Evaluating Video Prototypes with Emotion Analytics

Bruce Wan is an assistant professor at the School of Design at The Hong Kong Polytechnic University. He is interested in technology-mediated experience design that supports human flourishing in the field of tourism and leisure. He is currently working on experience-based design literacy, a knowledge base that facilitates the co-design process with multi-disciplinary researchers for smart tourism innovation.
Bruce.Wan@polyu.edu.hk

Emotions have an important influence on how we think and behave. Incorporating emotion analytics in the video prototype evaluation process allows service designers to gather feedback from viewers based on moments of high emotional arousal in videos. This article explains how to incorporate emotion analytics into semi-structured interviews when evaluating video prototypes. We conducted a study and found the proper application of this method provided valuable feedback for enhancing the service concept.

Video prototyping and emotion analytics

Video prototyping is a storytelling method for communicating service concepts. The process involves service designers creating an explanation video that highlights the user experience as developed through multiple service touchpoints under different situations. To evaluate the video prototype, designers conduct semi-structured interviews with viewers after they watched the video. Viewers then share their opinions and the emotions they felt while watching specific interactions, features and touchpoints.

A typical problem is that a viewer's feedback may be limited by the capacity of his or her short-term memory, which normally lasts for 20-30 seconds, and can store between five and nine items. An explanation video can easily exceed this duration and as well as present more things to react to emotionally. Service designers can benefit from a way to help viewers remember what they felt as they watched the video. Emotion analytics may help here, by capturing viewers' emotions while they watch a video, complementing conventional semi- structured interviews and enhancing the extent and quality of feedback gained during the subsequent interview.

Emotion refers to strong feelings caused by the situation someone is in or the people one is with. Such instinctive feelings, which are distinct from cognitive reasoning, play a vital role in our survival and help us avoid danger and strive for greater fulfilment in life. Emotion analytics is a non-intrusive technique used to record and measure viewer's emotional reactions. Data related to emotions can come from different sources, such as text, speech, facial expressions and body language. That data is analysed and interpreted using machine learning algorithms.

Emotion analytics is capable of identifying and distinguishing basic emotions such as happiness, sadness, anger and fear. The advertising, marketing and entertainment industries have increasingly relied on emotion analytics to gain insights into customers' emotions.

We used the emotion analytics platform 'Affectiva'[1] for our study. The platform allows users to set up an emotion analytics study by inputting a link to a video hosted on YouTube[2]. A camera above the monitor records viewers' facial expressions while the video plays. The platform can calculate – in real time – the likelihood that the viewer is feeling one or more of five basic emotions: joy, anger, disgust, contempt or surprise, and does so with an accuracy in the 90th percentile.

Facial expressions can have either a positive or a negative effect on the likelihood of an emotion. Calculating that likelihood requires the system to recognise and map certain facial expressions to emotions. For example, if the viewer smiles, the software will increase the likelihood of joy, whereas raised or furrowed eyebrows will decrease its likelihood[3]. The strength (i.e., likelihood) of each emotion is represented as a line chart plotted below the video. After the video ends, those carrying out the research can identify video segments that triggered a viewer's specific emotion and investigate them in the interview (Fig. 2). In sum, using emotion analytics can help service designers gather more

1 https://www.affectiva.com/
2 https://affectiva.github.io/youtube-demo/
3 https://developer.affectiva.com/mapping-expressions-to-emotions/

Segment	Feature	Description	Moment
1	Daily advice	Start your day with supportive advice	Morning
2	Mini-game	Get inspired by the wizards	Commuting
3	Reminder	Recall what you have learnt from others	Working
4	Emotion logging	Log your feelings	When you feel agitated
5	Strength developer	Get motivated to develop strength of character	Facilitate leisure activity
6	Guided journaling	Reflect on your day	Evening
7	Goal-setting	Set a goal	Evening
8	Choose the wizard	Decide on optimal ways to interact with others	Before bed

Table 1. Segments and features that appear in the video

precise and valuable feedback for improving the quality of the service concept under investigation.

Research design and data collection

This study used a video prototype of a capstone project designed by two undergraduate students[4] at the School of Design, at The Hong Kong Polytechnic University. The project, titled 'Refles', is a mobile application that promotes the psychological well-being of city dwellers via self-reflection and is based on scientifically proven studies of the 'Character Strengths and Virtues' framework[5]. It offers a guided reflection tool that turns self-reflection into a daily practice. When used purposefully, these character strengths can boost happiness. Refles personifies these character strengths as different virtual characters (known as the wizards). Users can interact with these wizards to learn about and develop their character strengths.

For example, the wizards invite users to reflect on their activities before going to bed. Using Refles has the potential for the users to build greater connections with themselves by recording their individual thoughts, behaviours and character strengths. Despite the positive potential, the designers were uncertain of the desirability of new behaviours. Therefore, they produced a four-minute explanation video[6], comprising eight segments and features to communicate key features, information and context of use, in order to elicit viewers' feedback.

Ten people participated in this study, none of whom had ever seen the video, because we wanted to gather a more instinctive emotional response from them. Before each session, we masked the line chart area with paper so it would not distract the participant's attention. We briefly introduced the project, then seated them in front of a computer and played the video via the Affectiva platform. After they watched it, we conducted a semi-structured interview with participants, inviting them to talk about specific segments that intrigued them. Then, we revealed the line chart that recorded the participants' emotions and invited them to elaborate on the segments that triggered their emotions. The interviews were audio recorded, transcribed, coded and thematically analysed.

4 The capstone project was designed by Nathan Hui and Chester Wan, mentored by Bruce Wan, BA (Hons) Interactive Media, School of Design, The Hong Kong Polytechnic University, Hong Kong.
5 https://www.viacharacter.org/
6 http://tiny.cc/refles

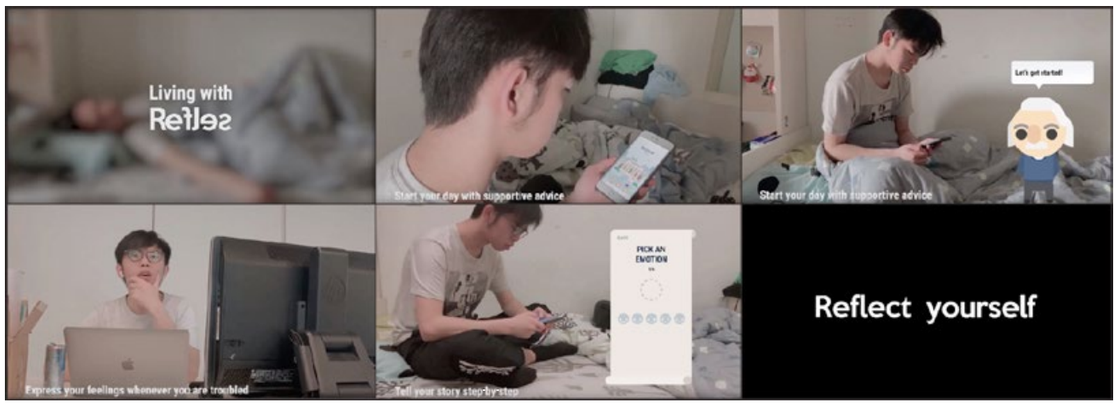

Fig. 1: Excerpt from the video[6]

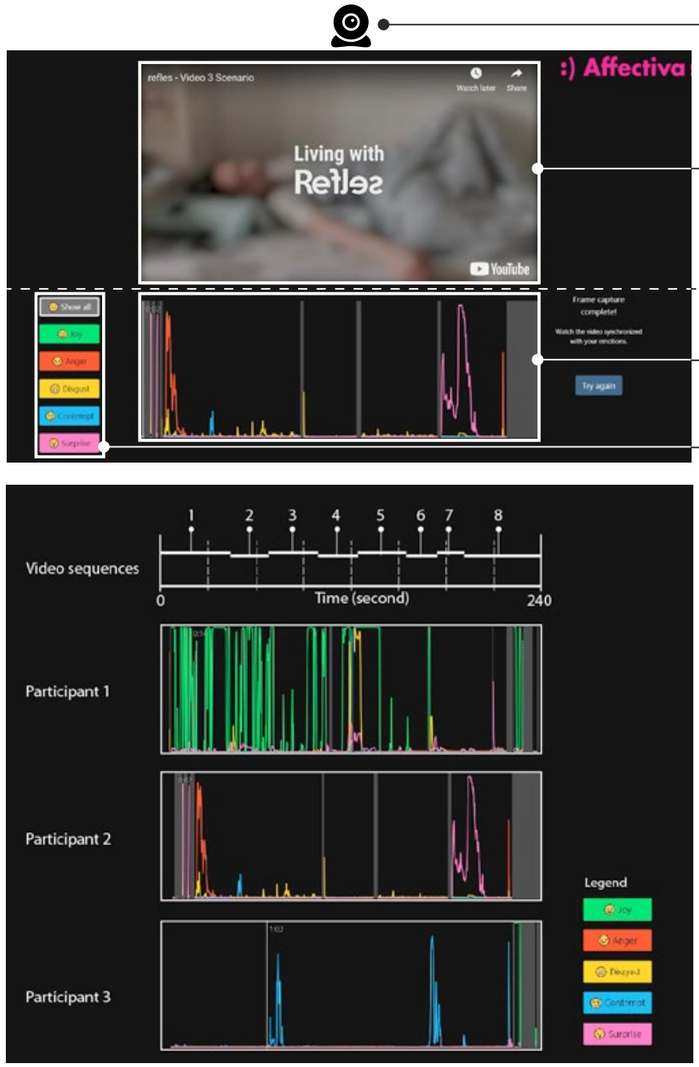

Fig. 2: (Top) The context of the study, in which a video was played and viewer's reactions were simultaneously captured and analysed
Fig. 3: (Left) Samples of emotion analytics of three participants mapped with the video sequences

Key findings

All participants were able to describe the key features that appeared in the video. At this stage, some appreciated these features because they believed the application could foster positive behaviours, whereas others questioned the feasibility and practicality of the system in a real-life setting.

We then proceeded to the second step of the interview in which we disclosed the emotion analytics results to the participants. Figure 3 shows the results of three participants who viewed the same video. The three participants were selected because their feedback was valuable and representative despite the diverse emotional reactions.

The analytics results were distinct in terms of the emotion (colour), likelihood (Y-axis), and occurrence (X-axis). For instance, Participant 1 mostly expressed joy on the first two sequences of the video, Participant 2 showed both surprise and anger at the first, seventh and the last sequences of the video, and Participant 3 expressed contempt at the second, sixth and the last sequences.

We believe the diverse emotional expressions might be due to the subjective and idiosyncratic nature of the emotional reaction. Certainly, the video content plays a role in triggering viewers' emotions; however, viewers can have very different emotional reactions towards the same content because of the individual's previous experience.[7]

We interviewed participants about each of the video sequences that had triggered a strong emotional moment (i.e., likelihood), and asked whether the type of emotion identified by the system was accurate. Because the goal was to improve the service concept, we were especially interested in negative emotional reactions such as the anger and disgust experienced by Participants 1 (sequence 2, 4 and 6) and 2 (sequence 1 and 3) during the video. However, both participants claimed that they did not feel particularly negative towards these sequences in the video, but were rather distracted by them (Participant 1) or confused and doubtful about the feasibility of the feature (Participant 2). More in-depth discussions with each participant revealed some moments of high emotional arousal to be highly valuable, whereas others represented little value in terms of feedback.

Signals: Valuable high emotional arousal moments

We found two types of feedback can fulfil our research goal: feedback on the features and content of the design solution, and feedback on the participants' personal interest. We found that the participants could elaborate on their strong emotional moments and thereafter make valuable suggestions regarding the service concept.

1. Features and content of the design solution

This is arguably the most valuable feedback the design team wanted to hear because the participants' emotional reactions were highly relevant to the service design concept, as shown in the video. Some participants expressed joy and surprise when they found that the features showcased were valuable. For example, some participants regarded the feature 'guided journaling' (sequence 6) as particularly valuable, although their emotions were distinct. Participant 2 provided more details on two moments she had not disclosed in the first interview. She clarified that she felt confused rather than angry during two intervals of the video because the feature's daily advice (sequence 1) and goal setting (sequence 7) did not appeal to her. In fact, she was worried these features would dampen her sense of autonomy. Here, the results of emotion analytics provided useful discussion prompts to facilitate in-depth interviews.

2. Participants' personal interest

Moments of high emotional arousal can be associated with participants' personal interest. For instance, one participant expressed joy (sequence 5, activity reminder) because she loved playing sports, just like the actor in the video. The sequence reminded her of good personal memories. In contrast, another participant felt disgusted by the same sequence because he had little interest in doing sport. Here, the negative emotion was not attributed to the feature but rather the content of the video. Designers should have a more in-depth discussion with participants to make sure viewer feedback helps to improve the product feature.

In sum, our interviews showed that participants appreciated the psychological benefits of using the application but also had concerns on the practicality of the system. Using emotion analytics allowed designers to probe into these concerns, such as adaptability to user's current habit and the intrusiveness of the reminders. The feedback allowed students to make adjustment on their final design.

Noise: Irrelevant high emotional arousal moments

We also found two types of high emotional arousal moments that contributed little or nothing to the video evaluation: participants' feelings about participating in the video evaluation and their feelings about the quality of storytelling.

1. Feelings about participating

Simply participating in the video evaluation session led to strong emotional reactions from some participants. For example, Participant 1 expressed excitement and joy at the beginning of the video because the experience was novel to her. However, her strong emotion was not related to the service concept. Interestingly, the video ended with a high emotional arousal moment for all three participants: They reported having an emotional reaction to finishing the task.

2. Reactions to the quality of storytelling

The quality of a narrative can also trigger participants' emotions. Inept acting, poor narrative, disturbing background music and unclear wording can distract viewers and thus increase noise. We found that too much noise (e.g., the first half of Participant 1) generates a chart that is a challenge for those carrying out the research to read. However, participant feedback on this issue can improve the quality of the narrative in future iterations. In this case, emotion analytics can be used to determine if the video is ready for a more formal evaluation.

Conclusion

Service designers can integrate emotion analytics into video prototype evaluation in four steps. First, by creating an explanation video that communicates a service concept. Secondly, by creating a list of video segments and features to help facilitate the interview. Thirdly, by using an emotion analytics platform to capture the viewer's facial expressions while they watch the video, and ensuring to conceal the results of the emotion analysis in order not to disturb the viewer. And finally, by conducting a semi-structured interview afterwards, using the line chart produced by emotion analytics tool as a discussion prompt.

On the whole, we recognised three important points when using analytics. First is that because emotional reactions are highly idiosyncratic, there may be little similarity between the line charts produced by different individuals. Second was that the result of emotion analytics contains both distracting noises as well as interesting signals, so researchers or designers need to distinguish between the two when collecting and analysing feedback. And last, compared with 27 human recognisable emotional experiences[8], current emotion analytics can only capture a handful of basic emotions. Therefore, researchers or designers should further discuss with viewers the emotion identified and, more importantly, flesh out the meaning behind the emotional reaction that was observed.

7 Bryce, E. (2017). How emotions are 'made': Why your definition of sadness is unlike anyone else's. Retrieved August 7, 2019, from https://www.wired.co.uk/article/lisa-feldman-barrett-emotions

8 Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. Proceedings of the National Academy of Sciences, 114(38), E7900–E7909. https://doi.org/10.1073/pnas.1702247114