AFTRS WHITE PAPERS

3

Biometrics: An AFTRS Applied Innovation Research Project

Author: Penelope Thomas (2018)

AFTRS Australian Film Television and Radio School

Contents

The Art and Science of Biometrics: Eight Pilot Studies on the Use of Biometrics across Screen and Broadcast for a Mixed Methodological Approach to Audience Engagement

Executive Summary

Biometrics technology enables researchers to measure physiological and emotional responses to stimuli in real time using biosensors. Although research in the field is proven and mature in academia (Jain, Nandakumar, & Ross, 2016), the film industry has only more recently locked onto its enormous potential as a way to improve entertainment products and create highly engaging content. NBC, Netflix and AMC/Sundance currently run biometric testing on their products, and Danish Broadcaster DR have used it successfully to inform their television series The Bridge, The Killing and Borgen.

In 2015, the Australian Film, Television and Radio School (AFTRS) set out to provide this powerful source of information for the Australian media arts. The AFTRS Applied Industry Research Unit worked with the platform provider iMotions to establish a Biometrics Lab and test its use in a variety of ways across screen and broadcast practices. The research outcomes would strengthen the School's innovation initiative and be shared with industry for the development of new business models and improved screen and broadcast content.

More recently, AFTRS partnered with Screen Audience Research Australia (SARA) to conduct Industry-commissioned projects for the Australian Broadcasting Commission. This research further refined how biometrics could be used for audience engagement in meaningful ways by content creators. The research team developed a mixed methodological approach that can be embedded into the editorial process of measureable media content.

This paper presents highlights from the studies and evaluates the use of biometrics and its data for audience engagement research across a variety of industry needs, including film, television, broadcast, casting, pitching, television commercials (TVCs), trailers and genre studies.

Summary of Findings

Research findings support the use of biometric measure for audience engagement in the Australian media arts. The data provides very specific information at the granular level: recommendations for a television pilot can be made about audience preferences, character development, storyline incongruities and dialogue misfires, to inform not only a pilot episode but an entire series.

A biometrics approach to content development enables content creators to refine their offering to the consumer rather than operating on traditional audience testing and what might be considered low-level instinct. The innovative aspect of this research is the design of a mixed methodological approach that can be embedded into the editorial process of measureable media content, combining qualitative and quantitative methods, and using biometrics as a diagnostic tool.

Introduction

Biometrics is a technology that enables researchers to measure physiological and emotional responses to stimuli in real time using biosensors. Danish broadcaster DR have used it successfully to inform their television series The Bridge, The Killing and Borgen (see Image 1). In 2015, the Australian Film, Television and Radio School (AFTRS) set out to provide the same powerful source of information for the Australian media arts. The AFTRS Applied Industry Research Unit worked with the platform provider iMotions to establish a Biometrics Lab and test its use in a variety of ways across screen and broadcast practices.

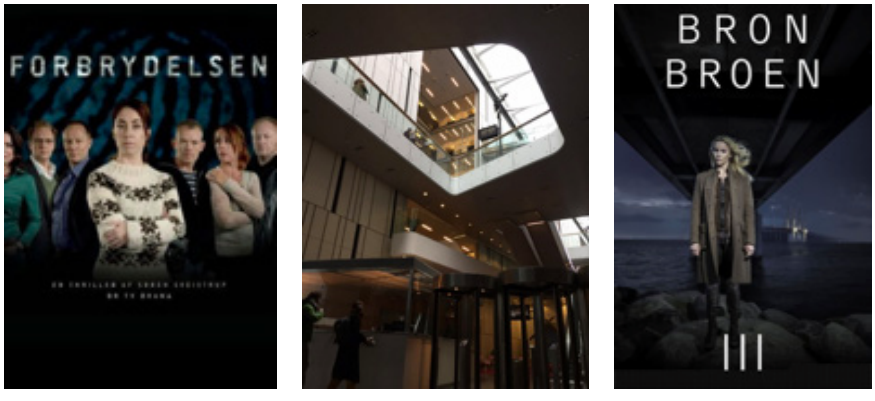


Image 1: Danish Group DR interior (centre) and biometrically informed television shows The Killing (left) and The Bridge (right).

More recently, AFTRS partnered with Screen Audience Research Australia (SARA) and conducted industry-commissioned projects on the pilot episodes of two national television dramas. This collaboration helped to refine the way biometrics could be used for audience engagement in meaningful ways. The team developed a mixed methodological approach that can be embedded into the editorial process of measureable media content, combining qualitative and quantitative methods, and using biometrics data as a diagnostic tool.

This paper presents eight studies that evaluate the use of biometrics and its modalities for a range of industry needs. The research outcomes will inform the School's curriculum and be shared with industry for the development of new business models and improved screen and broadcast content.

Ethical Considerations

AFTRS aligns with the Biometrics Institute and its global network to ensure best practice. Biometrics technology is an incredibly powerful tool, yet it also brings risks that must be acknowledged. The biometrics tracking technology being used by AFTRS is not invasive from a medical standpoint; however, physiological responses to stimuli are monitored and response data collected. There are no known risks with the processes and procedures being used for this research.

The Technology

Biometrics technology enables highly accurate testing of audience engagement by measuring physio-emotional responses to stimuli in real time using biosensors. With the iMotions platform, all sensor, stimuli and API data streams are real-time synchronised. Patterns in audience responses are quantified and data can be accessed at the individual participant level or aggregated across a sample group. The software provides a variety of data visualisations frame by frame or across a stimuli timeline.

The AFTRS projects use a combination of three biosensors simultaneously: galvanic skin response (GSR), eye-tracking and facial expression analysis. A self-reportage survey can also be built into the study (see Image 2).

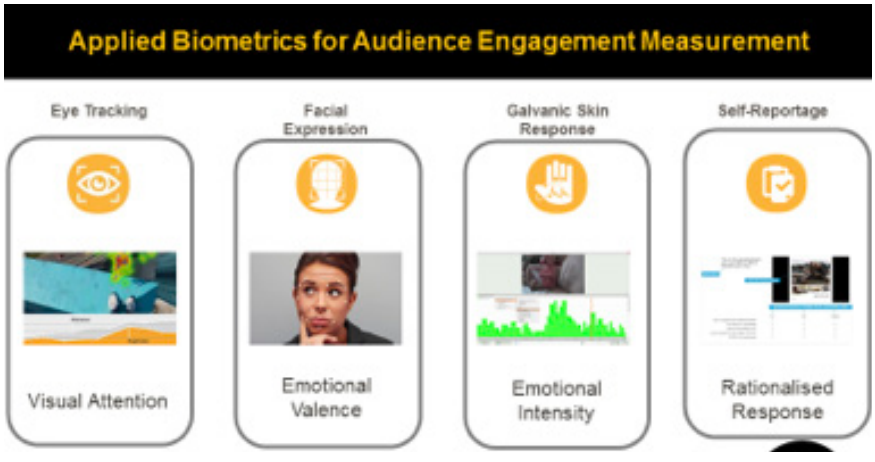


Image 2: AFTRS/SARA researchers use a combination of three biosensors and an optional embedded survey.

1. Galvanic Skin Response

Galvanic Skin Response (GSR), is a widely used method of research on human psycho-physiological phenomena (Boucsein, 2014). It measures heart rate and skin conductance, providing insights into emotional arousal, stress and engagement. The hardware is worn on the wrist/hand of research participants (see Image 3). It houses one sensor that measures heart rate and another two that transmit a low constant voltage between electrodes to measure electro-dermal activity or skin conductance (essentially, slight variations in sweat). Humans have around three million sweat glands, and as the primary interface between an organism and its environment, the skin is an incredibly dynamic and powerful source of information.

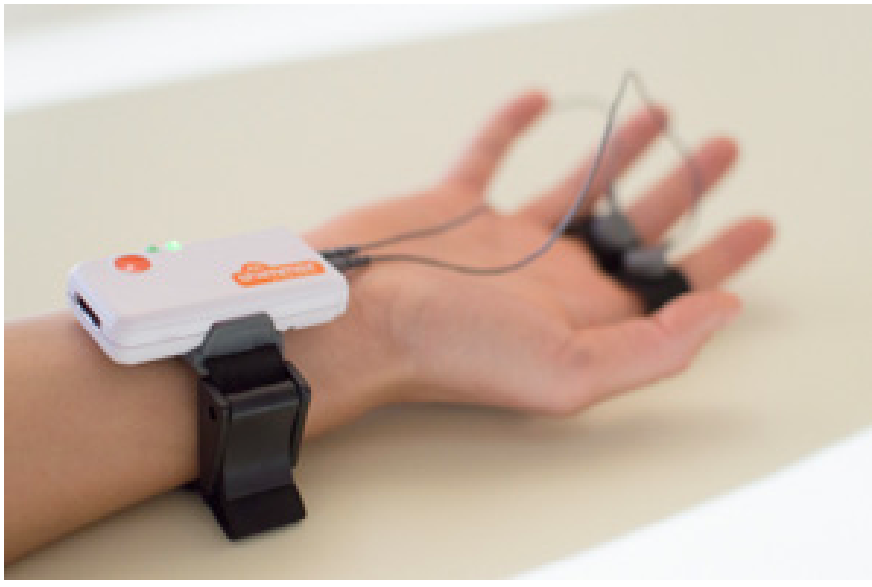


Image 3: GSR sensors measure heart rate and electro-dermal activity.

GSR technology for audience engagement generates quantitative data about autonomous, unconscious responses to what audiences are experiencing (see Image 4). Across a sample group, distinctive patterns in arousal can be quantified statistically. Metrics are in amplitude over time (uS), measuring peak onsets and offsets at around 0.5–1.0 seconds gathered at the tonic and phasic level. The metrics reflect four events: Latency, Peak Amplitude, Rise Time and Recovery Time. With audience engagement measure, researchers are interested in event-related GSR peaks against the timeline of the content.

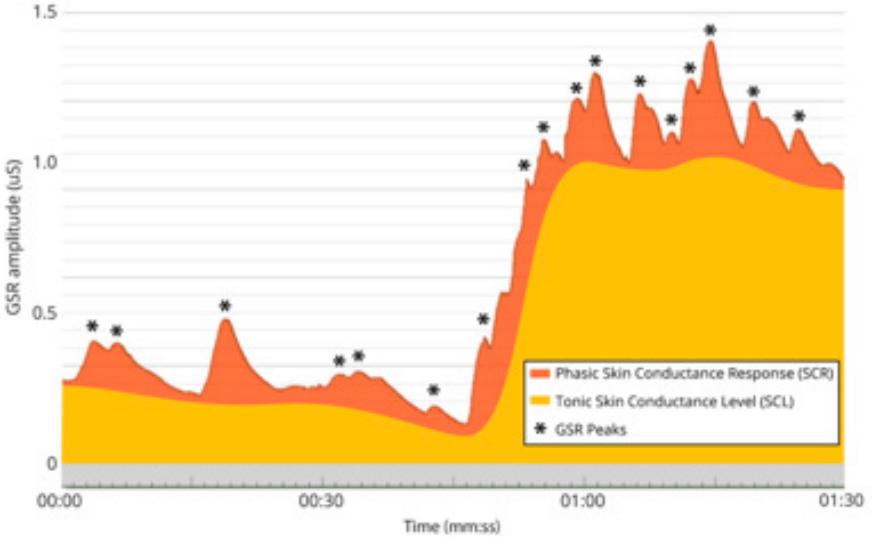


Image 4: GSR metrics in amplitude over time of phasic and tonic skin conductance levels.

GSR provides an empirical read of levels of engagement; however, the limitation is that GSR does not indicate the valence or type of engagement: [arousal = intensity], whereas [valence = quality]. Valence is an important concept in biometric research. To demonstrate, classifications can be visualised along two or more dimensions, with valence (positive versus negative emotions) as a horizontal axis and arousal (activating versus calming emotions) as a vertical axis (see Image 5). More nuanced emotional classifications are possible' for example, 'happiness' can be classified into a less aroused 'happy' state or a more aroused 'elated' state.

Facial expressions are core indicators of underlying emotional states and provide information about valence that GSR does not. This is why a multi-modal approach using a combination of biometric technologies is important.

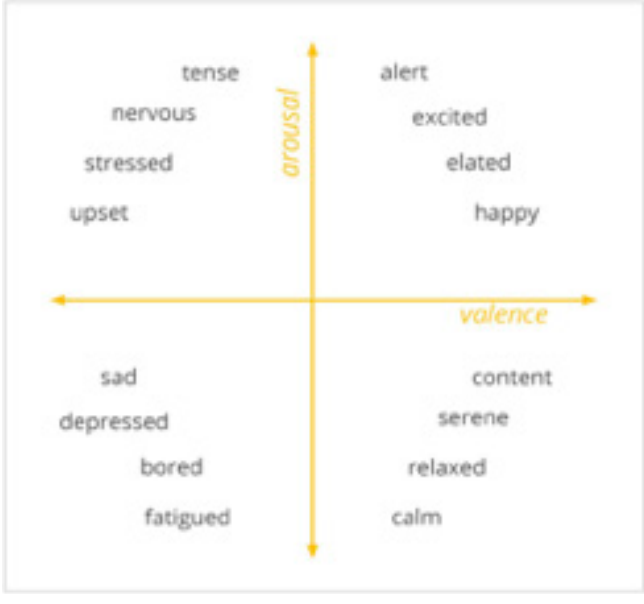


Image 5: Visualisation showing classifications of Valence (quality) and Arousal (intensity).

2. Facial Coding Analysis

Facial coding technology taps into something ancient and evolutionary within the same region of the brain that controls emotional processing and regulation: a complex system that allows for the exchange of important social information. Ekman and Friesen (1981) tested Darwin's (1872) universality hypothesis, and found a small set of distinctive facial configurations that most people associate with certain emotions irrespective of social, cultural, age and/ or gender factors. There are seven emotions in the set: joy, surprise, anger, contempt, sadness, disgust and fear (see Image 6).

Image 6: The seven facial configurations universally recognised as representing seven emotions.

Biometrics technology uses the Facial Action Coding System, or FACS (Ekman, Davidson, & Friesen, 2002) of 33 facial landmarks or feature points that map macro-expressions and micro-expressions. At between 0.5 to 4 seconds, macro-expressions are visible to the human eye, whereas at less than 0.5 seconds micro-expressions are unable to be seen by the naked eye. Metrics include: the Seven Basic Emotions; Valence — positive, negative and neutral sentiment; Emotion Channels, such as smile, lip pucker, brow furrow, eye closure, nose wrinkle chin raise and smirk; Head Orientation; and Interocular Distance.

In combination with other biosensors such as GSR and eye-tracking, facial coding analysis provides important information about the conscious feelings, emotions and moods likely to be accompanying unconscious engagement. The data represents probabilistic scores from [0 = no expression] to [100 = fully present expression], with the option to set absolute and relative thresholds. The aggregation of data across a sample group can be generated on the actual values, or on the 'thresholded' and binarised signal of individual respondents.

The limitation of facial expression analysis is that facial coding metrics are probabilistic, based on the likelihood that the expression captured is representative of the emotion being experienced.

3. Eye-Tracking

Eye-tracking hardware is usually attached to the screen being viewed by a participant. The technology measures pupil centre corneal reflection (PCCR), or visible reflections in the cornea, the outermost optical element of the eye, caused by near-infrared light directed at the pupil. The technology generates a variety of data for insights into visual attention and interest in real time (see Image 7).

Image 7: Eye-tracking technology generates a variety of metrics on visual attention in real time.

Eye-tracking metrics include:

- Pupil Size (dilation): indicating emotional arousal and cognitive workload.
- Ocular Vergence: indicating engagement where the distance between the pupils diverge when the attention attenuates.
- Gaze Fixation: measure gaze points from 100–300 milliseconds, indicating an item of interest.
- AOIs: quantitatively generates a heat map visualising areas of interest, where red is the highest.
- TTFF: measures time to first fixation.
- Distance to Screen: measures avoidance/approach behaviours; for example, participants moving closer to the screen indicates engagement.

The eye-tracking technology generates visual representations of data that can be viewed at the individual participant level or as an aggregate across a sample group (see Image 8). These include gaze replay and dynamic visualisation; static and dynamic areas of interest; heat maps showing colour-coded gaze distribution; and gaze plots indicating what participants look at and in what order.

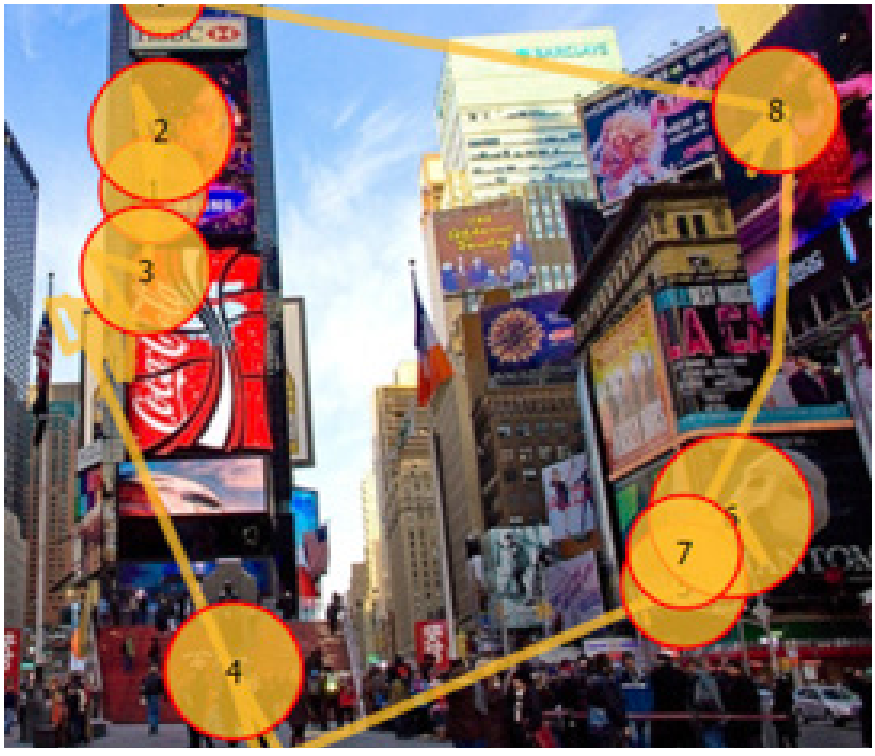


Image 8: Eye-tracking technology generates data visualisation of audience gaze plots.

The limitation of eye tracking is that it does not provide information about the cognitive processes and emotional states that drive eye movements: it indicates what participants see, but not what they perceive. In the case of scripted content, inferences can be made against the context of what is being shown.

As shown through a description of each biosensor, a multi-modal approach to biometric research is essential to capture the full range and complex interplay of unconscious modulations of psycho-emotional responses through to conscious thoughts and cognition. Rigour in experimental design is also essential, owing to the multitude of variables and limitations in human cognitive-behavioural research.

The Studies

A series of pilot studies were conducted to test the biometrics technology in a variety of ways across the media arts. The aim was to train in the application of the technology and evaluate methodological approaches to several disciplines including short film, pitching practices, binaural sound, casting practices, television commercials and genre. The pilot studies initiated two industry-commissioned projects for the Australian Broadcasting Corporation (ABC) that were conducted and delivered within a two-week time frame.

In the next sections, excerpts from each study will be presented to demonstrate a variety of methods and provide useful information, starting with the ABC studies and following with the pilot studies for further reading. It is beyond the scope of this paper to include every study in full detail; however, additional information is available at AFTRS.

The Industry-Commissioned Studies

PROJECT LEAD AFTRS: Penelope Thomas
PROJECT LEADS SARA: Dan Harrow, Tom Lobban

The pilot studies generated and informed two industry-commissioned projects for the Australian Broadcasting Commission. The process used was similar to that of DR (Danish Broadcasting): the film was received at final-cut stage on Day 1; researchers watched the film and designed the study with a focus on how the qualitative/interview stages might best work in combination with biometrics; the study was conducted over Days 2, 3 and 4; analysis and preparation were completed on Day 5 and a report delivered to the show's producers by video conference the same evening. After the findings and recommendations were presented, discussion opened up with filmmakers about options for the final edit.

The mixed methodological approach combined quantitative and qualitative methods: quantitative biometric measure obtained physio-emotional response data; qualitative self-reportage surveys obtained subjective self-reportage reactions; and individual interviews obtained further subjective information for depth of insight. A representative sample group of 20 participants individually watched the episodes while being measured with biometric biosensors, specifically galvanic skin response, eye-tracking and facial expression capture. The data output plotted engagement and valence against the episode timeline. Researchers were able to observe the data in real time, using it as a diagnostic approach to inform the post-viewing interviews. The in-build, post-viewing survey was designed to capture the participant's response to specific editorial considerations, including likes, dislikes, and areas of confusion/frustration, character and narrative preferences.

Research findings included empirical levels of engagement that mapped against the story arc of the script. The nuance and value of this can be demonstrated by looking at the aggregate reads of engagement levels across the two different drama episode timelines. Image 9 shows that one drama generated a strong engagement level at the beginning and end, and a steadily escalating rise of engagement level through the episode. By contrast, the second drama produced an engaging beginning with a sharp peak mid-way and a steady engagement build towards the end of the episode (see Image 10). In this case, the biometric engagement levels indicated a strong start and steady build to a strong finish, yet the need for a tightening to maintain interest at several points of the timeline.

Image 9: Aggregate of empirical engagement levels shown in real time against the episode timeline

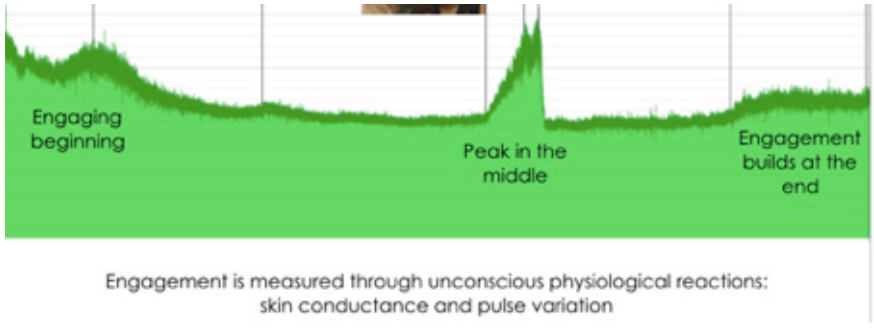


Image 10: Aggregate peaks of facial expression and associated emotions against the episode timeline

In addition, the combination of these engagement levels with facial expression and eye tracking analysis revealed several findings. Firstly, that specific types of emotional engagement aligned with the intention of the content: joy with dancing; excitement with characters on the move; relief with resolution; and surprise with misstep and revelation. Secondly, that key moments evoked significantly strong emotional responses, including humour, sadness, joy, surprise, anger, and high levels of tension (see image 11). This aggregate read also indicated minor dips in engagement that suggested the time to action needed tightening. Recommendations included the type of shot: reduction of extended shots, streamlining, focus and direction.

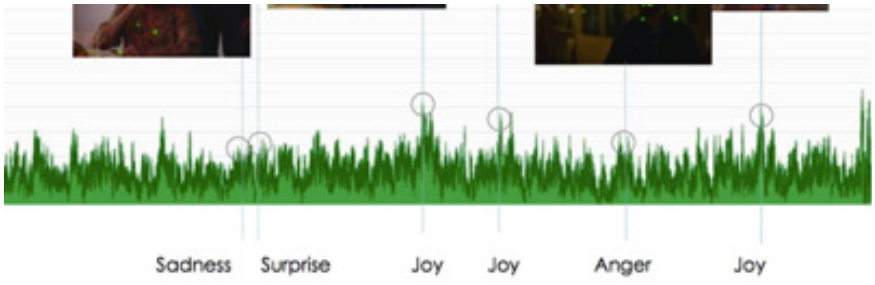


Image 11: Aggregate peaks of facial expression and associated emotions against the episode timeline

Findings and recommendations included a sharpened focus on preferred characters and scenes. Quantitative aggregates could be equated to the strength of engagement generated by main characters. It also indicated how immediate the characters were liked by audiences, which is a valuable finding for the first episode of a series. Specific characters were identified as generating strong emotional responses and this success could be mapped at the granular level of the script against lines of dialogue. Evaluations were made about how to improve character engagement, such as increasing screen time for specific supporting characters to increase the emotional investment and raise the stakes within the narrative.

A thematic hierarchy was designed to explain results to the show's producers at multiple engagement levels: narrative, cast, community, familial dynamics, and character relationships. Specific elements within these thematic levels were evaluated against the biometric data. Recommendations were made to tighten pacing and central tension, build greater intrigue at specific moments, and improve the clarity of core relationships. Researchers were also able to recommend ways to tighten the pace and elevate the drama. These techniques included removing non-critical scenes, cutting low-drama moments, and adding music to elevate the emotional response and increase engagement.

The findings demonstrated that the biometric data provides very specific information at the granular level that can be quantified and used to apply insights in a meaningful way. With both television drama pilot episodes, recommendations were made about refinements to specific audience preferences, character development, storyline incongruities, dialogue ‘misfire’ around a sensitive topic, continuity, relatable and realistic intrigue, tension, sub-plot, angle, roundedness of character, and the desired amount of screen time for individual characters. These recommendations not only informed the creative team about the pilot episode but the entire series. This granular information enables experimentation and strengthens the argument for more risk-taking, iterative methods in production where versions can be assessed at any stage of the edit, before locking the edit or, in some cases, maybe even finding a strong enough case to reopen the edit.

The value of the AFTRS/SARA collaboration was the team's use of an iterative approach to develop a mixed methodology for the research process. Mid-week of the first study, an interim analysis generated the idea to invert the methodology. Rather than look at the data first, the team tested the diagnostic use of biometrics in combination with the qualitative interviews. This led to an important shift. Analysing what audiences self-reported against the biometric data provided fascinating insights into audience engagement in the product. For example, a participant may have reported feeling anger, whereas the biometrics read detected joy; a participant may have stated that they preferred a specific character, whereas the biometrics read identified their gaze as predominantly focused on another. The qualitative interviews generated hypotheses about the content that could then be tested and analysed against the biometrics data.

Client feedback on the research was positive: compared with previous audience test studies from service providers in the UK and USA, the AFTRS/SARA study was endorsed by as by far the most thorough and useful. For researchers, the experience demonstrated that the relationship with producers and commissioning editors requires trust and diplomacy, particularly as recommendations may empower some decisions over others. Overall, the more seamlessly the research can be embedded into the editorial process, the more manageable it will be to address client needs and refine the production.

The Pilot Studies

Prior to the ABC commissioned research, a series of pilot studies were conducted to train practitioners and test the biometrics technology in a variety of ways across the media arts: Two Short Films, the Science of Pitching, Sound and Sense, Casting Call, TVCs, Tapping into Fear. The pilot studies initiated two industry-commissioned projects for the Australian Broadcasting Corporation that were conducted and delivered within a two-week time frame.

In this section, excerpts from each study will be presented to demonstrate a variety of methods and provide useful information. It is beyond the scope of this paper to include every study in full detail; however, further information is available at AFTRS.

Pilot Study 1: Two Short Films – A Biometrics Approach

ASSOCIATE INVESTIGATOR: Dr Elvira Fischer

This pilot study evaluates the use of biometrics to measure audience engagement in film.

BACKGROUND

Why study emotional responses to films? Literature on the subject shows that there are evolutionary, complex processes happening in multiple areas of the brain that cause us to engage with films. One example is involuntary flinching, where audiences are emotionally involved and move with a character as they might in real life. Another is mirroring behaviour, such as audiences laughing when a character laughs or mirroring stress-induced behaviours when a character is distressed. These are people-centered, social behavours that generate a connectedness between us.

These deep-rooted and involuntary responses to film are processed by our brain as real-life stimuli. Daniel Kahneman (2013), a behavioural economist and Nobel prize winner, has described two 'thinking' or behavioural systems: System 1 involves everyday decisions and is a fast, unconscious, automatic and error-prone system. System 2 involves complex decisions and is more rationalised – it is a slow, conscious, effortful and reliable system (see Image 12).

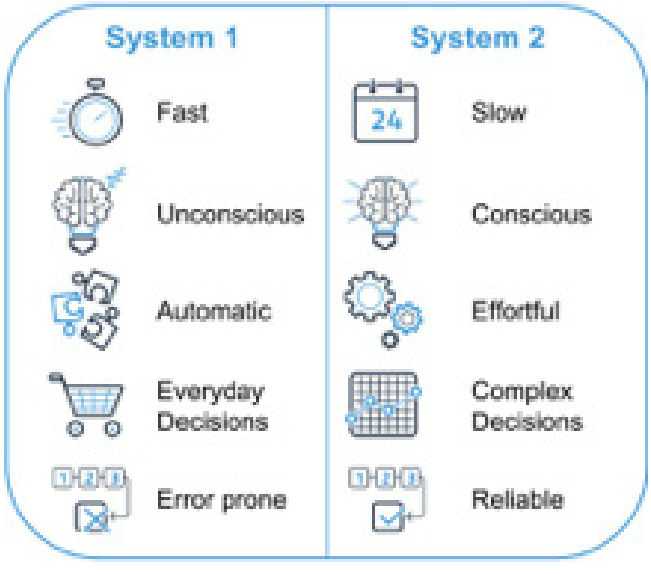


Image 12: Kahneman's (2013) two ‘thinking’ systems.

When watching films and expressing emotions, audiences mostly use System 1 and the first response to content is usually involuntary. Biometric measure enables researchers to access and measure this unconscious behaviour and associated automated responses to film. Arousal, for example, is an indicator of tension points in film. Unlike tradition audience testing, biometrics allows us to tap into that system and better understand why we respond in certain ways.

PURPOSE

To test the quantitative biometric measurement of audience engagement against qualitative audience self-reportage for predicting the short-term and long-term effects of film.

OBJECTIVES

1. To gather and analyse biometric response data and self-reportage data from a representative sample group viewing two short films.
2. To evaluate which metrics are most suitable for the drama genre.
3. To identify correlations between emotional responses and recall rates.

METHOD

Applied multi-modal biometrics: Eye Tracking, GSR and Facial Expression Analysis.

Three Self-Reportage Surveys: one following each film and one final comparative survey.

Sample Group: non-random, 20 participants recruited through a School-wide call for participants.

Location: AFTRS Biometrics Lab, Sydney

Using biometric measure, a range of physiological and emotional responses to two short films are evaluated against a qualitative self-reportage survey.

STIMULI

Two AFTRS student short films – drama

1. *A Boy Called Su* (2014)
   Director: Vedrana Musić
   Producer: Rachel Argall.

   Su, a kindergarten boy, must beat the language barrier and find a means of communication to get his ball back.

2. *Object* (2013)
   Director: George-Alex Nagle
   Producer: Sabrina Organo.

   When a model arrives at a studio for a photo shoot, she discovers how the charismatic photographer always gets what he wants... no matter what.

SELECTION OF RESULTS

1. Biometric Results for the Film *Object*

Facial Expression Analysis: results indicate that respondents registered high levels of emotional engagement during the positive scene and displayed mirroring of positive expressions. Respondents registered less emotional engagement during the negative scene. There is a high registration of the specific emotion ‘anger’ across the film.

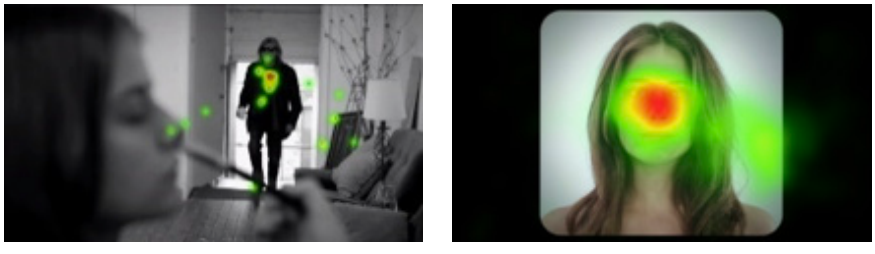


Image 13: Scene from *Object* (2013) and eye-tracking heat maps indicating aggregate gaze-fixation points.

Eye-Tracking: results show that respondents' attention increases and gaze fixation becomes more intense with the close-up that occurs during the negative scene (see Image 13). This differs significantly from the mid-shot.

GSR: results indicate that the level of engagement across the film correlates with emotional intensity. Aggregate levels of intensity are higher across participants that identified as 'male' than participants that identified as 'female' (see Fig. 1).

Fig. 1: GSR result for the film *Object*.

2. Biometric Results for the film *A Boy Called Su*

Facial Expression Analysis: respondents registered high levels of emotional engagement during the positive scene and displayed a mirroring of positive expressions. Unlike *Object*, fewer facial expressions were registered during the negative scene in *A Boy Called Su*.

Eye-Tracking: as shown in Image 14, the aggregate gaze-point focus is intensified during a scene where the victim is in frame compared with the group of abusers in frame. These results are better understood when combined with GSR. Unlike *Object*, an increased attention on the victim did not elevate emotional responses in *A Boy Called Su*, suggesting that responses are more abuser-driven than victim-driven. Participants are more emotionally aroused looking at the abusers.

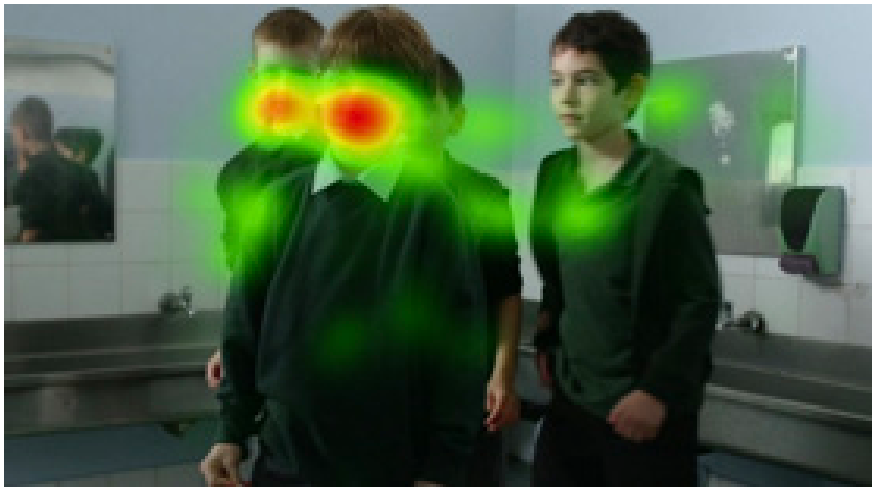
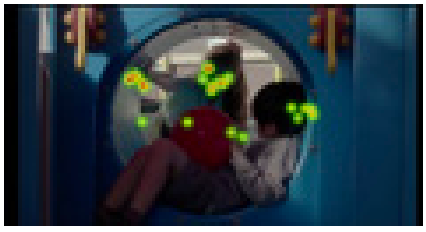


Image 14: Aggregate heat map demonstrating scattered focus in *A Boy Called Su*.

GSR: results indicate that the level of engagement across the film correlates with emotional intensity. Event-related peaks show specific scenes and frames that registered higher results (see Image 15).



(Friendship scene at positive engagement peak)

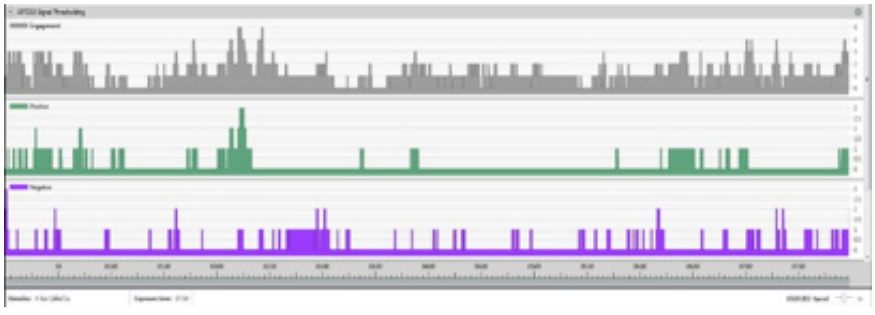


Image 15: Biometric aggregate across timeline of *A Boy Called Su*: purple=negative; green=positive; grey=engagement.

As with *Object*, the aggregate levels of intensity shown in Image 14 are higher in participants that identified as 'male' than participants that identified as 'female' for *A Boy Called Su*.

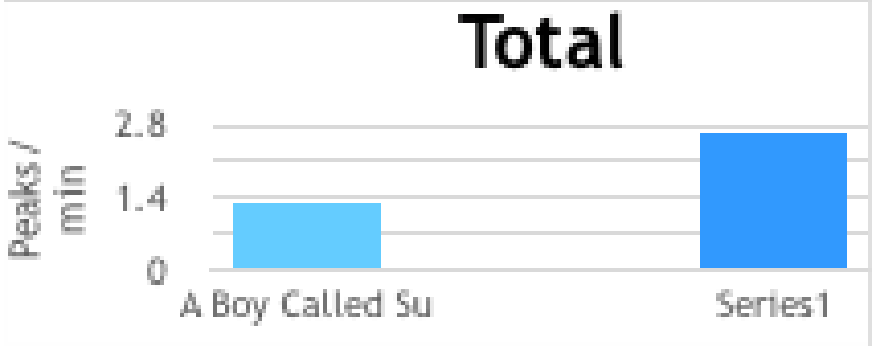


Fig. 2: Aggregate levels of intensity are higher across the male participants for *A Boy Called Su*.

COMPARING FILMS

In both films, respondents registered high levels of emotional engagement at specific 'events' and displayed a mirroring of expressions. For *Object*, facial expression metrics indicate a high level of arousal across the entire film, correlating with a high level of frustration and very little registration of positive emotions. With *A Boy Called Su*, engagement was high during positive scenes and less facial expressions were registered during negative scenes (see Figures 3 & 4).

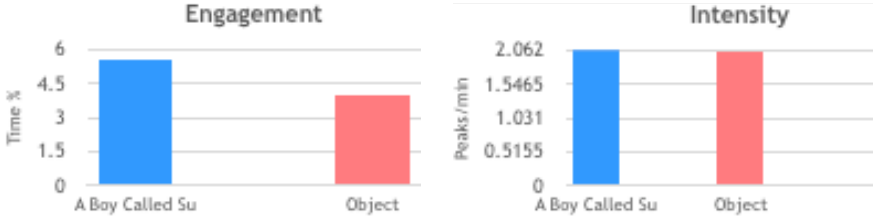


Fig. 3 & 4: Results showing a comparison of the engagement and intensity levels for the two films.

SUMMARY OF FINDINGS

Pilot Study 1 demonstrates that the biometric measure of audience engagement provides strong evidence about overall emotional engagement in short films.

Findings indicate that both films were equally arousing overall; however, *A Boy Called Su* (2014) was more emotionally engaging and *Object* (2013) had the more immediate effect. The results generated the hypothesis that a higher level of engagement correlates with a longer duration of an emotional abuse scene. In addition, close-ups may generate higher engagement levels depending on the context.

The self-reportage results indicate that recall rates are higher with *A Boy Called Su* than *Object*. In terms of audience preference, *A Boy Called Su* was the preferred film, which could be indicative of a greater impact on short-term memory (see Image 16).

Image 16: *A Boy Called Su* (2014) registered more emotional engagement and better recall rates.

Pilot Study 2: The Science of Pitching – A Biometrics Approach

ASSOCIATE INVESTIGATOR: Holly Lyons

This study tests biometrics for the evaluation of pitching techniques in the screenwriting industry.

BACKGROUND AND PURPOSE

Screenwriters, directors and producers are preoccupied with the question: What makes a media pitch a success? Specific pitching techniques are often fed into a pitch, such as the use of a gimmick or an early 'hook' to capture attention. The purpose of this study is to improve our understanding of what makes the best pitch for current reception and offer findings to inform pitching practices.

OBJECTIVES

1. To evaluate if biometric measure of audience engagement can provide information about what makes a pitch effective and identify which techniques are the most effective.

2. To test the biometric measure of conventional pitching as practiced in the screenwriting industry, and evaluate and compare which techniques have stronger impact.

METHOD

Applied multi-modal biometrics: Eye Tracking, GSR and Facial Expression Analysis.

Self-Reportage Survey: one final comparative survey.

Sample Group: non-random, 20 participants recruited through a targeted call out to relevant Industry producers, broadcasters and commissioning editors.

Location: AFTRS Biometrics Lab, Sydney

Using biometric measure, a range of emotional responses to the pitching of an original dramatic concept using three different approaches are evaluated against the success or failure of each pitch and compared with a qualitative self-reportage survey. Specific techniques are identified on the timeline of each pitch to try and ascertain which techniques are the most effective.

STIMULI

The project uses three 3-4 minute video pitches of the same original television series concept delivered by a professional Australian scriptwriter. Each pitch uses different pitching techniques and reveals information in a different order, as follows:
Version 1: Story-Focused

A 3.26-minute pitch told from a story perspective. The pitch opens with a strong hook/cold open and describes the opening of the series with a question. It provides a logline, focusing on story, and follows with a discussion of the protagonist and antagonist, their key internal conflicts and how this relates to the story. An outline of the entire series arc then details the main story points. It concludes by answering the cold open raised at the start of the pitch.

Tone: Formal

Version 2: Character-Focused

A 3.35-minute pitch told from a character perspective. It opens with a joke and then describes the premise by putting the listener in the character's shoes. It outline key research that informs the series, from the protagonist's point of view, and describes the protagonist's flaws, from their own point of view. Key elements from the pilot episode are then presented from the character's point of view and it concludes with a relevant gimmick designed to make the content memorable.

Tone: Casual

Version 3: Writer-Focused

A 3.5 minute pitch told from the writer's perspective. The opening describes the current and topical significance of the concept, and asks: why should this story be told now? The premise is then described by stating why the writer cares and where their passion lies. Relevant research and statistics are used to reinforce the concept. Main and secondary characters are described and why they are important/what meaning they have for the writer. The writer's background and a brief summary are delivered to support why they are the best writer to execute this concept. The entire series arc is described, including a meaningful title of the series.

Tone: Sincere and personal

SAMPLE TIMELINE OF PITCHING TECHNIQUE

Version 3: Writer/creator focused

Time Code	Pitch Technique
00:10-00:17	Clear premise with tagline, 'No one goes home after the bell'.
00:17-00:27	Discuss the topical significance of the concept.
00:27-00:34	Describe the tone: Humour, not a dark series.
00:34-00:42	Raise the problem of the Indigenous population in Australian prisons.
00:42-00:49	Discuss the aspirational nature of the series.
00:49-01:02	Provide an uplifting view of the premise.
01:03-01:13	Explain that the aim is not to glorify prison.
01:20-01:21	Discuss prison as a test of the human spirit.
01:21-01:28	Discuss the mix of characters and indicate their arcs and journeys.
01:29-01:34	Ask a direct question about the state of Australian youth prisons.

SELECTION OF FINDINGS

The findings indicate that a biometric measure provides strong evidence about overall emotional responses to pitches when combined with traditional survey methods. Pitch Versions 2 and 3 elicited high emotional arousal and strong levels of negative valence. Self-reportage favoured Versions 1 and 2 over Version 3, and recall rates were highest for Version 2.

Image 17: Pitch video still showing aggregate heat map of gaze fixation across the entire sample.

The combination of GSR, facial expression analysis and self-reportage surveys provide unprecedented insight into pitch delivery engagement levels, indicating which specific pitch techniques are more effective than others. The eye-tracking was useful for arousal and cognitive load metrics; however, gaze path had limitations as the visual remains constant throughout the videos (see Image 17). Recommendations include to compare study design options for the stimuli, such as randomising/non-randomising the video sequencing.

Pilot Study 3: Sound and Sense – Testing Biometrics for Audience Engagement of Binaural Sound

ASSOCIATE INVESTIGATOR: Lisa Sweeney

This study trials biometrics for audience engagement of binaural sound.

BACKGROUND AND PURPOSE

Sound design has been identified as one of the most important elements of new immersive mediums such as virtual reality, augmented reality and binaural podcasting. Why is this, and what does sound provide that is critical to the success of immersive productions?

The purpose of this study is to improve our understanding of what makes the best immersive sound content for current audiences and offer quantitative, empirical data on audience engagement of binaural sound.

METHOD

Applied biometrics: GSR (galvanic skin response) only.

Sample Group: non-random, 20 participants recruited through a School-wide call out.

Recruitment Exclusion: Respondents with previous experience listening to the stimuli.

Location: AFTRS Biometrics Lab, Sydney

The study applies a GSR biosensor to measure autonomous engagement levels of individual participants as they listen to a 20-minute binaural podcast drama via noise-cancelling headphones from a recliner chair, in-studio. The sample group's heart rate and electro-dermal activity data is collected and aggregated against the timeline of the stimuli.

STIMULI

One 20-minute previously unheard binaural podcast drama titled *Precipice*.

SELECTION OF FINDINGS

The aggregate video against the timeline shows a consistently high measure of heart rate, indicating a high level of engagement across the full 20-minute binaural narrative (see Image 18). In the same aggregate visual, GSR is also consistently high and clearly indicates areas of the experience that generated arousal, particularly in the first quarter of the episode and the final scene. Event-related peaks in GSR can be examined against the narrative and sound design techniques, to identify which moments in the experience generated the highest engagement.

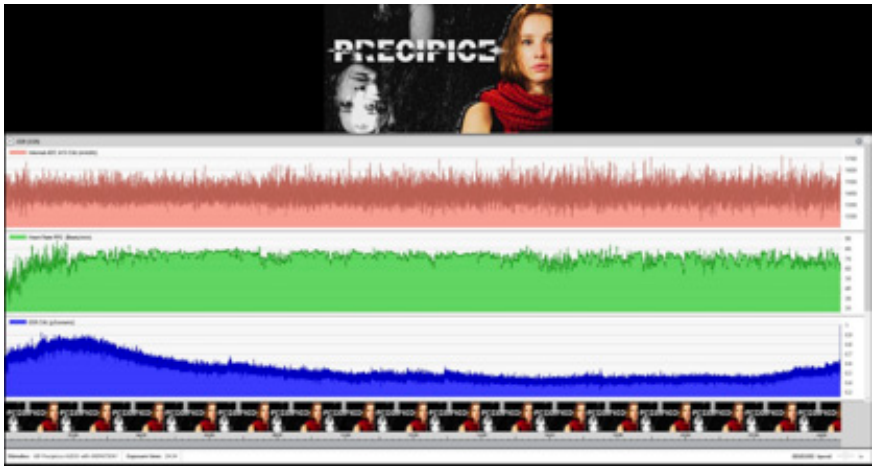


Image 18: Immersive sound - The GSR aggregate shows consistently high levels of engagement and event-related peaks.

The results of this study demonstrate that GSR provides a highly accurate measure of audience engagement with binaural sound and generates valuable information about the potential of immersive sound productions.

Pilot Study 4: Casting Call – Testing the Biometric Measure of Audience Engagement to Identify Correlations between Emotional and Physiological Responses to Auditions and Casting Outcomes

ASSOCIATE INVESTIGATOR: Barrett Casting

This pilot study tests biometrics for casting auditions as practiced in the Australian film industry.

BACKGROUND AND PURPOSE

What happens during a casting session? Are we able to identify objective responses to an actor's audition? Can we quantify physiological responses that are indicative of personal preferences? Are there correlations between these responses and casting outcomes? This research project studies elements of audience engagement in talent casting, as practiced in the screen and broadcast industry.

METHOD

Applied multi-modal biometrics: Eye Tracking, GSR and Facial Expression Analysis.

Self-Reportage Survey: one final comparative survey.

Sample Group: non-random, 20 participants recruited through a targeted call out to relevant Industry producers, broadcasters and commissioning editors.

Location: AFTRS Biometrics Lab, Sydney

The study applies multi-modal biometrics to measure a range of physiological and emotional responses to three videos of three live auditions. The videos are of different actors auditioning for the same role in studio. Biometric results are compared with a qualitative self-reportage completion survey demonstrating respondent recall and preference.

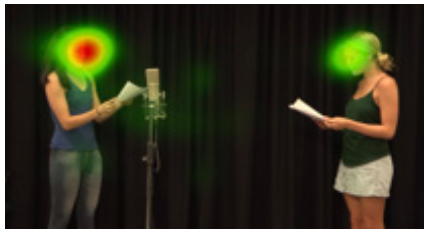
STIMULI

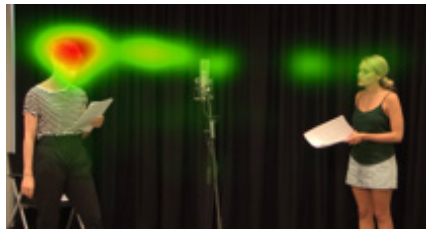
Three videos of three live auditions, each approximately 90 seconds in length.

SELECTION OF FINDINGS

The results clearly indicate that Candidate X differs from Candidates Y and Z (anonymised results) in terms of emotional arousal, engagement and attention. GSR and facial expression metrics prove to be the strongest biometric predictors for this topic, providing deep insights into respondents' emotional responses.

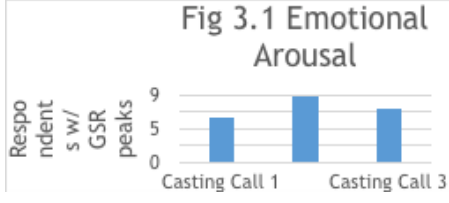
The results correlate well with self-reportage (see Images 19 and 20). The limitation to these findings is that observation of the auditions by respondents is via video and not live.

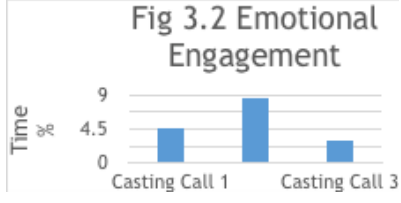




Images 19 & 20: Heat maps showing aggregate gaze fixation on casting candidates

The findings demonstrate that a biometric measure of audience engagement provides insights about the overall impact of casting candidates and casting call performances.





Figures 5 & 6: Quantitative measures of emotional arousal and engagement across the three auditions.

Pilot Study 5: Television Commercials

ASSOCIATE INVESTIGATOR: Martin Brown

This study is very similar in design to that of the two films. A brief summary will be provided.

PURPOSE

The purpose of this study is to test the biometric measure of audience engagement in television commercials (TVCs) with a focus on branding recognition and recall.

METHOD

Applied multi-modal biometrics: Eye Tracking, GSR and Facial Expression Analysis.

One Self-Reportage Completion Survey

Sample Group: non-random, 20 participants that responded to a School-wide call for participants.

Location: AFTRS Biometrics Lab, Sydney

Using biometric measure, a range of physiological and emotional responses to two different TVC campaigns are evaluated against a qualitative self-reportage survey. The survey tests brand recognition and recall, and asks respondents to decide which campaign they think is more successful.

STIMULI

The stimuli consists of two Australian TVC video campaigns, under three minutes in length: 1) Legacy, an Australian charity for families of injured or deceased defecnce force service members; and 2) New Idea, an Australian household magazine distributed weekly.

SELECTION OF FINDINGS

Findings indicate that faces compete with attention to brand when presented simulatenously (see Image 21). Attention to brand is higher for the Legacy campaign (see Image 22). Emotional engagement and arousal are higher for the New Idea campaign. The self-reportage survey results indicate that emotional engagement and arousal can lead to a higher recall of the commercial, but not necessarily of brand.

Biometrics provide strong evidence about overall positive engagement of a TVC as well as attention to brand and brand recollection.

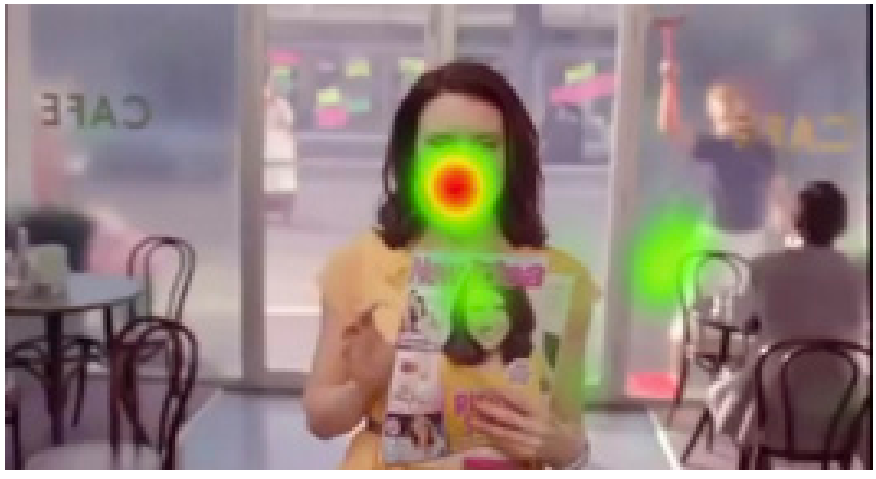


Image 21: Heat map showing aggregate gaze fixation on New Idea TVC branding

Image 22: Heat map showing aggregate gaze fixation on Legacy TVC branding.

Pilot Study 6: Tapping into Fear – Exploring the Potential of Biometrics for Genre Studies

Prepared for ACMI (Australian Centre for Moving Images)

This pilot study evaluates the use of biometrics for genre studies.

PURPOSE

This study is an exploration of the potential of biometrics for genre studies, specifically the fear genre. The main purpose is to explore ways to compare what audiences think they feel about a genre against biometric measure.

METHOD

Applied multi-modal biometrics: Eye Tracking, GSR and Facial Expression Analysis.

One Self-Reportage Completion Survey

Sample Group: non-random, 20 participants recruited through a School-wide call for participants.

Location: AFTRS Biometrics Lab, Sydney

The study uses five fear-inducing video clips on a variety of subjects selected by a web search of videos that appeared using the superlative term 'scariest'. Three survey questions are used: the first question establishes which clips have been seen previously by the participant in order to include that variable in any evaluation of results; the second question asks participants to rate clips according to how fear-inducing they found them; and the third question asks respondents to list their top three fears to provide additional information about participant attitudes to fear.

STIMULI

The five videos of varying lengths are titled Water Cave, Spider Ear, Home Horror, Car Cat and Scenic Zombie. Participants are provided with appropriate content warnings.

SELECTION OF FINDINGS

Results indicate that less than 5 per cent of participants had previously seen one or more of the stimuli videos. Participants ranked Home Horror as the most fear-inducing video and Scenic Zombie as the second most fear-inducing video. The top results of the self-reportage survey regarding participants' top three fears are: 'drowning, 'darkness' and associated fears, including a combination of the two with 'deep dark waters'. This result is incongruent with the Water Cave clip ranking lower. Similarly 'spiders' and associated fears had a high number of responses; however, the Spider Ear clip was not rated highly as fear-inducing by the sample group.

BIOMETRICS RESULTS

GSR results quantifying heart rate and electro-dermal activity across the sample group activity indicate that the Home Horror clip has the highest level of autonomous engagement: statistically significant at 45.5% (see Fig. 5).

GSR	CAT CAR	WATER CAVE	SPIDER EAR	HOME HORROR	SCENIC DRIVE
PEAK COUNT	3	1	1	5	3
PEAK %	27.3	9.1	9.1	45.5	27.3

Fig. 5: GSR results showing the heart rate and electro-dermal activity measures of five fear-inducing clips.

As shown in Figure 6, biometric sentiment results (valence) indicate that Scenic Drive generated the highest levels of sentiment: joy (7.755), disgust (4.182) and anger (1.285). The low level of 'surprise' with the Spider Ear clip, given the high ranking of 'spider' in the top three fears question, was also unusual, as well as the low level of 'sadness' and the high level of 'joy' sentiment in the Car Cat video.

SENTIMENT	CAR CAT	WATER CAVE	SPIDER EARS	HOME HORROR	SCENIC DRIVE
ANGER	0.107	0.137	0.806	0.048	1.258
SADNESS	0.334	0	0	0	0
DISGUST	3.351	0.061	3.318	0.778	4.182
JOY	4.666	0.189	6.196	2.456	7.755
SURPRISE	0	0	0.059	0.004	0
FEAR	0	0	0	0	0
CONTEMPT	0	0	0.049	0	0.105

Fig. 6: Biometric sentiment results for the fear genre study.

This study serves as a first exploration into the biometric measure of audience engagement for genre studies. The results generate further research questions and hypotheses. What is the role of narrative? Can sentiment nuances be identified, such as fear versus fright? Why is there a high result with the sentiment 'joy and a low result with 'fear'? Is this indicative of a defence mechanism? If so, how does this affect the interpretation of the sentiment findings? What is the role of music and sound? Would the Water Cave clip have generated different results with different music?

As with the previous pilot studies, these results show the importance of a mixed methodological approach to biometric testing for audience engagement where, in many cases, the limitations of one biosensor are compensated by another.

Conclusion

The research outcomes support the use of biometric technology to measure audience engagement in the media arts. The pilot studies demonstrate how it can be used in a variety of ways across screen and broadcast. The short film and TVC studies generated very specific findings such as correlations between audience engagement levels and shot types. The pitching study enabled levels of engagement and emotional responses to be pinpointed to the effectiveness of pitching techniques, as well as at the macro-level with a correlation between pitch-styles and pitch outcomes. A biometric read of casting sessions generated recommendations for future research on audition techniques and traditional casting practices. The use of GSR for an immersive sound study revealed that the medium maintained a high level of autonomous engagement over a 20-minute audio entertainment experience.

Finally, conducting the industry studies helped to demonstrate the way biometrics can be used for audience engagement in meaningful ways for content creators. A biometrics approach to content development enables content creators to refine their offering to the consumer rather than operating on traditional audience testing and what might be considered low-level instinct. The innovative aspect of the AFTRS/SARA collaboration is the curation of a mixed-methodological approach that can be embedded into the editorial process of any media content.

AFTRS would like to acknowledge the contributions of Martin Brown, Steve Vidler and Kim Batterham in the preliminary stages of this initiative. We also thank directors Vedrana Musić and George-Alex Nagle for permission to use their films, and all contributors to the studies.

Media Image by Tristan O'Donnell for AFTRS

References

Boucsein, W. (2012). *Electrodermal Activity* (2nd ed.,). New York, NY: Springer Science + Business Media.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Ekman, P., Hager. J. C., & Friesen, W. V. (1981). The symmetry of emotional and deliberate facial actions. *Psychophysiology, 18*(2), 101–106.

Ekman, P., Davidson, R.J., & Friesen, W. (2002). *Facial Action Coding System investigator's guide*. Salt Lake City, UT: A Human Face.

Jain, A., Nandakumar, K. Ross, A. (2016). 50 Years of biometrics research: Accomplishments, challenges and opportunities. Pattern Recognition Letters, Vol.79, pp. 80-105.

Kahneman, D. (2013). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Images:

Pilot study images: AFTRS copyright.

iMotions images: granted permission to use.

Student films: AFTRS copyright/granted permission to use.

DR interior photograph: granted permission to use (M. Brown).

Legacy and New Idea TVC: granted permission to use.

DR television show images: adhere to copyright for final.

Australian Film Television
and Radio School

Australian Government

AFTRS