Game or Watch: The Effect of Interactivity on Arousal and Engagement in Video Game Media

Joshua Juvrud, Gabriel Ansgariusson, Patrik Selleby, and Magnus Johansson

***Abstract*—This study examined arousal and engagement during video game play for interactive players and passive spectators who watched a recording of gameplay. We examined interactive (online) and noninteractive (offline) events in the game design. We collected psychophysiological data (eye movements, pupil dilation, heart rate) and self-reported measures (general anxiety, game experience, levels of fear). Overall, playing the game was associated with general increases in heart rate and a significant increase in the mean fixations distance. During offline events, players and spectators showed similar levels of arousal as indicated by pupil dilation. During online events, however, players showed increased pupil dilation—more so than someone who is watching the exact event without interactive control. These findings suggest that both players and spectators show increases in arousal and engagement, but there is a unique level of arousal and engagement for interactive play, as indicated by changes in heart rate and eye-tracking metrics. We discuss the application of the psychophysiological findings to recent trends in consumer media such as Twitch and YouTube, as well as the importance of game design, as it relates to human behavior.**

***Index Terms*—Arousal, engagement, eye tracking, game design, heart rate, interactivity, pupillometry.**

I. Introduction

VIDEO games are often perceived through a unique lens of interactive (or ergotic expression) [1]. Aarseth first described the nonpassive experience of video games, but the exact interactive nature of video games, and what is included in such a definition, continues to be a source of contention [1]. It is true that the majority of games contain interactive elements and indeed many scholars define video games by their interactive elements; whether it is through using a joystick to navigate a character or object, or selecting text options to progress a narrative. However, as games have advanced beyond several square pixels moving on a screen from the 1970s to more complex structured and segmented experiences, what qualifies as interactivity and its role in how a player experiences a video game have become more unclear. Moreover, qualifying what constitutes an interactive video game experience through the relationship between its interactive elements and the player fails to capture the effects of video games that are frequently enjoyed by those that might be considered as nonplayers, or "spectators," that exert no direct control via the game controls. Video games media have established its place in the sphere of spectator entertainment with the rising popularity of online streamers, e-sports, and let us plays (a play through of a game often paired with commentary that is uploaded online). While the arcades and living rooms have long offered spectator experiences, today millions of viewers login and subscribe to individuals who play video games for others' enjoyment on popular platforms such as Twitch and YouTube, offering a new way to experience a game like never before. In fact, at the time of this publication, there were 1.5 million live viewers on Twitch, nearly 50 000 live channels, and on average 2.5 million viewers per week [2]. In just the eight years since Twitch's launch, it has grown from 300 000 monthly broadcasters to 6.35 million, and from on average 100 000 viewers per month to now over 2 million. The content of the channels on Twitch is dominated by video games (see Fig. 1), with nearly 20 000 games being streamed per week.

The growth of streaming platforms and media featuring games, however, has been blurring the traditional lines between active and passive media consumption. With a huge increase in games being consumed by spectators, the interactive methods of experiencing a video game have been changing. New media platforms for video games have led to a rapid increase in the different ways that their content creators are engaging both players and spectators, breeding new kinds of interactivity and different degrees of control for both players and spectators. This raises a couple of important questions: 1) how do different degrees of interactivity effect player engagement; and 2) does having direct interactive control offer a larger degree of engagement than spectating? The answers to these questions could potentially help media platforms and content creators construct innovative ways to further engage both players and spectators and even help game designers decide how to balance interactive elements within the game.

A. State of the Art

The state of current research and industry trends suggests that the question of how interactivity effects player and spectator engagement a timely subject for examination. Specifically, a handful of prior studies have examined spectating in games and what makes spectating engaging and enjoyable [3]–[6]. This body of work suggests that, despite high levels of enjoyment

Manuscript received November 12, 2020; revised February 11, 2021; accepted April 9, 2021. Date of publication April 13, 2021; date of current version June 16, 2022. *(Corresponding author: Joshua Juvrud.)*

The authors are with the Games & Society Laboratory, Department of Game Design, Uppsala University, Visby 621 57, Sweden (e-mail: joshua.juvrud@psyk.uu.se; gabbish379@hotmail.com; patrik.selleby@student.uu.se; magnus.johansson@speldesign.uu.se).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TG.2021.3073084.

Digital Object Identifier 10.1109/TG.2021.3073084

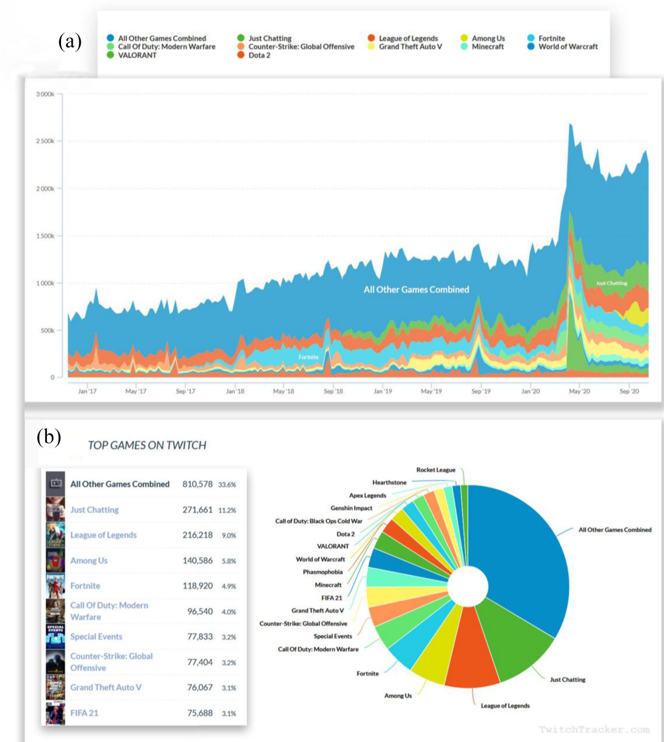

Fig. 1. (a) Top games on Twitch from January 2017 to September 2020. (b) Top games on Twitch during week 43 of 2020.

for both players and spectators, there may still be important differences in how players and spectators experience games. For example, a study using functional magnetic resonance imaging showed a selective sensitivity to playing versus watching first-person shooter games in the striatum region of the brain, an area associated with rewarding stimuli [4]. Other projects have taken these questions one-step further, examining the interaction dynamics and experiences between players and spectators [7]–[9]. In addressing the "in-between" design space [7] of players and spectators, both researchers and creators in the industry have been designing new and innovative ways to integrate the interactive player experience with the passive observer.

1) What Makes Spectating Engaging and Enjoyable?: There is extensive research that has relied on a distinction between active and passive consumption of media, such as the active consumption of video games and the passive consumption of television or sports [10]–[12]. These studies and their conclusions are often based on the theoretical assumption that when playing a video game, participant modeling is assumed, in which the person playing a video game virtually becomes the character of the video game. Due to this participant model, owing to the direct control over the character's behavior, a larger effect for something like arousal is often assumed in the player when playing a video game than a spectator watching television.

Previous work, which has largely relied on self-report and observational measures to explore spectator engagement and enjoyment [3], [5], [13], have identified key aspects as to why people enjoy the spectating activity in which they have no control. These studies have pointed to elements of the games and gameplay, such as the spectacle of graphics and a feeling of presence [14], [15], escapism [16], the emotions of the players [3], social and affective motivational factors [13], [17], [18], as well as both the building [3] and release [13] of tension. Information asymmetry, or the imbalance of information between the player and spectator, is another aspect that builds tensions and contributes to a feeling of suspense [3]. These studies contribute to our understanding of how spectators engage, and find enjoyment, in watching video games without directly interacting themselves. What is more, game designers are taking innovative steps to blur the line between player and spectator, and in some cases, integrating these experiences.

2) Spectator: Only a "Passive" Observer?: Both researchers and industry members have challenged what it means to be a passive observer. In games research, studies have explored the social relationship between spectators and players, such as how audience members have an impact on a player's performance [8], [9]. Additional work has integrated the spectator and player experiences, such as incorporating the spectator's eye gaze into a player's gameplay, changing the game experience for both the player and the spectator [7]. Such integrations have been suggested to be effective tools for treatments and interventions with special populations, such as children with ADHD, as a way to social motivation and collaboration [19].

Prior work notes the formative role of spectators in digital play, such as in the context of Twitch [17], [20] and eSports [3], [16]. In the games and tech industry, services that offer video games as spectator entertainment have designed innovative ways to increase spectator interaction with the game and game stream player. One of the most notorious examples of applying interaction with spectating was a global phenomenon and social experiment in February 2014 on the Twitch channel called Twitch Plays Pokémon [21]. In over 16 days of continuous streaming, up to 80 000 concurrent viewers (and over 1.6 million total participants) simultaneously controlled and completed the game Pokémon Red by inputting commands into the chat. The success of the event led to an increase in many more streams with interactive elements and led to the Twitch platform itself increasing interactivity options for streamers and watchers.

Indeed, platforms such as Twitch and Amazon now implement a wide range of interactive features, such as chat and emojis, creating and sharing polls, interactive overlays, live statistics, mouse click heatmaps, and ability to directly impact events in the game via chat inputs. Amazon's game engine Lumberyard, for example, comes equipped with native Twitch integration that allows for features in real time, such as spectators voting on game outcomes, chat commands, and the levels of the game actually changing and being affected by the number of viewers watching the player [22].

However, to what extent do different degrees of interactivity affect levels of engagement? After all, film and television have existed as forms of spectator entertainment for over half a century. Video game streaming, however, represents a unique kind of spectator entertainment where the events are largely unscripted, unedited, and the decisions are the result of a real live person who has control, while the spectator does not. What is more, platforms such as Twitch offers an average player the opportunity to broadcast their mundane, everyday experiences with a video game [23], in contrast to more polished video game media broadcasts like e-sports and YouTube channels. This knowledge that the spectator holds, that the player is not an actor following a script, but a regular person experiencing the game as they are, may create a unique sense of empathy and unpredictability that may result in levels of engagement qualitatively different to that of watching a film or television. Studies in television media have shown such differences when comparing scripted shows versus more unscripted shows, such as in reality television [24] or live sport events [25].

B. Online and Offline Engagement

An important aspect to note about games is that they are not always interactive, even for the player. Therefore, a remaining question is to what level of engagement a spectator, who is watching someone else control and experience a game through interactive input, is also affected by a spectator experience that does not have a traditionally interactive experience with the video game.

To help answer this question, we can examine players and spectators during scripted video game events in which the player has control of the actions on screen and scripted events in which that control is taken away (also referred to as cutscenes). Today, many games do not have a singular interactive experience. While they often have very intense segments that require some kind of input, they also contain segments of far more limited control or interactivity (looking at a map or reading text), or none at all (a cinematic sequence that breaks up the gameplay). Cutscenes offer advantages to game designers, such as through gameplay functions like transitioning the player to a new level that perhaps could not be achieved through other means [26]. Games are often structured into discrete events or sequences, and cutscenes can serve as a way to connect these into a coherent experience. These noninteractive experiences are often integral parts of a game; as previous studies have found, a cutscene can affect the rhythm of gameplay, providing moments of release from the action, or providing helpful information to the player [27].

Newman [27] described the state of interactive participation, or what we might simply think of as "playing the game," as the *online* state of the game. In contrast, *offline* engagement could be seen as equating with noninteractivity, periods where no registered input control is received from the player (such as a cutscene). Under this model, online and offline engagement are considered polar extremes of an experiential continuum. Games may require players to input degrees of commands during cutscenes (quick-time events) or provide responses to events that the player has little control over. Nonetheless, more polar offline segments in games, where the player is engaged with the action but not actually exerting direct control through the interface, remain a common tool for game designers that continue to be an integral part of players' experiences.

1) Cutscenes in Games: An additional question for game designers is therefore to consider how much time players themselves might spend during a game in sequences that requires some kind of active input when "playing" a video game. There are no hard and fast rules for the use of the cutscene in game design, or what clear benefits or disadvantages noninteractive moments can provide for the player. Furthermore, we do not know how cutscenes differ from a traditional passive observing of a game. Would a passive spectator watching someone else play a game on a streaming platform such as Twitch experience a similar level of engagement during a cutscene as the player? This lack of understanding comes from a lack of research that properly examines how both players and spectators respond to these moments of stripped interactivity.

One could argue, and indeed some have, that cutscenes are another form of interactivity [27], [28]. They are an integral part of the configurative experience. Even if the player is denied any active input, this does not mean that the interactive experience and effort are paused. In fact, cutscenes are often used to propel the player into a more dramatic or intense sequence, presumably building the tension or engagement of the player. It could be that cut-scenes provide a chance to focus attention between sequences of frenetic, high velocity, and volume play, and these sequences maintain interest and attention with frequent and often frantic suggestions, advice, and warnings. In contrast, it could be that by stripping control away from the player, the game is also breaking down an experience of engagement. Perron [29], for example, has suggested that using cutscenes may pull the player out of their "magic circle," or decrease their immersion in the game. He argued that a cutscene, or any form of noninteractive media, implicitly tells the audience that there is only one outcome and that it most often is not one where the player loses. By stripping control, the player has a sense of security that whatever happens is supposed to happen and that the player cannot fail. As a result, Perron suggested that we might get different reactions from players when comparing interactive and noninteractive scenarios. In a particularly arousing moment of gameplay (indicated by increased excitement and heart rate), a cutscene may unintentionally reduce overall levels of arousal for the player. How is a game designer to know when these scripted experiences, moments of limited or no control, are effective? It is important to know, as a designer, when active input is important, and when removing player control might actually disengage the player from the moment or impact player engagement.

C. Measuring Engagement

1) Arousal: Most of the prior work examining engagement in spectators relies on explicit self-report measures. However, response and sampling biases, the introspective ability for individuals to accurately assess themselves after spectating, and the lack of moment-to-moment information as to the changes in engagement, all present challenges for researchers and game designers. Research across several fields has shown that the automatic responses of spectators, or psychophysiological data, can be a valid representation of audience engagement [30]–[32]. This work demonstrates that arousal, as measured by biofeedback (e.g., galvanic skin response, heart rate, and eye-tracking), shows strong associations with explicit ratings of emotional reactions and engagement.

2) Eye Tracking: One way to measure player or spectator arousal with a game is through measuring the player or spectator's eyes. This is easily achieved through nonintrusive cameras that record an individual's eyes. We can extract two key measurements from the eyes that can give us insights into their cognitive processing: eye movements and pupil dilation. When the eyes move from point to point, these points are called fixations, where they remain stationary and take information in [33]. One prior study looking at games suggested that the number of fixations of a player may be an indicator of engagement, where eye movement during game play would increase as the subject was less engaged with the game on screen [34]. Other studies have suggested that an increase in the mean distance from one fixation point to another may be a reliable measure of cognitive processing [35]–[37].

Other studies in the field of psychology have shown that the pupil of the eye can be an important measure of arousal. According to Bradley *et al.*, test subjects showed further increased pupil dilation when exposed to emotionally positive and negative arousing images, compared to when exposed to emotionally neutral images [38]. Measurements of pupil change independent of luminance have been demonstrated to be a unique window into the mind, providing momentary, involuntary, and unbiased measures of arousal, attention, and cognitive load [32], [39]–[41]. Increased pupillary change has been shown to be a response to arousal, comparable with other measures in the peripheral nervous system, such as measures of heartrate and skin conductance response [32], [38], [42].

D. Current Study

This study examined player arousal levels during interactive (online) and noninteractive (offline) events within a horror-themed video game. The aim was to compare levels of arousal in players versus spectators, that is, individuals who watched a recorded video from the players, similar to watching a YouTube or Twitch recording. To measure arousal, we collected both psychophysiological data and subjective ratings of arousal during either play or spectating of a novel horror-themed game specifically designed for this study. The three measures of psychophysiological data we chose were eye movements, pupil dilation, and heart rate. We predicted that eye movements, measured as mean fixation distance from the center of the screen, would be an indicator of engagement with the scenario, where a higher mean distance from the center would suggest increased engagement because of more visual scanning.

II. Method

A. Participants

A total of 23 participants (males = 10, female = 11, nonbinary = 2; mean age = 23.88) participated from the Uppsala, Sweden area. One additional participant (male = 1) was tested but excluded from the final data set due to experimental error. Three participants' heart rate data were excluded from the final analyses ($n = 20$) due to technical error of the device. The target sample size was selected based on a power analysis and studies using similar experimental methods and analyses [32], [43], [44].

Fig. 2. Images of the (a) game environment and (b) the monster that chased the player.

All participants provided informed consent. This research was supported by a grant from the Swedish Research Council. The study was conducted in accordance with the standards specified in the 1964 Declaration of Helsinki and approved by the local ethics committee. All participants received a gift voucher.

B. Game Scenario

The current study used the horror genre to elicit strong arousal responses in participants. The experiment used a novel horror game specifically designed for this study (Sailor's Grave, Fig. 2). A horror-themed game was chosen because horror games are a specific type of game that are typically designed to create a sense of tension, anxiety, and arousal through tension and fear. Most horror games, through their design, have the aim of creating arousal in their players—particularly at key moments in the gameplay experience.

Horror games create a sense of engagement and arousal through tension and fear. Most horror games, by their design, have the aim of creating arousal in their players at key moments in the gameplay experience [45]. Some game designers choose to make arousing moments, such as encountering a monster, an online experience where the player must respond to a tense or startling event through inputs. Others choose to create these moments as offline cinematic cutscenes, where the player loses control and watches as the game plays out the proceeding events. Understanding the impact that this particular design choice has on the player is important for game designers who aim to evoke a particular state of arousal in their players.

The game was programmed in C++ using Unity 3-D real-time engine 4.22.0 (Unity Technologies SF, San Francisco, CA, USA) for application with a PC running Windows 10. The game was a first-person controlled character exploring a sunken shipwreck while a monster antagonist searched for the player and chased the player. The antagonist monster AI ran by using a behavior tree within the Unreal Engine See Fig. 2 for screenshots of the game.

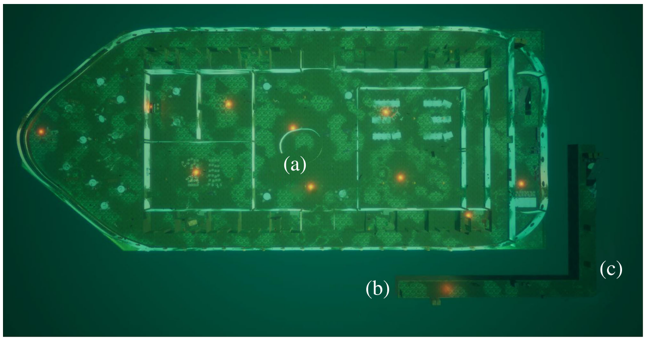

Fig. 3. Image of the game map with (a) marked player starting position, (b) appearance of the online shark event, and (c) offline antagonist event. The red dots indicate the transmitters that the player could collect.

Fig. 4. Player and spectator room setup.

During the gameplay, two key events were chosen for data analysis (see Fig. 3 for a map of the game environment). The first was a scripted event that was triggered when the player turned the first corner down a narrow corridor. A large shark would appear and swim past the player, accompanied by audio of high-pitch string instruments playing, lasting approximately 3 s. This event was online, in that the player maintained full control during this event. The second event occurred when the player would approach one central transmitter in the center of the map, which triggered a scripted offline cutscene in which the player loses control for 5 s. During the cutscene, the monster would suddenly appear from behind the player, grab the player, and then launch them across the room, lasting approximately 3 s. Player control would then resume. Timestamps were logged for each event allowing for the synchronization of game events and testing measures. An oxygen meter served as a timer and limited each gameplay session to 10 min.

Before running the experiment, we conducted a quality control of the playable scenario ($N = 5$). The aim of the quality control was to ensure that the game scenario was played as intended, that there were no major bugs, and that the game induced arousal (evaluated through self-reports).

For testing, half of the participants were placed in the interactive condition and half were in the spectator condition.

C. Conditions

1) Interactive Condition: The gameplay consisted of collecting 11 tokens (“transmitters”) hidden in the ship while also avoiding a monster that roamed and chased the player. The player won the game when they collected all 11 tokens. The player would lose if they were hit twice within 5 s by the monster’s weapon, in which case they would start again from the beginning. The player controlled the player avatar using a keyboard and mouse. The player pressed and held the E key (prompted on screen) for 2 s while in the proximity of a transmitter.

2) Spectator Condition: The gameplay film clips in the spectator condition were taken from the prior participant’s recorded gameplay, therefore making them identical to the interactive scenario in both content and time, only removing player control and interactivity from the experience. The motivation for letting the spectating participants watch recorded sessions of the interactive condition was to give the spectators the same outcomes as the players.

D. Room Setup

Participants in both conditions sat roughly 60 cm in front of a 52’ monitor (see Fig. 4). The room was unlit with blackout curtains. The participants were wearing a pair of headphones with directional sound. The researchers sat at a table behind the participants in the same room but did not interact with the participants during testing.

E. Measures

1) Mean Fixation Distance: We measured participants’ eye movements using an eye tracker. We used a Tobii 4C Eye tracker with Tobii Pro Lab software for data collection. A prior study on immersion using eye tracking found that the more a player of a game gets immersed, their eyes tend to move less [34]. Therefore, we tracked the mean fixation distance during each gameplay session from the center of the screen for each group.

2) Pupil Dilation: We recorded the pupil dilation of the participants as a measure of arousal. We used two events: the online shark encounter and the offline forced encounter with the antagonist. During the online encounter, participants would turn a corner in a narrow corridor, where a giant shark would suddenly appear and swim at the player. During the offline encounter, whenever the player picked up the selected transmitter, this would cause the antagonist to appear behind the player and grab them forcing them to face the antagonist, lift them up and then throw them in a straight line. During this scripted event, the player lost control in the interactive condition, which always occurred after the player picked up one of a few selected transmitters.

3) Heart Rate: Based on previous studies that have found a relationship between arousal intensity and heart rate [46], we calculated the Δ heart rate between participants' resting heart rates and their heart rates during the scenario as an additional measure of arousal. Participants wore a heart monitor wristwatch (Activity armband, INF company AB) that recorded their heart rate. Prior to testing, we measured each participant's resting heart rate for 3 min. The participants would then have their heart rate measured during the entire length of the scenario. We synchronized the heart monitor with an android smartphone via bluetooth and used the application FlagFit 2.0 to remotely monitor the heart rate. Prior studies have shown this method to be a valid measure of heart rate [47].

4) *Self-Report Measures:*

a) *General anxiety:* We measured participants' levels of general anxiety using the GAD-7 screening tool to measure the test subjects' general anxiety ($\alpha = 0.92$) [48]. The self-report questions were answered on a range from 1 to 5. Responses were aggregated to form a general anxiety score.

b) *Gaming Experience:* In the scenario quality control, we noticed that the amount of previous gaming experience affected how people played the game scenario. We, therefore, suspected that gaming experience could potentially affect the arousal of the participants. As a result, we added a question that assessed prior video game experience ("On a scale from 1 to 5, how experienced with computer games do you consider yourself?"). The test subjects rated their subjective gaming experience on a Likert scale from 1 to 5, with 1 being "not at all experienced" and 5 being "very experienced."

c) *Horror Familiarity:* We measured participants' subjective familiarity with horror media ("On a scale from 1 to 5, how familiar with horror media do you consider yourself?"). Participants rated their subjective horror familiarity from 1 to 5, with 1 being "not familiar at all" and 5 being "very familiar."

d) *Levels of Fear:* After testing, we measured participants' subjective fear, which is an emotion associated with arousal ("On a scale from 1 to 5, how scared would you say that you were during the scenario?"). The test subjects rated their subjective fear during the scenario from 1 to 5, with 1 being "not scared at all" and 5 being "very scared."

F. Procedure

The participant first entered the room and gave informed consent. The experimenter explained the procedure to the participant and was then seated at a table. The participant completed the pretest questionnaire, which included the GAD-7 questionnaire for general anxiety. Participants were then seated at another table in the back of the room and the researcher fitted the heart monitor bracelet and measured the participant's resting heart rate. The participant was then moved to the TV and a researcher explained the nature of the game scenario. If the participant was in the interactive condition, the researcher would then inform them of the game controls and the goal of the game. If the participant was in the spectator condition, they were explained the goal of the game. The researcher then calibrated the eye-tracker for the participant using a standard 9-point calibration procedure [49]. Testing began as soon as the game scenario started. Participants in the interactive condition would play the game either until they won, or until they had played for 10 min. Participants in the spectator condition would watch one entire recording from the previous participant. Once the participant was finished with the testing, the researcher would then bring the participant back to the table where they would complete the postquestionnaire, which included the subjective measure of fear, subjective gaming experience, and subjective horror familiarity.

G. Data Analyses

For pupil dilation data, Tobii Pro Lab software automatically placed event markers for analyzing the pupil dilation of the participants during the shark encounter and antagonist forced encounter events within the scenarios. The dependent variable for the event was the difference between mean pupil size during the 5 s following the onset of the event and mean pupil size during the baseline period. The baseline period began 1000 ms prior to the onset of each event, which was identical across participants. The baseline period was controlled for luminance by ensuring that all participants had the same image on screen prior to each analysis period (both events were triggered only after a particular moment in gameplay that was identical across all players). Analysis occurred during the 5 s following the events. All events were visually inspected for normal pupillary light reflex response. The analysis was performed in the open-source analysis program TimeStudio version 3.03 running in MATLAB version 7.12 (www.timestudioproject.com) [50]. We used a moving-average filter and gap interpolation across all data. The pupil dilation mean of the participants in the interactive condition was then compared to the pupil dilation mean of the participants in the spectator condition using SPSS version 24.

The mean fixation distance was analyzed using the Tobii fixation filter by comparing the average screen distance from the center of screen of all participants watching the scenario and all participants playing the scenario from start to finish of the game scenario.

The participant's average resting heart rate was subtracted from their average scenario heart rate, resulting in a heart rate Δ (change). The mean Δ heart rate of the players was then compared to the mean Δ heart rate of the spectators.

We had two within-subjects independent variables of interest, Encounter (shark or antagonist) and Interactivity (online and offline), and one between-subjects independent variable, Condition (interactive or spectator). The outcome measures of interest included mean pupil changes to the specific encounters and mean fixation distance and change in heart rate over the entire session. Separate repeated measures ANOVAs with Greenhouse-Geisser correction were used for the pupil measures to study the interaction between encounter, interactivity, and condition, and paired t-tests were used to test the comparisons. See Table I for a summary of mean statistics.

III. Results

Across all players, 80% succeeded at winning the game with an average number of 0.50 deaths per game and an average playtime of 8 min and 5 s. Using Pearson's correlation coefficient, we found no relationship between success, number of deaths, or

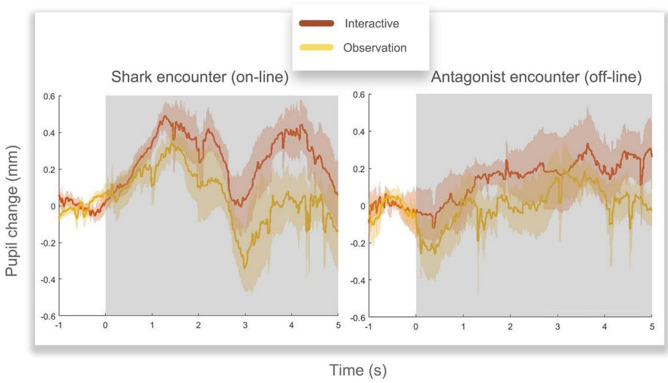

Fig. 5. Change in pupil size from baseline during the online shark encounter for the interactive and spectator conditions.

playtime with any of the investigated dependent variables and were therefore not included in final the analyses.

A. Pupil Dilation

We examined differences in pupil dilation between participants in the interactive and spectator conditions. See Table I for a summary of participant measurements. In the online event in which participants encountered a shark, there was a significant difference between playing and spectating during the shark encounter, $t(19) = 2.35$, $p = 0.031$, with more pupil dilation for participants in the interactive condition than the spectator condition (see Fig. 5). In contrast, there was no significant difference between playing and spectating during the offline forced antagonist encounter ($p > 0.50$).

B. Eye Movements

There were no differences between the interactive and spectator condition in the total number of fixations, number of fixations during the online event, or number of fixations in the offline event. We did find, however, significant differences in the average fixation distance between conditions. We compared the mean fixation distance across all participants in the interactive and spectator condition and found an increase in the fixation distance from the center in the interactive participants compared to the spectating participants ($p = 0.009$).

C. Heart Rate

There was a significant difference in the change in heart rate between the interactive and spectator condition, $t(16) = 2.56$, $p = 0.021$. The average heart rate change (Δ) from baseline to playing the game in the interactive condition was higher than change from baseline in the spectator condition. Heart rate was calculated as an average during the entire session. Due to pupil dilation being measured during key events, we could not examine the relation between heart rate and pupil dilation.

D. Self-Report Measures

See Table II for a summary of participant responses. We conducted Pearson's correlation coefficient to examine the statistical relationship among our variables. When examining self-report measures, we found that across all participants regardless of condition, there was a negative correlation between the game experience and how scared they were during the scenario ($r = -0.55$, $p = 0.01$). That is, the more experienced participants were with playing games the less fear they reported during the session. For participants in the spectator condition, there was a significant negative correlation between game experience and change in heart rate ($r = -0.93$, $p = 0.006$). That is, the more game experience the spectator had, the lower their mean heart rate during spectating sessions. For participants in the interactive condition, change in heart rate during play was significantly correlated to increased levels of self-reported anxiety ($r = 0.68$, $p = 0.04$).

TABLE I
Summary of Descriptive Statistics

	Condition	**M**	**SD**
On-line event pupil change (shark encounter)	Interactive	0.28	0.15
	Spectator	0.07	0.25
Off-line event pupil change (antagonist encounter)	Interactive	0.11	0.26
	Spectator	0.01	0.21
Scenario heart rate	Interactive	84.81	10.39
	Spectator	78.28	9.42
Heart rate change (Δ)	Interactive	8.09	6.81
	Spectator	-1.00	8.12
Total Fixations	Interactive	1468	425
	Spectator	1370	387
On-line event fixations	Interactive	29.63	5.18
	Spectator	30.54	5.35
Off-line event fixations	Interactive	25.60	8.35
	Spectator	27.50	4.99
Fixation distance	Interactive	600	296
	Spectator	515	273

Means and standard deviations in each condition for measures of: change in pupil dilation during the online and offline events (mm), heart rate over the entire scenario, the Δ (change) between the participant's average resting heart rate and their average heart rate during the scenario, and the participants mean fixation distance from center during the scenario (pixels).

TABLE II
Summary of Self-Report Measures

	M	**SD**
General anxiety (GAD-7)	6.85	3.88
Interactive	7.66	4.84
Spectator	6.18	2.96
Amount of game experience	3.38	1.46
Interactive	3.20	1.39
Spectator	3.54	1.57
Levels of fear during scenario	2.86	0.96
Interactive	3.00	1.39
Spectator	2.75	0.86
Horror genre familiarity	2.73	1.32
Interactive	2.36	1.20
Spectator	3.08	1.37

IV. General Discussion

While both playing and spectating a video game can be both engaging and arousing, it is evident that it is a uniquely arousing and engaging experience particularly for individuals who have interactive control during online events. When video game players have interactive control during a potentially arousing event (online event), they show increased pupil dilation in response to that event—more so than someone who is watching the exact same event without interactive control. However, this does not mean that spectating is not an engaging and arousing experience. If an arousing event is presented through a scripted cutscene (offline event), players and spectators both show similar (increased) levels of arousal. It is just that adding an element of control significantly increases levels of arousal in players.

When examining the eye movements of players, no differences were found in the number of fixations, but instead, we found a significant increase in the mean fixations distance for those that were playing the game compared to spectators. In addition, playing the game was associated with general increases in heart rate. Together, these findings suggest that online interactive experiences may provide increased levels of arousal, as indicated by the change in pupil size. Increases in fixation distance and increases in heart rate may also suggest higher levels of engagement for those playing the game versus watching.

Video games have the capacity to affect both players and spectators, but the findings from the current study suggest that having input control creates a unique increase in arousal and engagement. These findings are consistent with prior studies using biofeedback, which suggest that levels of heart rate are associated with levels of cognitive engagement in video games [35]–[37]. Koutepova *et al*. [36] have shown that over time, players adjust to the stimuli of the game and thus decrease engagement, as reflected through heart rate. The results from the current study suggest that one method for adjusting player biofeedback and modifying engagement is through managing player interactivity during play.

In another study, Jennett *et al*. [34] used the total number of fixations to measure the engagement of player during an "immersive" and "nonimmersive" game. They showed that during the immersive game the number of fixations decreased, and therefore suggested an individual's eye movements will decrease as their attention becomes more focused during an immersive task. In the current study, we did not find a difference between players and spectators in the number of fixations. This does not mean that Jennett *et al*.'s findings were not supported; it is likely that the current game was immersive and engaging enough for both players and spectators. Unlike Jennett *et al*., the current study did not have the equivalent of a "nonengaging" condition. Indeed, as discussed in the introduction, the phenomenon of spectating games on Twitch is a highly engaging activity [17], [23]. What we did find, however, is that the mean distance of fixations from the center of the screen was much higher for players than spectators. This may suggest a qualitatively different kind of engagement for players. Studies that have examined visual scanning patterns and information processing have shown that differences in fixation patterns related to length and distance reflect increased cognitive processing during tasks [51]–[54]. While both players and spectators were engaged with the media, players were likely showing increased cognitive load and engaging in more information processing while playing than an individual spectator. This may be comparable to an "active" versus "passive" type of engagement [55], and related to an increased sense of empathy and unpredictability with being a spectator to the unique, unscripted nature of video game players. More research is needed, however, to test this interpretation.

In the offline event, as soon as the antagonist forced encounter starts, player interactivity is removed. In the online event, when the player encounters the shark, interactivity remains. We find that the offline event shows similar levels of pupil dilation between players and spectators, while the online event shows an increase in pupil dilation uniquely for the players. Change in pupil dilation is a reliable indicator of levels of arousal, meaning that we can conclude from the current study that players are showing a heightened level of arousal during the online event when they maintain interactive control.

The self-report data from the current study reveal several important findings for content creators and game designers. First, the negative correlation between the game experience and the self-reported fear suggests that an unfamiliarity with games added an increased self-reported arousal in the game. It could be that unfamiliarity with video games increases the unpredictability when spectating by not knowing what to expect, and as a result, an increase in self-reported arousal. In line with this finding, self-report measures showed that higher levels of game experience were associated with decreased heart rate for spectators. These results were not evident in the actual players. This is interesting because it suggests that the interactive component of games, and the arousal that comes from increased cognitive engagement, is a more powerful determinant of biological responses than prior experience with games and self-reported levels of arousal. No matter an individual's experience with games, if the game is sufficiently engaging and arousing, it will be arousing and engaging for the majority of players. Furthermore, certain individual differences in players are still important. This is demonstrated by the association between general anxiety and change in heart rate amongst game players.

A. Limitations and Future Directions

This study is one of the first to use pupil dilation as a measure of arousal in players and spectators. As such, there are a few limitations worth noting and directions for future research.

The correlations among self-reported data in the current study suggest an important role of individual differences, such as with general anxiety and change in heart rate. It is worth noting that it is possible that higher levels of general anxiety in participants could contribute to increased heart rate and eye movements. More research is needed examining how individual differences might interact with the effects of games on individual levels of engagement and arousal.

In a study by Yannakakis *et al.* [56], they describe how biofeedback measures, such as heart rate, can be used to distinguish among different types of media and their entertainment value. The findings here not only support using biofeedback as an indicator of qualitative experiences with a game but also demonstrate through pupil dilation that we can measure and compare discrete time-locked events within a game, such as online and offline events. The results of the current study further support Yannkakis *et al*.'s suggestion that biofeedback and psychophysiological feedback may have the additional potential to be used dynamically within the game design to adjust in real-time the features of a game to the player. Both scripted online and offline events could react to a player's pupil dilation or heart rate in real-time. Therefore, future studies should take advantage of the time-frequency data that is offered by data from the pupil.

Due to perceptual differences between the online and offline events, we could not directly compare the pupil change within conditions. When examining changes in pupil size, controlling for perceptual differences such as light are necessary for direct comparisons. We did, however, control for these factors between conditions and found significant effects for playing versus spectating. Additionally, the game created for the current study reflects just one of many artistic styles, aesthetics, and types of gameplay. It is possible that these findings may not generalize across all types of games.

As previously mentioned, online and offline engagement as defined by Newman [27] are polar opposites on a continuum. It would be interesting to examine to what degree of interactivity is needed to create an increase in arousal. For example, would the popularly used "quick-time events" (where players input simple commands during cutscenes) be enough to increase arousal. One could also create conditions for measuring the level of engagement in a game and connect that to self-reported levels of perceived engagement from the respondents.

It is possible that other indirect forms of interactivity, such as player face-cams and spectator chat with players, may increase arousal for individuals watching someone else play. The phenomenon of streamers playing games for others to watch is a popular form of entertainment on sites such as Twitch or YouTube, and a common practice for these streamers is to have a camera filming them as they play to showcase their reactions. Future studies should examine how being able to see the player and the player's emotional responses might influence individuals watching someone play a game (for example, by providing an overlay that displays biometric and webcam-derived data for observers [57]). It is possible that, even across the internet, facial emotion cues will increase similar levels of emotional responses, or empathy, and thus increased levels of arousal, in those watching. Similarly, differences between commentated and uncommentated play sessions would be another interesting aspect to study. These are just a few of many features that are popular on streaming platforms, and more work is needed to explore how chat features, donations, spectator inputs, and other innovative interactive features affect both players and spectators. For game designers, future research should specifically examine the range of interactivity that exists between online and offline events, as interactivity is always on a continuum.

V. Conclusion

This work contributes to both our understanding of arousal and engagement when both playing and spectating games. This is the first study to measure arousal with both a change in the pupil dilation and heart rate in both video game players and spectators. The fact that the pupil and heart rate showed similar differences for both players and spectators supports prior work suggesting autonomic arousal as a valid reflection of engagement [30]. More specifically, a novel contribution of this study was the finding that both spectators and players show similar patterns of autonomic arousal when interactivity is stripped from the players during cutscenes. In this study, we have also found a connection between online and offline engagement, and we have also used a method to actually examine such a connection.

Altogether, the findings presented in the current study further demonstrate that levels of arousal for players and spectators have potential implications for media creators who have a vested interest in creating arousing and engaging experiences for their audiences. This work opens up possibilities for future studies that can further examine different conditions for engagement. Using pupil dilation and fixations as metrics, both content creators and game designers can reliably measure, and possibly manipulate players' levels of arousal and engagement with game media. Whether it is through interactive features on streaming platforms or design choices related to levels of interactivity, these elements have clear psychophysiological effects on game players and spectators.

References

[1] E. J. Aarseth, *Cybertext: Perspectives on Ergodic Literature*. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1997.
[2] Twitch Tracker, "Twitch statistics & charts," Accessed: Oct. 21, 2020. [Online]. Available: https://Twitchtracker.com/statistics
[3] G. Cheung and J. Huang, "Starcraft from the stands: Understanding the game spectator," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, May 2011, pp. 763–772.
[4] J. Kätsyri, R. Hari, N. Ravaja, and L. Nummenmaa, "Just watching the game ain't enough: Striatal fMRI reward responses to successes and failures in a video game during active and vicarious playing," *Front. Hum. Neurosci.*, vol. 7, pp. 1–13, 2013.
[5] N. T. Taylor, "Now you're playing with audience power: The work of watching games," *Crit. Stud. Media Commun.*, vol. 33, no. 4, pp. 293–307, 2016.
[6] M. Palaus, E. M. Marron, R. Viejo-Sobera, and D. Redolar- Ripoll, "Neural basis of video gaming: A systematic review," *Front. Hum. Neurosci.*, vol. 11, pp. 1–40, 2017.
[7] B. Maurer, I. Aslan, M. Wuchse, K. Neureiter, and M. Tscheligi, "Gaze-based onlooker integration: Exploring the in-between of active player and passive spectator in co-located gaming," in *Proc. Annu. Symp. Comput.-Hum. Interaction Play*, Oct. 2015, pp. 163–173.
[8] B. Maurer, M. Lankes, B. Stiglbauer, and M. Tscheligi, "EyeCo: Effects of shared gaze on social presence in an online cooperative game," in *Proc. Int. Conf. Entertainment Comput.*, Sep. 2016, pp. 102–114.
[9] S. Shahid, E. Krahmer, and M. Swerts, "Video-mediated and co-present gameplay: Effects of mutual gaze on game experience, expressiveness and perceived social presence," *Interact. Comput.*, vol. 24, no. 4, pp. 292–305, 2012.
[10] S. Natkin, *Video Games and Interactive Media: A Glimpse at New Digital Entertainment*. Boca Raton, FL, USA: AK Peters/CRC Press, 2006.
[11] H. Polman, B. O. de Castro, and M. A. van Aken, "Experimental study of the differential effects of playing versus watching violent video games on children's aggressive behavior," *Aggressive Behav., Official J. Int. Soc. Res. Aggression*, vol. 34, no. 3, pp. 256–264, 2008.

[12] N. S. J. M. Schutte, P.-J. C. G. Malouff, and A. L. Rodasta, "Effects of playing videogames on children's aggressive and other behaviors 1," *J. Appl. Social Psychol.*, vol. 18, no. 5, pp. 454–460, 1988.

[13] M. Sjöblom, and J. Hamari, "Why do people watch others play video games? An empirical study on the motivations of twitch users," *Comput. Hum. Behav.*, vol. 75, pp. 985–996, 2017.

[14] P. Skalski, R. Tamborini, A. Shelton, M. Buncher, and P. Lindmark, "Mapping the road to fun: Natural video game controllers, presence, and game enjoyment," *New Media Soc.*, vol. 13, no. 2, pp. 224–242, 2011.

[15] S. Joeckel and N. D. Bowman, "Graphics and gratification: Exploring the link between technology and enjoyment in video games," *J. Gaming Virtual Worlds*, vol. 4, no. 1, pp. 25–43, 2012.

[16] J. Hamari and M. Sjöblom, "What is eSports and why do people watch it?," *Internet Res.*, vol. 27, pp. 211–232, 2017.

[17] Z. Hilvert-Bruce, J. T. Neill, M. Sjöblom, and J. Hamari. "Social motivations of live-streaming viewer engagement on twitch," *Comput. Hum. Behav.*, vol. 84, pp. 58–67, 2018.

[18] W. A. Hamilton, O. Garretson, and A. Kerne, "Streaming on twitch: Fostering participatory communities of play within live mixed media," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Apr. 2014, pp. 1315–1324.

[19] K. Arrambide *et al.*, "The development of 'orbit' the collaborative BCI game for children with AD (H) D," in *Proc. Extended Abstr. Annu. Symp. Comput.-Hum. Interaction Play Companion Extended Abstr.*, Oct. 2019, pp. 341–348.

[20] T. L. Taylor, "The rise of massive multiplayer online games, esports, and game live streaming," *Amer. J. Play*, vol. 12, no. 2, pp. 107–116, 2020.

[21] D. Ramirez, J. Saucerman, and J. Dietmeier, "Twitch plays pokemon: A case study in big g games," in *Proc. DiGRA*, 2014, pp. 3–6.

[22] Amazon web services (n.d.). Amazon lumberyard. Retrieved Feb. 1, 2021. [Online]. Available: https://aws.amazon.com/lumberyard/details/

[23] T. L. Taylor, *Watch Me Play: Twitch and the Rise of Game Live Streaming*. Princeton, NJ, USA: Princeton Univ. Press, 2018.

[24] R. M. Huff, *Reality Television*. Westport, CT, USA: Greenwood Publishing Group, 2006.

[25] B. Buraimo, D. Forrest, I. G. McHale, and J. D. Tena, "Unscripted drama: Soccer audience response to suspense, surprise, and shock," *Econ. Inquiry*, vol. 58, no. 2, pp. 881–896, 2020.

[26] D. Říha, "Cutscenes in computer games as an information system," *in Proc. Int. Conf. Des., User Exp., Usability*, Jun. 2014, pp. 661–668.

[27] J. Newman, "The myth of the ergodic video game," *Game Stud.*, vol. 2, no. 1, pp. 1–17, 2002.

[28] R. Klevjer, "In defense of cutscenes," *in Proc. CGDC Conf.*, 2002.

[29] B. Perron, "Horror video games," in *Essays on the Fusion of Fear and Play*. Jefferson, NC, USA: McFarland, 2009.

[30] C. Latulipe, E. A. Carroll, and D. Lottridge, "Love, hate, arousal and engagement: Exploring audience responses to performing arts, " in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, May 2011, pp. 1845–1854.

[31] S. Steiner, E. Abbruzzese, R. L. Marca, and U. Ehlert, "Autonomic stress responses elicited by watching a live broadcast soccer game: A pilot study," *Gazzetta Medica Ital. Archivio per le Scienze Mediche*, vol. 172, no. 6, pp. 443–448, 2013.

[32] J. Juvrud *et al.*, "The immersive virtual reality lab: Possibilities for remote experimental manipulations of autonomic activity on a large scale," *Front. Neurosci.,* vol. 12, pp. 1–10, 2018.

[33] M. F. Land, "Eye movements and the control of actions in everyday life," *Prog. Retinal Eye Res.*, vol. 25, no. 3, pp. 296–324, 2006.

[34] C. Jennett *et al.*, "Measuring and defining the experience of immersion in games" *Int. J. Hum.-Comput. Stud.*, vol. 66, no. 9, pp. 641–661, 2008.

[35] R. M. Flynn and R. A. Richert, "Cognitive, not physical, engagement in video gaming influences executive functioning," *J. Cogn. Develop.*, vol. 19, no. 1, pp. 1–20, 2018.

[36] T. Koutepova, Y. Liu, X. Lan, and J. Jeong, "Enhancing video games in real time with biofeedback data," in *Proc. ACM SIGGRAPH ASIA Posters*, pp. 1–1, 2010.

[37] H. Monkaresi, N. Bosch, R. A. Calvo, and S. K. D' Mello, "Automated detection of engagement using video-based estimation of facial expressions and heart rate," *IEEE Trans. Affect. Comput.,* vol. 8, no. 1, pp. 15–28, 2016.

[38] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, "The pupil as a measure of emotional arousal and autonomic activation," *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 2008.

[39] S. E. Kuchinsky *et al.*, "Pupil size varies with word listening and response selection difficulty in older adults with hearing loss," *Psychophysiology*, vol. 50, no. 1, pp. 23–34, 2013.

[40] T. Piquado, D. Isaacowitz, and A. Wingfield, "Pupillometry as a measure of cognitive effort in younger and older adults," *Psychophysiology*, vol. 47, no. 3, pp. 560–569, 2010.

[41] P. van der Wel and H. van Steenbergen, "Pupil dilation as an index of effort in cognitive control tasks: A review," *Psychon. Bull. Rev.*, vol. 25, no. 6, pp. 2005–2015, 2018.

[42] M. J. Wieser, P. Pauli, G. W. Alpers, and A. Mühlberger, "Is eye to eye contact really threatening and avoided in social anxiety?—An eye-tracking and psychophysiology study," *J. Anxiety Disorder*, vol. 23, no. 1, pp. 93–103, 2009.

[43] D. T. J. Knoepfle, T. Y. Wang, and C. F. Camerer, "Studying learning in games using eye-tracking," *J. Eur. Econ. Assoc.,* vol. 7, no. 2–3, pp. 388–398, 2009.

[44] K. R. Segal and W. H. Dietz, "Physiologic responses to playing a video game," *Amer. J. Dis. Child.*, vol. 145, no. 9, pp. 1034–1036, 1991.

[45] B. Perron, (Ed.). *Horror Video Games: Essays on the Fusion of Fear and Play.* Jefferson, NC, USA: McFarland, 2014.

[46] A. Azarbarzin, M. Ostrowski, P. Hanly, and M. Younes, "Relationship between arousal intensity and heart rate response to arousal," *Sleep,* vol. 37, no. 4, pp. 645–653, 2014.

[47] A. G. de M. Zambotti, S. Claudatos, I. M. Colrain, and F. C. Baker, "A validation study of fitbit charge 2^{TM} compared with polysomnography in adults," *Chronobiol. Int.,* vol. 35, no. 4, pp. 465–476, 2018.

[48] R. L. Spitzer, K. Kroenke, J. B. Williams, and B. Löwe, "A brief measure for assessing generalized anxiety disorder: The GAD-7," *Arch. Intern. Med.,* vol. 166, no. 10, pp. 1092–1097, 2006.

[49] G. Gredebäck, S. Johnson, and C. von Hofsten, "Eye tracking in infancy research," *Devlop. Neuropsychol.,* vol. 35, no. 1, pp. 1–19, 2009.

[50] P. Nyström, T. Falck-Ytter, and G. Gredebäck, "The timestudio project: An open source scientific workflow system for the behavioral and brain sciences," *Behav. Res. Methods,* vol. 48, no. 2, pp. 542–552, 2016.

[51] K. Rayner, "Eye movements in reading and information processing," *Psychol. Bull.,* vol. 85, no. 3, pp. 618–660, 1978.

[52] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychol. Bull.,* vol. 124, no. 3, pp. 372–422, 1998.

[53] M. Ryan, N. Krucien, and F. Hermens, "The eyes have it: Using eye tracking to inform information processing strategies in multi-attributes choices," *Health Econ.,* vol. 27, no. 4, pp. 709–721, 2018.

[54] M. Shojaeizadeh, S. Djamasbi, and A. C. Trapp, "Density of gaze points within a fixation and information processing behavior," in *Proc. Int. Conf. Universal Access Hum.-Comput. Interaction*, 2016, pp. 465–471.

[55] N. P. Pitterson, S. Brown, J. Pascoe, and K. Q. Fisher, "Measuring cognitive engagement through interactive, constructive, active and passive learning activities," in *Proc. IEEE Front. Educ. Conf.*, Oct. 2016, pp. 1–6.

[56] G. N. Yannakakis, J. Hallam, and H. H. Lund, "Entertainment capture through heart rate activity in physical interactive playgrounds," *User Model. User-Adapted Interaction,* vol. 18 no. 1–2, pp. 207–243, 2008.

[57] R. Robinson, Z. Rubin, E. M. Segura, and K. Isbister, "All the feels: Designing a tool that reveals streamers' biometrics to spectators," in *Proc. 12th Int. Conf. Foundations Digit. Games*, Aug. 2017, pp. 1–6.