SMPTE

MOTION IMAGING JOURNAL

Covering Emerging Technologies for the Global Media Community

Biometric Signals Reveal How Audiences Engage With Stories

By Clayton P. Mosher and Brian Wellner

Introdução:

Em tempos de TV 3.0, este artigo (que, na verdade, é muito mais um tutorial) discute um *framework* capaz de coletar dados biométricos dos usuários e como estes dados podem ser incorporados as experiências de visualização de conteúdo, tornando-se inclusive em *insights* para os próprios criadores de conteúdo. Para os mais céticos (que não é o meu caso), que vêem isso muito mais como um problema de invasão de privacidade, o artigo também lança componentes do que eles chamam de *frameworks* éticos. O que particularmente achei muito importante, pois como sempre digo quanto mais precisarmos dos dados dos usuários, mais éticos devemos ser para com eles. Sendo assim tenham todos uma boa leitura em um início de ano desafiador!

Tom Jones Moreira

Abstract

Consumer-wearable Bluetooth-enabled biometric devices are becoming more reliable for the collection of physiological data. They are also more accessible to the average consumer. Building a technology platform with the right devices, components, and algorithms can enable content creators to gather consumer insights on how viewers are engaging with their content. Additionally, content can be created with the intention of allowing viewers to interact and control storylines by utilizing their physiological responses. This article outlines the framework by which this type of data can be collected, what types of stimuli can be applied to insight gathering and content creation, and how this data can be processed.

Keywords

Advanced narratives, algorithms, artificial intelligence, biometrics, facial action coding system (FACS), galvanic skin response (GSR), gaze tracking, heart rate variability (HRV), signal processing

Introduction

Biometric data used as measurements of physiological responses to stimuli can reveal quantifiable metrics on how audiences engage with storylines and characters. This type of data can be used in focus group research environments and to power consumer-level content viewing experiences (see examples in **Fig. 1**). There are several devices and standards by which this type of data can be collected. As we will explore in depth in the upcoming sections, Bluetooth-enabled galvanic skin response (GSR), heart rate variability (HRV), gaze tracking, *electroencephalography (EEG),* and facial action coding devices can be utilized to gather physiological responses from people while they are watching content. These types of responses can be gathered in the form of raw data that can be processed and used to gather insights on how people engage with content. This data can also be used to create personalized and interactive content viewing experiences. When gathering biometric data from people watching video-based content, consideration must be given to the various types of devices and screens that viewers use to consume content in today's landscape. Additionally, consideration must be given to cloud connectivity constraints and challenges. Finally, data privacy is paramount to a successful implementation of this technology.

Biometric data used as measurements of physiological responses to stimuli can reveal quantifiable metrics on how audiences engage with storylines and characters. This type of data can be used in focus group research environments and to power consumer-level content viewing experiences.

Biometric Signals

The last several decades have seen a surge in the development of biometric technology, ranging from applications within the medical field (remote patient monitoring) to security and authentication (unlocking your smart phone with a fingerprint), law enforcement (lie detection testing), and user experience (UX) design and research (eye-tracking to gauge attention). For the entertainment industry, we believe a core toolkit of sensors could together offer a vast array of information about a viewer's experience and reaction to content (**Fig. 2**). While this toolkit is by no means exhaustive, it is a reasonable starting point for any biometric research program.

GSR to identify timepoints in content that are highly emotional: Electricity is more conductive in water, and thus our skin becomes more conductive as we sweat. The palms of our hands, the soles of our feet, and our foreheads have specialized sweat glands that release tiny amounts of sweat when we feel strong emotions. These sweat glands are ideal for measuring engagement since they exist in the only part of the body that is exclusively controlled by the sympathetic nervous system, which readies the body for action, also referred to as "fight or flight."[1–5]

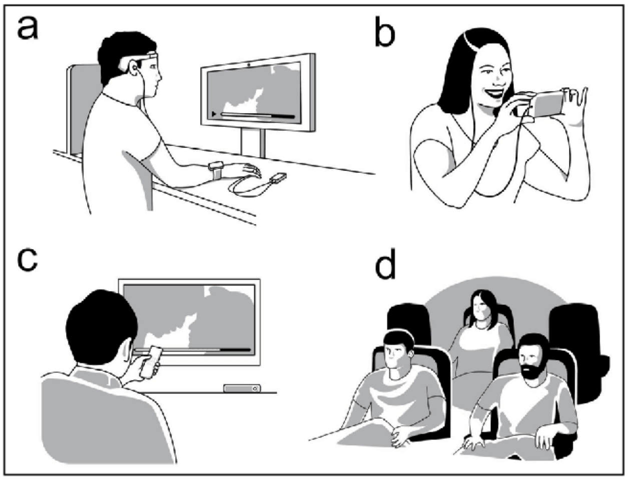

FIGURE 1. Illustration depicting the various platforms where biometric data might be used to better understand audience responses to stories. (a) Focus group research setting. (b) Remote testing with smartphones. (c) 10-ft TV experience. (d) Theater setting.

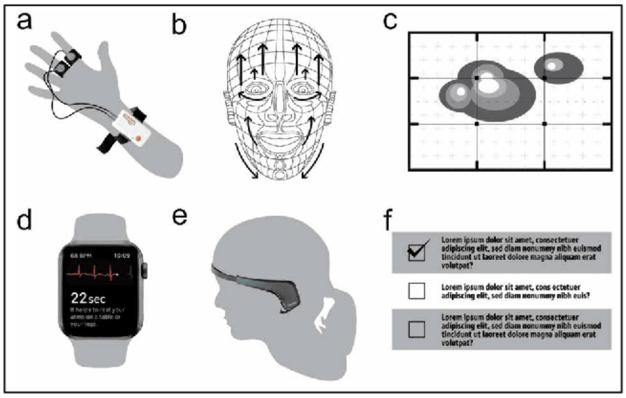

FIGURE 2. Examples of devices that form a biometric toolkit and how they are worn. (a) GSR. (b) Facial action coding. (c) Eye-gaze tracking, illustrated as a heatmap showing where users look on a screen that is divided into nine segments. (d) HRV measured through a sensor embedded in a watch. (e) EEG. (f) Survey.

Facial Action Coding System (FACS) to identify when people respond to content with facial expressions, including subtle microexpressions: Facial expressions communicate social signals and can be used to infer a person's internal emotional state. GSR tells us when a person experiences a strong emotion, whereas facial expressions tell us the valence (negative or positive) of that emotional experience. The FACS was developed to operationalize the science of facial expression by breaking down facial expressions into individual muscle movements. Current approaches use computer vision technology to identify these movements automatically from video footage. However, the accuracy of the technology depends on the images used to train the algorithm—to avoid bias, it is critical that the algorithm is trained on a diverse population of faces with images taken from multiple viewing angles.[6–9]

Gaze tracking to detect the focus of a person's attention: Using an infrared camera or computer vision applied to a video of a face, we can track where a person is looking on a screen. This technique tells us where a person is paying attention, if there are multiple features that compete for users' attention, and which features they may remember later on. Gaze tracking can help inform a content creator on how to guide the user's attention, for example, by facilitating gaze following, a social skill where a user will look in the same direction as an actor on screen to gather more information. While gaze tracking is a powerful technique, it is limited on the quality of the video signal of the face (e.g., head direction and lighting) and requires a calibration procedure.[10]

HRV to differentiate content that is stressful or anxiety provoking versus peaceful and relaxing: Our heart rate increases for a variety of reasons—exercise, attention, when we are angry or excitedly joyful—making it very difficult to make high-level inferences. HRV, however, measures the variance of the heart rate and infers the balance of the parasympathetic (rest/relax) and sympathetic tone (fight/flight). Through this metric, we can measure on the scale of minutes how content may make someone stressed or anxious (as in a horror or thriller film) versus calm or peaceful (meditation). HRV is easy to record from numerous publicly facing devices that measure the electrocardiogram or photoplethysmography with a bracelet around the wrist, a ring on the finger, or incorporated into a headset or other wearable on the skin surface.[11,12]

EEG to infer decisions of participants and their willingness to approach or avoid content: EEG is measured on the scalp and represents the summed electrical activity of the brain. The signal varies at different locations, depending on the function of the brain tissue at that contact, the folds of the brain, and the thickness of the bone, skin, and hair the signal traverses. A variety of metrics can be calculated from EEG such as frontal alpha asymmetry to infer whether participants find content to be approachable or something they would prefer to avoid, or the P300, which signals whether content is perceived as novel or unique. While EEG promises rich data, it suffers from a number of artifacts (eye movements, head movements, blinking, and heartbeat) that must be taken into account. Current EEG sensors can be expensive and sometimes other biometric sensors may provide equivalent or better information with a more user-friendly interface.[13,14]

Survey or dial testing to measure self-report of a participant's experience with content: Still a gold standard in the field of UX research, surveys provide a meaningful self-report before or after an experience while dials can be used by a user to provide feedback in realtime (e.g., how much a user is currently enjoying content). Our surveys focus on participants' demographics (age, gender, and income), their genre affinity, and questions specific to the study of interest (e.g., memory of content and enjoyment of content). These tools are excellent for complementing biometric sensors and encouraging participants to reflect on and relate their experiences with content to others.

Use Case: GSR Implementation

GSR Background: What Is it? What Insights Can it Provide?

As a key component of the biometric toolkit, GSR measures changes in skin conductance caused by sweating and can be used to identify events in content that evoke strong emotions in viewers. Here, we focus specifically on GSR to outline an organizing framework for designing a biometric research study, establishing best practices for collecting data and reporting the results.

Also termed electrodermal activity or skin conductance response, GSR is widely used in the field of psychology and is one of the most well-established biometric sensors in the field. The specialized eccrine sweat glands that give rise to the GSR signal are one of the few parts of the body that are exclusively innervated by the sympathetic nervous system and thus provide a unique avenue to studying "fight or flight" behaviors (i.e., "calls to action" versus "rest and relax"). Unlike the well-known apocrine sweat glands, which help regulate the temperature when the body is overheated, these sweat glands are located in the palms of the hand, the soles of the feet, and the forehead and are activated during emotional experiences. While GSR excels at identifying content with strong emotional activation (e.g., sex and violence), it is less active and therefore less useful for identifying subtle emotional activations (e.g., seeing a tasty cake versus seeing pollution). It cannot indicate whether an emotional experience was good or bad, only that it happened (Edelberg, 1995) [5].[1,2,4]

GSR can be used in a variety of applications, a few relevant to the motion picture industry are as follows:

- GSR data could be incorporated into existing computational models to improve the ability to predict a movie's box-office success or the engagement of a movie trailer.
- GSR could provide feedback to content creators who wish to compare the emotional impact of elements in their content, for example, effectiveness of two competing scenes, characters, and situations.
- Importantly, GSR provides a tool that is unbiased by culture or language (including children), allowing it to provide insights into the emotional events that resonate with audiences that may lack strong self-reporting abilities.
- If a timepoint in content is known to elicit a strong GSR emotional response, this could be used to trigger the modification of the content, for example, to deliver a relevant advertisement or product placement to maximize emotional impact or to modify a storyline.

GSR Methods: Study Admin Tool, Processing Pipeline and Signal Processing Algorithms

Setting Up a Study, Recruiting Participants, and Obtaining Informed Consent

We created a study admin tool that allows us to organize research studies and recruit participants through an internet browser. We designed the tool, called *Peak*, to provide realtime and post-study reporting and to allow for scalability with participant capacity, sensor types, content types, and study location. We opted to create our study admin tool so that we could integrate the procedure of designing the study, with access to the raw data, and tools that a researcher can use to easily and quickly visualize the results. While other study admin tools are available on the market (e.g., iMotions and MindProber), they often come with a costly subscription and restricted access to data. Often these tools do not provide transparency about the data signal or analysis procedures, making it difficult to appreciate some of their outcome metrics.

In our study admin tool, each study is assigned specific content (e.g., a series of trailers, a movie, and a TV show with ads), and if there are multiple content, then the order should be randomized within each participant, a viewing platform (Android smart phone, iOS smart phone, TV, and theater), and the biometric sensors that will be used (e.g., GSR) (**Fig. 3**). The study designer also has the option to include survey questions before or after viewing the content to obtain self-report information on the participant's demographics, genre affinities, and experience (e.g., memory of content, enjoyment of content, etc.) (SurveyMonkey).

After designing a study, a study admin will then assign one or more sessions within a study. For each session, the user selects a study location (either an address of a physical location or remote at-home for each participant) and a list of participants. When a session is created, the participants will receive an email notification providing instructions on how to participate in the study. Participant demographics and contact information are stored in a directory that allows us to identify specific participants of interest for a study session.

We obtain informed consent from all participants prior to their participation in the study, delineating the participant agreement to be confidential about the study, the use of their biometric data for emotion research, and their right to withdraw from the study at any time. Just like patient health information in the medical field is protected and made private, we ensure that the participants' biometric data will remain private by unlinking it from their personal information (name, contact info, etc.) or any other information that point the data back to the participant (see sections below on industry standards and data privacy).

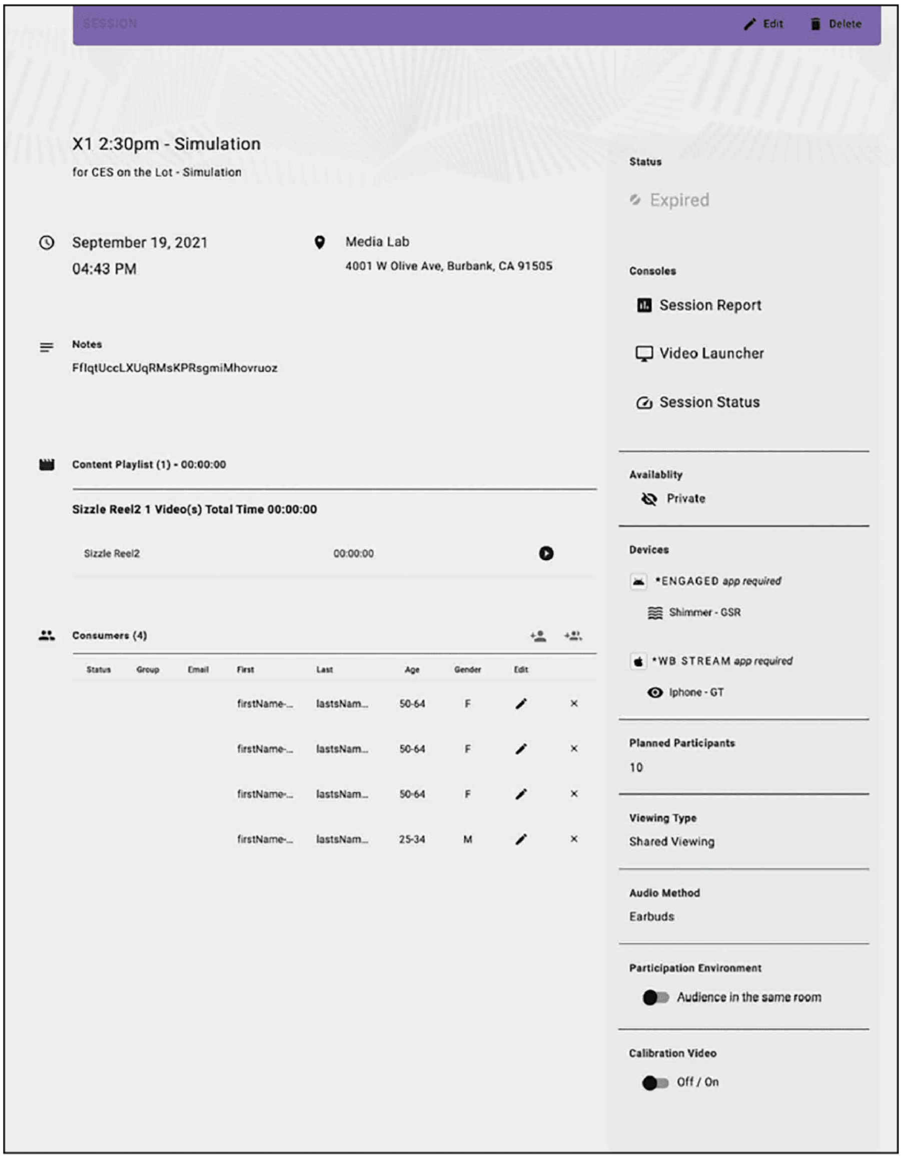

FIGURE 3. Screenshot of the study admin tool illustrating how a study designer can create and launch a study. (Left) The study is assigned a name, a location (virtual or in person), a time the study must be completed by, the content (video trailers, etc.), and a list of participants (consumers) who are invited to participate in the study via email. (Right) The admin selects the platform, biometric sensor, number of participants, and a description of the viewing environment.

Identifying a Sensor and Correctly Placing the Sensor

Since the advent of Bluetooth and other wireless technologies, GSR research has seen an increase in the number and types of available devices. While some wireless GSR devices are more research-oriented and best used by a trained researcher in a focus group (e.g., Shimmer GSR+, Mindprober James 2, and Empatica), others are consumer facing and publicly available for purchase (e.g., Fitibit Charge 5, Moodmetric Ring, to name a few). We have vetted multiple GSR devices in our research and have found that it is important that the company provides an accessible software development kit (SDK)/ application programming interface (API) to efficiently retrieve data from the device on the platform of interest (e.g., iOS, Android, Windows, etc.), that it is a user-friendly device that could be scalable and used in an at-home remote setting, and records data with the highest quality signal informed by biological principles and engineering. A GSR device will produce the best signal if placed on the underside of the hands or feet (palm/sole) or the forehead, as these areas of the skin have the highest density of GSR-specific sweat glands. Movement of the skin in the area of the sensor or the application of pressure to the sensor can cause spurious events that look like GSR but are in fact noise. To help detect these noisy events, some devices incorporate accelerometers or gyroscopes that detect motion and attempt to clean the signal and improve its quality.

In our initial research presented here, we started with Shimmer GSR 3+ because it was one of the first wireless GSR devices introduced to the market and is a standard in the field of academia and consumer neuroscience research. This device contains a wristband with a Bluetooth emitter, which the participant first places on their wrist like a bracelet. The participant then wipes their hands with an alcohol swab to clean the palms of the hands and fingers and to remove any sweat. A researcher applies one silver chloride gel electrode (Lafayette Model 76642R) to the pointer finger and a second electrode to the middle finger of the non-dominant hand. The non-dominant hand is selected to minimize motion artifact. These electrodes are attached with wires to the Bluetooth emitter on the wrist. The participant is asked to limit the amount of movement of their hand during the study.

Establishing a Data Pipeline for Recording and Processing the GSR Signal

To integrate GSR data with our admin tool, we developed an Android app ("Engaged") that could communicate with Shimmer through Bluetooth and relay the signal to our cloud platform in Peak. Here, we present our solution to GSR integration from both the view of the backend data flow as well as the participant's user experience:

- *Requirements:* A GSR sensor with Bluetooth technology, an Android smartphone with Engaged app installed, the participant has been registered for a study, a phone has Bluetooth turned on, a phone has stable WiFi connection.
- *From the Backend:* Once the participant has attached the sensor, the Bluetooth connection is established with the Android smartphone. The raw GSR signal is transferred via Bluetooth alongside timestamps. To accurately calculate GSR peak timepoints, we require a successful minimum transfer rate of 8 samples/sec, for at least 20 sec. Once the app has connected with the sensor, the participant starts the study. The app continuously creates a local record of the raw GSR data on the participant's smartphone. At the same time, the app processes the data in 20-sec packets, an important step that allows visualizing the streaming of fully processed data in near realtime (total of a 30-sec delay). Each packet of data is processed by a peak detection algorithm (see section that follows), and the GSR peaks are recorded. If there is a stable internet connection, the peak times are sent via WiFi to the cloud for storage and the results are displayed in the Study Admin Tool. If there is no stable WiFi connection, the peak times are locally stored on the smartphone until the phone successfully connects to WiFi. Buffering and data streaming are supported by Amazon Web Services Kinesis.
- *From the Participants Perspective:* Each user receives an e-mail notice indicating the schedule for the session and instructions to install the latest version of the Engaged app, to charge their GSR sensor, and directions on when and where to attend (virtually or in-person). At the time of the session, the participants ensure they are connected to WiFi and, if it is an in-person event, they are logged into our server. They then log into the Engaged app on their phone using their email account (**Fig. 4**). They connect their sensor via Bluetooth to the app and receive an indication of successful connection. If they connect successfully and the app is receiving quality data (acquiring at a sufficient sampling rate and data is within range), they can begin the study. If not, the app indicates to try reconnecting the Bluetooth. If the data quality is poor, the participant is instructed to try repositioning the sensor or replacing the gel electrodes with new ones. When they are ready to begin the study, the participant uses their unique ID displayed in the Engaged app to synchronize their GSR data with the content. If the participant is seated at a computer

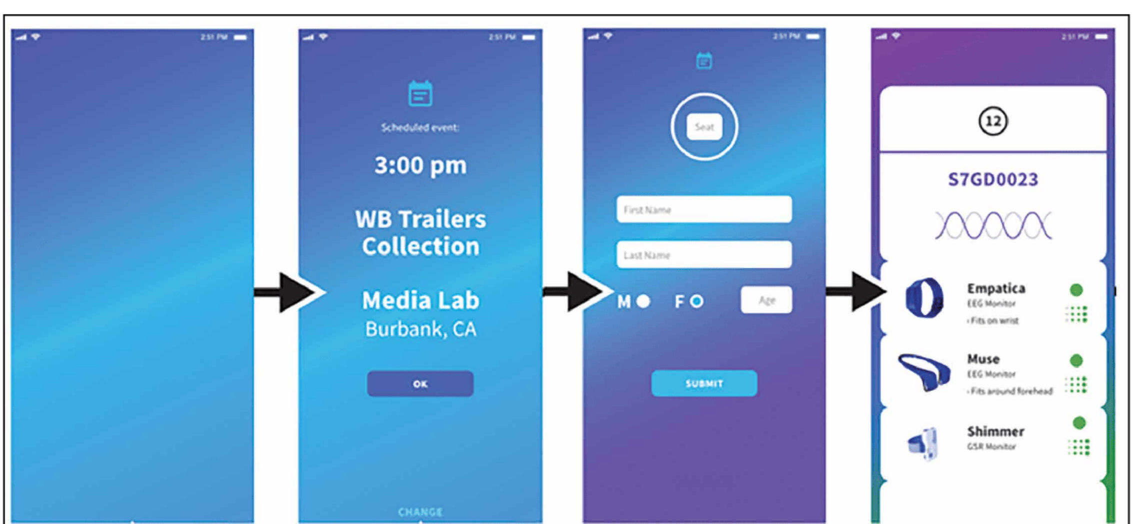

FIGURE 4. User experience flow. A user logs into their account via a WiFi connection. The user selects a device to connect with via Bluetooth. Once connected, the study begins and the content is displayed on a PC monitor.

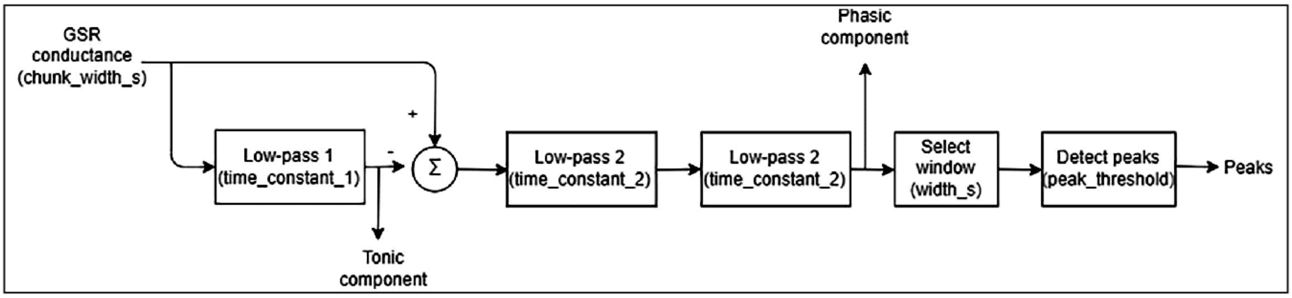

FIGURE 5. Signal processing of the GSR signal. A low-pass-filtered signal (tonic component) is subtracted from the signal to uncover the phasic component of the GSR. Peaks in the phasic component are detected and identified as meaningful timepoints in the content.

monitor, they will open a web browser and navigate to the site to view content. They will type in their unique ID and begin the study. Once the study is complete, they will be notified through the web browser and the Engaged app on their phone and the GSR device will be disconnected. If there are multiple participants watching a single screen, the synchronization to content is predetermined by the study administrator who identifies when all participants are linked to Bluetooth and internet and begins displaying the content.

Algorithm for Calculating GSR Timepoints

GSR measures how the electrical conductance of the skin changes as we sweat and has unit microSiemens (the inverse of electrical resistance, $1/\Omega$). The signal is composed of a phasic component (~0.5-sec duration) that rides on top of a much slower tonic component (changes over minutes). The phasic component is most useful for identifying emotionally laden events. The tonic component is less well understood and may represent mood states but is difficult to interpret due to the slow accumulation of sweat under the sensor. There is a physiological latency of 1–3 sec from the time the neural command is sent, the sweat glands open, and the sweat diffuses to the skin, so this latency must be taken into consideration when determining which emotional event triggered the specific response (Edelberg, 1995).[15,16]

In our initial implementation for the detection of GSR timepoints, we used a series of filters and peak detection algorithms that are well established in the field of emotion research. We first coded the algorithm in R to test the metric for research purposes and then incorporated it into an Android app to provide realtime feedback of the GSR. The key steps are listed below and summarized in **Fig. 5**.

```
Algorithm Steps to Identify GSR
Timepoints
>> Access the last 10 sec of the GSR
signal through Bluetooth.
>> Identify the tonic component of the
GSR by applying low-pass filter to the
signal (low-pass FIR moving average
filter, 8-sec sliding window).
>> Subtract the tonic component from
the GSR signal to obtain the phasic
GSR component.
>> Low-pass filter the phasic GSR com-
ponent (first-order Butterworth fil-
ter, filter cutoff = 5 Hz).
>> Identify the onset of a GSR peak as
the timepoint when the filtered phasic
GSR component exceeds 0.01 uSiemens.
>> Identify the offset of a GSR peak
as timepoints when the filtered signal
falls below 0 uSiemens.
>> Exclude onset and offset times that
are too short to be considered GSR
events (<0.5 sec).
>> Determine the maximum peak between
each onset and offset period. If the
maximum peak amplitude is at least
0.005 uSiemens and no larger than 100
uSiemens, consider it as a GSR peak
event.
>> Report all GSR peak events that are
>0.5 sec, <100 uSiemens, and >0.005
uSiemens as GSR timepoints.
```

Displaying Results in Near-Realtime

By designing the study admin tool, the GSR processing pipeline allows us to display results to a focus group leader or any other researcher in realtime, with a 30-sec delay (see outcome for discussion on this delay). To display the results, a visual-overlay display is used where the processed GSR activity can be displayed per participant or averaged across a selection of participants (**Fig. 6**). Since demographics like age and gender are built into the study admin tool by design, this also provides the ability to create video overlay reports to compare the response to content by different viewers. In post-production, a study admin can identify timepoints in the movie that evoke strong emotional responses (high GSR counts among participants) and create short clips centered on the response to identify the emotional trigger in the content.

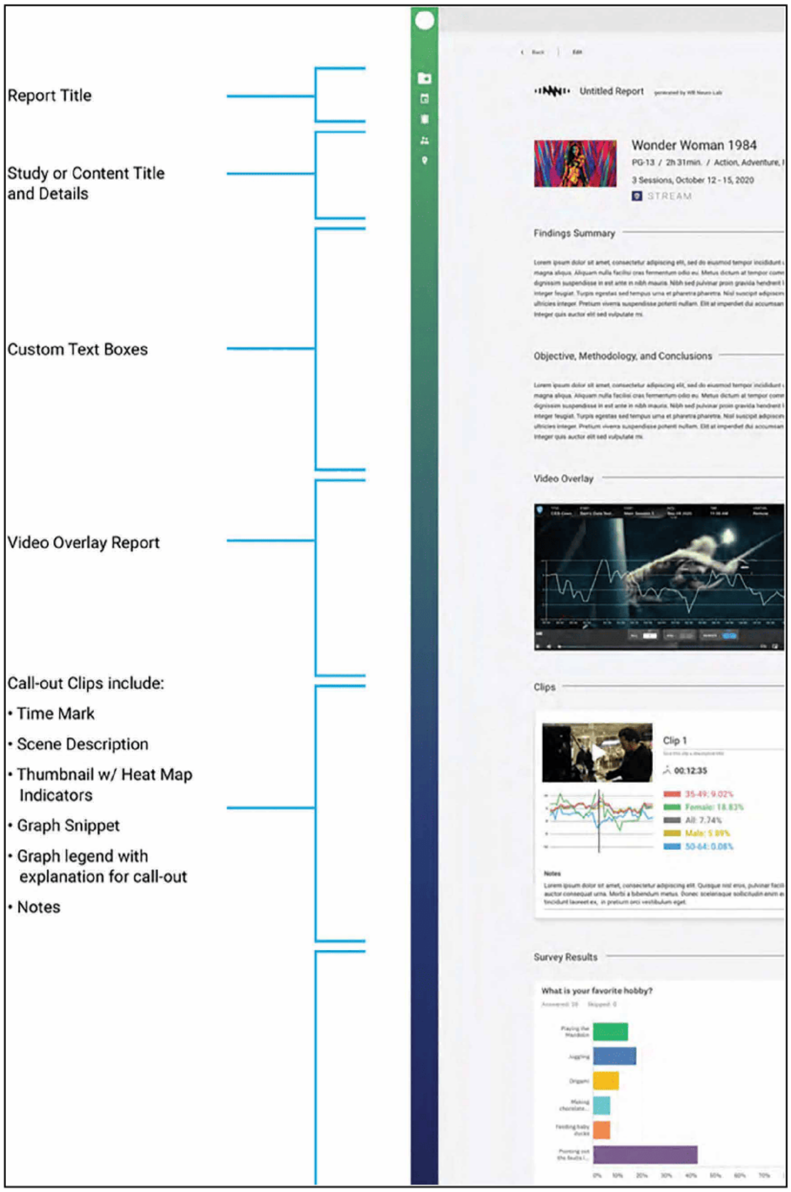

FIGURE 6. Screenshot of the report viewer in the study admin tool. A study admin can label the report and type key insights into open-field textboxes. Admins can identify key timepoints in the content and create video overlay reports where the GSR data is overlaid on top of the video content. GSR data can be separated by demographics to identify timepoints where content differentially drove engagement.

Statistical Comparisons of GSR Timepoints

For each participant, the number of GSR events is summed up in a 4-sec sliding window (binned). Two-sample *t*-tests or analysis of variance (ANOVA) with post-hoc comparisons are used to statistically compare between two or more conditions (alpha level = 0.05). These standard statistical tests can be used to identify if different demographics respond differently to the same content (e.g., are there gender differences?) or how one type of content performs relative to another (e.g., comparing between two different trailers, ads, or scenes). Because these statistics are performed at multiple points in time, there is the potential for obtaining false positives (e.g., if the test is run at 100 time points, there is a chance that at least five of these timepoints will be significant; alpha = 0.05). To correct for multiple comparisons, a cluster-based bootstrapping analysis can be used.[17]

GSR Results and Outcomes | Displaying Results Through Admin Tool

To evaluate the speed, accuracy, and scalability of our study pipeline in realtime, we showcased our tools at an event on the Warner Bros. studio lot and recruited individuals to participate. We did this as a proof of concept, to prepare ourselves for future, more exhaustive and carefully controlled studies with consumers.

We created a short sizzle reel (140 sec) with highlights from movies and TV series we expected to be emotionally engaging and measured how participants' GSRs were activated. We presented the reel on a TV screen while four participants at a time watched, each wearing a GSR monitor (and an EEG headset, see the section titled "GSR Future Directions"). At the same time, we displayed the GSR overlaid on the video on a separate TV screen for other people passing by to view.

This test of our pipeline revealed that we could efficiently record GSR data from multiple individuals in the same room simultaneously with fast turnaround. Over a two-day period, this setup produced data from 134 individuals. Many individuals showed synchronized responses to specific content. For example, a scene of a person getting shot in the chest reliably produced a response in ~90% of individuals. Importantly, consistent with the literature, 10% of the participants exhibited little or no GSR activity. These "non-responders" are known to occur in the population and highlight the need for alternative metrics in the biometric toolkit to capture their emotional responses. As a proof of concept, we compared the GSR response between male and female recipients and observed significant timepoints when these two genders exhibited different GSR activity (**Fig.** 7). Two scenes elicited more GSR activity, and therefore more engagement, in female viewers: a violent scene of a gunshot and a comedic scene from a sitcom.

GSR Future Directions | Other Platforms? New Metrics? Relation to Other Sensors?

Here, we presented a first step toward building a scalable framework for measuring biometrics in a real-world setting. In the future, we aim to test the limits of how many individuals we can record GSR from simultaneously, either watching the same content in a theater venue or at home on users' phones.

Other Platforms: In addition to the Android Engaged app phone-video experience setting we have developed for focus groups, this pipeline can be expanded to apply the same technology in a theater setting, in a "ten-foot" experience where individuals at home are watching TV, or a phone app, where content is streamed through the mobile device. Each of these platforms presents its own set of challenges that we pinpoint in the "Industry Standards" section. For example, one challenge with applying biometrics to multiple platforms is that each platform has different requirements for using biometric data: some restrict biometric data use to health services, others restrict access to certain types of devices, and many platforms work with different coding languages, making it challenging to recode biometric algorithms at scale.

Other Metrics of GSR: Our algorithm currently reports GSR events to identify emotional timepoints in content. However, we can also calculate the amplitude (peak) of these events, which correlates with emotional intensity: content that is highly emotional evokes larger amplitude GSR events. Recent metrics perform phasic decomposition of the GSR signal to estimate the sympathetic nerve activity that gives rise to the sweat response. These metrics combine GSR timepoints and amplitudes into a single metric that reports both when an emotional event occurs and the intensity of the emotion.[15,16,18]

Other Sensors: GSR represents a single aspect of emotional affective space (emotional intensity) and is biased toward very strong emotions. It does not, however, tell us whether an emotional event was good or bad. Combining GSR with other sensors allows for deeper inferences about an individual's emotional experience. Facial action coding coupled with GSR has the potential to indicate the type of emotion a person experiences (e.g., activity of the zygomatic muscle to form a smile would indicate that the event that triggered the GSR event was a positive experience). Eye tracking can indicate what features in the content a person was paying attention to when the GSR event occurred (e.g.,

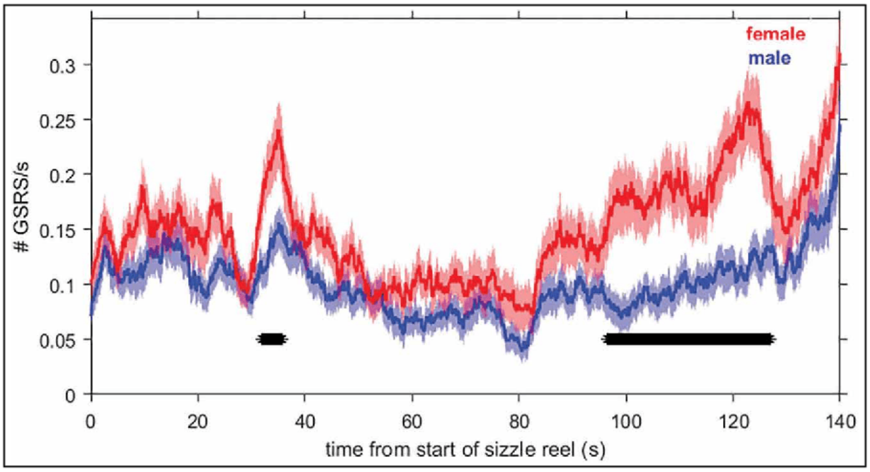

FIGURE 7. Female participants exhibited higher GSR rates at two timepoints during the sizzle reel: an early timepoint (40 sec) depicting a violent gunshot and a later timepoint (120 sec) depicting a comedic episode from a sitcom.

maybe they made eye contact with one of the characters in the content and became more engaged). HRV can provide information on whether the person was in a relaxed state or in a stressed and anxious state just prior to when the emotional event occurred. While EEG has the potential to provide insights into the valence of the emotional event, whether the content is approachable or a participant would prefer to avoid it. It is likely that no single sensor will provide complete insight into the emotional and cognitive state of an individual (not even EEG); only by looking at these sensors simultaneously can we formulate a deeper understanding of how content impacts the individual.

Scaling Up: In our GSR on-the-lot study, we recorded from a small group of people simultaneously, though with high throughput. To record from a large group of people simultaneously or a large pool of people remotely, we face multiple challenges: making sure (1) that a user is able to put the device on and record with little assistance, (2) that the sensor has a wearable design and the user wants to wear it, (3) that multiple sensors in the same room do not block each other's signals (e.g., Bluetooth limitations to how many sensors a device can connect to), (4) that the cost of the sensor is at scale, and (5) that the sensor has universality. Each of these challenges presents the need for developing industry standards to create efficient products that can work on multiple platforms for different user groups.

Signal Processing Time: As we scale up or apply GSR for other purposes, there may be a need to reduce the processing time of our pipeline. Currently the ~30-sec delay for visualizing the content is suitable for a Focus Group study as it gives time for our Focus Group leader to help participants start the study and then walk to the viewing room and analyze the results. While this processing could be reduced, one of the main limitations is the need to have a sufficient amount of data (20-sec packet) to ensure high-level quality of continuous data and to accurately calculate the GSR timepoints. This packet size could potentially be reduced, but alternative catches to determine signal quality would need to be introduced. The packet size cannot be smaller than the duration of a GSR (0.5–1 sec) or the filtering window. Other limitations to the processing time are data transfer from the sensor to the app and from the app to the cloud, buffering times, and visual display of the content with GSR data overlaid.

Discussion

Future of Biometrics in Storytelling

Integrating biometric data into the motion picture industry has the potential to reveal unique insights into how viewers respond to content as well as the ability to create new and unprecedented interactive experiences. These tools offer the ability to observe how viewers respond to content: whether they are engaged with a storyline, whether they will remember a scene in a story, if they are annoyed with a cut or an edit, whether they are paying attention to features a content creator wants them to, or whether they empathize with a character on screen. Importantly, they allow administrators to quantify and observe in realtime the cognitive and emotional state of viewers without disrupting their experience with a self-report tool like a survey or dial. Instead, biometric tools supplement self-report to reveal the underlying perspective of the viewer that they might not report or may even be unaware of. Biometric tools allow access to the minds of viewers who are unable to fully express themselves in a survey (a child, a non-native language speaker). Not only can biometrics be used to provide suggestions on how to adjust and improve content, but they also introduce the ability to create interactive content. Biometric signals could be used to help guide a user through an experience, whether the user is a child viewing an education program, a gamer playing in augmented reality, or an individual at home playing on their phone or looking for the next best series to watch.

Developing Standards

For biometric data to be useful across multiple platforms and accessible to a variety of members of the motion picture industry, it is important to develop standards for devices and processing pipelines that promote interoperability. Based on our experience developing a study admin tool and an app for recording GSR, we recommend the following standards, as new devices are incorporated:

- Developing a single study admin tool capable of integrating data from multiple devices and platforms.
- Using common coding tools and languages so biometric algorithms do not have to be completely rewritten for use on new devices.
- Deciding on standard signal processing tools and metrics for each biometric signal.
- Defining what values of a biometric signal are considered acceptable data quality (e.g., what values of GSR amplitude are out of range to be feasible).
- Identifying biometric devices that are usable across multiple platforms (e.g., some devices only have an SDK for Android, making them unusable on iOS).
- Determining a best approach to synchronizing data on a mobile device with a second device displaying content (e.g., synchronizing to a movie in a theater, audio tagging?).
- Ensuring data privacy of biometric signals and preventing misuse (see the following segments).
- Requiring informed consent to protect both the participant and the company.

Norms

As we build products and begin collecting large datasets from consumers, it will be important for the industry to agree on an established set of minimum normative data so that biometric inferences are interoperable throughout different departments within and across agencies. This will allow us to use biometric data more effectively to predict values in other normative databases related to financial success, merchandise sales, film reviews and rankings, customized streaming, and advertising. In addition to the benefits of linking up databases, norms about consumer demographics help categorize participant information, so we can begin to anonymize their data and protect their personal privacy. We have identified the following list of minimum norms as a starting point for consideration:

- Content type (trailer, movie, series, etc.).
- Metadata about content (IMDB identifier).
- Viewing environment (watching alone or in a group, what size group).
- Screen type (theater, laptop, and phone).
- Consumer profile (age range, gender, ethnicity, parent, and genre affinity).
- Consumer's content evaluation through self-report (enjoyable, memorable, etc.).

Ethical Considerations

Biometric data has the potential to carry sensitive information. A heart rate recording could reveal a participant's underlying health conditions. An eye-gaze recording reveals what features people pay attention to in content and, similar to predictive models that use social media interactions of individuals, could be used to infer sexual orientation, pregnancy status, or other potentially private information.[19] Advanced techniques in brain recording could one day have the potential to reveal mental health concerns, intelligence, interests, and, in the distant future, read our moment-to-moment thoughts in realtime. If this information is not protected, it could have indelible consequences to a person's life and to society at large.

As an industry, we have an ethical obligation to protect personal data and educate our participants in the type of data we collect. We currently require informed consent prior to enrolling a participant in one of our studies. Similar to how every U.S. academic institute and medical hospital has an Institutional Review Board to protect patient rights in research, it may be of interest to establish a similar Review Board of scientists, ethicists, and members of the public when considering biometric data use in the motion picture industry.[20] As we build up our research tools, we will continue to build a framework that encrypts biometric data and unlinks it from participant IDs, names, and other information. A patient's face uniquely identifies them (so much that it is used as superior security log-in on some devices) and as such, we must make efforts to ensure the privacy of any video recording (e.g., as required for eye-gaze tracking or facial expression detection). This may mean that we elect not to save the video footage and only stream and save processed data extracted from the video that does not reveal participant identity (e.g., timestamps of facial action units). The more comfortable our participants feel about their data privacy, the more willing they will be to engage in our research studies and interact with our content.[21,22]

Conclusion

In this article, we discussed a framework by which we can detect and collect biometric data. We also discussed how we can apply this data to consumer research-based insight gathering along with interactive content creation. We outlined the types of devices and biometric signals that can be used to collect data: GSR, HRV, gaze tracking, *EEG*, and facial action coding. We then provided an explanation on how these devices and their data can be incorporated into content viewing experiences. We discussed the need and method by which we are processing the data from these devices to translate into actionable insights. We foresee the need for evolving ethical frameworks to coincide with the evolution and advancements of the components of the technical framework. We also foresee ongoing advancements in the hardware space that will allow biometric data gathering to improve over time.

References

1. W. Boucsein, *Electrodermal Activity*. 2nd ed., Springer, 2012 .
2. M. M. Bradley, M. Codispoti, B. N. Cuthbert, and P. J. Lang, "Emotion and Motivation I: Defensive and Appetitive Reactions in Picture Processing," *Emotion*, Washington, DC, 1(3): 276–298, 2001. [Online]. Available: https://doi.org/10.1037/1528-3542.1.3.276
3. M. Dawson, A. Schell, and D. Filion, *The Electrodermal System*, Cambridge University Press: Cambridge, U.K., 2000. [Online]. Available: https://doi.org/10.1017/CBO9780511546396.007
4. M. van Dooren, J. J. G. (Gert-Jan) de Vries, and J. H. Janssen, "Emotional Sweating Across the Body: Comparing 16 Different Skin Conductance Measurement Locations," *Physiol. Behav.*, 106(2): 298–304, May 2012. [Online]. Available: https://doi.org/10.1016/j.physbeh.2012.01.020
5. R. Edelberg, "Electrical Properties of the Skin," *Methods in Psychophysiology*, Williams and Wilkins: Baltimore, MD, pp. 1–53, 1967.
6. P. Ekman, "Expression or Communication About Emotion," *Uniting Psychology and Biology: Integrative Perspectives on Human Development*, American Psychological Association: Washington, DC, pp. 315–338, 1997. [Online]. Available: https://doi.org/10.1037/10242-008
7. P. Ekman and W. V. Friesen, "Measuring Facial Movement," *Environ. Psychol. Nonverb. Behav.*, 1(1): 56–75, 1976. [Online]. Available: https://doi.org/10.1007/BF01115465
8. M. Kunz, and S. Lautenbacher, "'The Need for Diversity in Research on Facial Expressions of Pain': Reply to Dildine and Atlas," *Pain*, 160(8): 1902–1903, 2019. [Online]. Available: https://doi.org/10.1097/j.pain.0000000000001594
9. S. Stöckli, M. Schulte-Mecklenbeck, S. Borer, and A. Samson, "Facial Expression Analysis With AFFDEX and FACET: A Validation Study," *Behav. Res. Methods*, 50, Dec. 2017. [Online]. Available: https://doi.org/10.3758/s13428-017-0996-1
10. A. Yarbus, *Eye Movements and Vision*, Plenum Press: New York, NY, 1967.
11. A. Malliani, M. Pagani, F. Lombardi, and S. Cerutti, "Cardiovascular Neural Regulation Explored in the Frequency Domain," *Circulation*, 84(2): 482–492, Aug. 1991. [Online]. Available: https://doi.org/10.1161/01.cir.84.2.482

12. F. Shaffer and J. P. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," *Front. Public Health*, 5: 258, 2017. [Online]. Available: https://doi.org/10.3389/fpubh.2017.00258
13. J. Polich and A. Kok, "Cognitive and Biological Determinants of P300: An Integrative Review," *Biol. Psychol.*, 41(2): 103–146, Oct. 1995. [Online]. Available: https://doi.org/10.1016/0301-0511(95)05130-9
14. E. E. Smith, S. J. Reznik, J. L. Stewart, and J. J. B. Allen, "Assessing and Conceptualizing Frontal EEG Asymmetry: An Updated Primer on Recording, Processing, Analyzing, and Interpreting Frontal Alpha Asymmetry," *Int. J. Psychophysiol.*, 111: 98–114, Jan. 2017. [Online]. Available: https://doi.org/10.1016/j.ijpsycho.2016.11.005
15. M. Benedek and C. Kaernbach, "A Continuous Measure of Phasic Electrodermal Activity," *J. Neurosci. Methods.*, 190(1): 80–91, Jun. 2010. [Online]. Available: https://doi.org/10.1016/j.jneumeth.2010.04.028
16. T. Nishiyama, J. Sugenoya, T. Matsumoto, S. Iwase, and T. Mano, "Irregular Activation of Individual Sweat Glands in Human Sole Observed by a Videomicroscopy," *Auton. Neurosci.*, 88(1): 117–126, Apr. 2001. [Online]. Available: https://doi.org/10.1016/S1566-0702(01)00229-6
17. E. Maris and R. Oostenveld, "Nonparametric Statistical Testing of EEG- and MEG-Data," *J. Neurosci. Methods*, 164(1): 177–190, Aug. 2007. [Online]. Available: https://doi.org/10.1016/j.jneumeth.2007.03.024
18. S. Gerster, B. Namer, M. Elam, and D. R. Bach, "Testing a Linear Time Invariant Model for Skin Conductance Responses by Intraneural Recording and Stimulation," *Psychophysiology*, 55(2): e12986, 2018. [Online]. Available: https://doi.org/10.1111/psyp.12986
19. M. Kosinski, D. Stillwell, and T. Graepel, "Private Traits and Attributes Are Predictable From Digital Records of Human Behavior," *Proc. Natl. Acad. Sci.*, 110(15): 5802–5805, Apr. 2013. [Online]. Available: https://doi.org/10.1073/pnas.1218772110
20. C. Nebeker, J. Harlow, R. E. Giacinto, R. Orozco-Linares, C. S. Bloss, and N. Weibel, "Ethical and Regulatory Challenges of Research Using Pervasive Sensing and Other Emerging Technologies: IRB Perspectives," *AJOB Empirical Bioethics*, 8(4): 266–276, Oct. 2017. [Online]. Available: https://doi.org/10.1080/23294515.2017.1403980
21. S. J. Elliott, S. A. Massie, and M. J. Sutton, "The Perception of Biometric Technology: A Survey," 2007 IEEE Workshop on Automatic Identification Advanced Technologies, pp. 259–264, June 2007. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4263251
22. M. B. Kugler, "From Identification to Identity Theft: Public Perceptions of Biometric Privacy Harms," U.C. Irvine L. Rev. 107, Oct. 2019. [Online]. Available: https://scholarship.law.uci.edu/ucilr/vol10/iss1/5

About the Authors

Clayton P. Mosher received his PhD in neuroscience from The University of Arizona and is the lead scientist of neurotechnology at the Warner Media Neuro Lab, Burbank, CA. His research focuses on social behavior and emotion and his studies have been published in a variety of journals spanning a breadth of topics: from how neurons in the brain facilitate eye-contact to how emotions cause our hearts to race and our palms to sweat. He is an active member of the Society for Neuroscience and the Society for Psychophysiological Research.

Brian Wellner is the director of digital platforms at Warner Brothers, Burbank, CA, where he leads the Neurotechnology Team, which is responsible for building and growing technology solutions for research and development and immersive content development purposes. The team consists of software engineers, UI/UX designers, product managers, neuroscientists, and data scientists. Wellner's career spans more than 20 years in the media and entertainment industry. All of his past roles have been focused on productizing new and innovative technologies and applying them to businesses that are going through disruption due to evolving technology and changes in consumer behavior.

SMPTE