, North Carolina, 2007.
- [19] D. L. Roberts, H. Narayanan, and C. L. Isbell, “Learning to influence emotional responses for interactive storytelling,” *AAAI Spring Symposium*., 2009.

- [20] M. Seif El-Nasr, D. Milam, and T. Maygoli, “Experiencing interactive narrative: A qualitative analysis of façade,” *Entertainment computing*, vol. 4, pp. 39–52, Feb. 2013.
- [21] B. Mallon and B. Webb, “Stand up and take your place: Identifying narrative elements in narrative adventure and role-play games,” *Comput. Entertain.*, vol. 3, pp. 6–6, Jan. 2005.
- [22] M. Cataldi, R. Damiano, V. Lombardo, and A. Pizzo, “Representing dramatic features of stories through an ontological model,” in *Interactive Storytelling*, pp. 122–127, Springer Berlin Heidelberg, 2011.
- [23] N. Szilas, “Requirements for computational models of interactive narrative,” in *Computational Models of Narrative. Papers from the 2010 AAAI Fall Symposium* (M. Finlayson, ed.), aaii.org, 2010.
- [24] M. A. Finlayson, “The story workbench: An extensible Semi-Automatic text annotation tool,” *Proceedings of the 4th Workshop on Intelligent Narrative Technologies (INT4)*, pp. 21–24, 2011.
- [25] D. K. Elson, “DramaBank: Annotating agency in narrative discourse,” *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12)*, pp. 2813–2819, 2012.
- [26] M. A. Finlayson, “A survey of corpora in computational and cognitive narrative science,” *Sprache und Datenverarbeitung (International Journal for Language Data Processing)*, vol. 37, no. 1-2, pp. 113–141, 2013.
- [27] F. Murtagh and A. Ganz, “Pattern recognition in narrative: Tracking emotional expression in context,” *JDMDH*, vol. 2015, 2015.
- [28] N. Szilas and I. Ilea, “Objective metrics for interactive narrative,” *The Seventh International Conference on Interactive Digital Storytelling (ICIDS 2014)*, pp. 91–102, 2014.
- [29] M. Wallentin, A. H. Nielsen, P. Vuust, A. Dohn, A. Roepstorff, and T. E. Lund, “Amygdala and heart rate variability responses from listening to emotionally intense parts of a story,” *NeuroImage*, vol. 58, pp. 963–973, Oct. 2011.
- [30] J. H. Kim, D. V. Gunn, E. Schuh, B. Phillips, R. J. Pagulayan, and D. Wixon, “Tracking real-time user experience (TRUE): A comprehensive instrumentation solution for complex systems,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’08*, (New York, NY, USA), pp. 443–452, ACM, 2008.

- [31] B. Medler, M. John, and J. Lane, “Data cracker: developing a visual game analytic tool for analyzing online gameplay,” in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 2365–2374, dl.acm.org, 2011.
- [32] R. Robinson, J. Murray, and K. Isbister, ““You’re giving me mixed signals!”: A comparative analysis of methods that capture players’ emotional response to games,” in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA ’18, (New York, NY, USA), pp. LBW567:1–LBW567:6, ACM, 2018.
- [33] C.-H. Hjortsjö, *Man’s Face and Mimic Language*. Studen litteratur, 1969.
- [34] E. Friesen and P. Ekman, “Facial action coding system: a technique for the measurement of facial movement,” *Palo Alto*, 1978.
- [35] B. Martinez, M. F. Valstar, B. Jiang, and M. Pantic, “Automatic analysis of facial actions: A survey,” *IEEE Transactions on Affective Computing*, vol. PP, no. 99, pp. 1–1, 2017.
- [36] T. Baltrušaitis, P. Robinson, and L. P. Morency, “OpenFace: An open source facial behavior analysis toolkit,” in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–10, Mar. 2016.
- [37] D. McDuff, A. Mahmoud, M. Mavadati, M. Amr, J. Turcot, and R. e. Kaliouby, “AFFDEX SDK: A Cross-Platform Real-Time Multi-Face expression recognition toolkit,” in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’16, (New York, NY, USA), pp. 3723–3726, ACM, 2016.
- [38] B. Farnsworth, “Facial action coding system (FACS) - a visual guidebook.” <https://imotions.com/blog/facial-action-coding-system/>, Dec. 2016. Accessed: 2018-3-26.
- [39] W. V. Friesen and P. Ekman, “EMFACS-7: Emotional facial action coding system,” *Unpublished manuscript, University of California at San Francisco*, vol. 2, no. 36, p. 1, 1983.
- [40] L. E. Nacke, “An introduction to physiological player metrics for evaluating games,” in *Game Analytics: Maximizing the Value of Player Data* (M. Seif El-Nasr, A. Drachen, and A. Canossa, eds.), pp. 585–619, London: Springer London, 2013.
- [41] F. Bousefsaf, C. Maaoui, and A. Pruski, “Remote assessment of the heart rate variability to detect mental stress,” in *Proceedings of the 7th International*

- Conference on Pervasive Computing Technologies for Healthcare, Pervasive-Health '13*, (ICST, Brussels, Belgium, Belgium), pp. 348–351, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013.
- [42] K. Alghoul, S. Alharthi, H. A. Osman, and A. E. Saddik, “Heart rate variability extraction from videos signals: ICA vs. EVM comparison,” *IEEE Access*, vol. 5, pp. 4711–4719, 2017.
 - [43] P. Mirza-Babaei, L. E. Nacke, J. Gregory, N. Collins, and G. Fitzpatrick, “How does it play better?: Exploring user testing and biometric storyboards in games user research,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, (New York, NY, USA), pp. 1499–1508, ACM, 2013.
 - [44] L. E. Nacke, “Games user research and physiological game evaluation,” in *Game User Experience Evaluation* (R. Bernhaupt, ed.), pp. 63–86, Cham: Springer International Publishing, 2015.
 - [45] G. N. Yannakakis and J. Togelius, “Experience-Driven procedural content generation,” *IEEE Transactions on Affective Computing*, vol. 2, pp. 147–161, July 2011.
 - [46] G. Chanel, C. Rebetez, M. Bétrancourt, and T. Pun, “Emotion assessment from physiological signals for adaptation of game difficulty,” *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 41, pp. 1052–1063, Nov. 2011.
 - [47] M. K. X. J. Pan, G. J.-S. Chang, G. H. Himmetoglu, A. Moon, T. W. Hazelton, K. E. MacLean, and E. A. Croft, “Galvanic skin response-derived bookmarking of an audio stream,” in *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, (New York, NY, USA), pp. 1135–1140, ACM, 2011.
 - [48] C. Bateman and L. E. Nacke, “The neurobiology of play,” in *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, pp. 1–8, ACM, May 2010.
 - [49] J. Laaksolahti, K. Isbister, and K. Höök, “Using the *Sensual Evaluation Instrument*,” *Digital Creativity*, vol. 20, no. 3, pp. 165–175, 2009.
 - [50] LANG and P. J., “International affective picture system (IAPS) : Digitized photographs, instruction manual and affective ratings,” *Technical report: NAVTRADEVCEEN. Naval Training Device Center*, 2005.

- [51] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton, “Measuring and defining the experience of immersion in games,” *Int. J. Hum. -Comput. Stud.*, vol. 66, pp. 641–661, Sept. 2008.
- [52] M. Mateas and A. Stern, “Writing facade: A case study in procedural authorship,” *Second Person: Role-Playing and Story in Games and Playable Media*, pp. 183–208, 2007.
- [53] R. Mckee, *Story: Substance, Structure, Style and The Principles of Screenwriting*. It Books, kindle ed., Dec. 1997.