Using physiological signals to measure the Quality-of-Experience of Health Care Professionals when interacting with a clinical guideline mobile app

Digital content adaptation and personalisation is a crucial component in increasing user engagement, and becoming of interest to designers/developers in areas related to clinical information delivery. In order to achieve this, new data-intensive methods are required that go beyond traditional user-centred design approaches. In this position paper, we discuss how although user-centred design has shown to be useful for generating generalised design guidelines (predominantly driven by qualitative data collection techniques), more quantitative methods and the use of measures such as Quality of Experience, could not only augment standard user research methods but also provide data to inform the adaption and personalisation of interfaces. In this paper we propose a solution-by-design to gather personal preferences through users’ physiological data (using pupillometry) and how it would be useful for applications such as mobile apps for clinical guidelines, where access to in-situ data collection is increasingly more challenging.

1. INTRODUCTION

Digital content adaption and personalisation is becoming more prevalent (Madan 2021). A popular example is the identification and delivery of interest-based digital content and services. Digital content providers (Netflix, Amazon Prime, Disney Plus etc.) use large amounts of visitor data to train deep learning algorithms, which in turn, predict a new user’s preferences based on their age group, geolocation, and other personal information. However, due to legal, ethical, and privacy requirements, using individuals’ information to even train machine / deep learning models is not recommended. Therefore, the research community is actively seeking different ways to address these issues, including federated learning, pseudonymisation, anonymisation, and encryption, as well as more traditional user research methods such as interviews and questionnaires.

In this paper we propose a solution-by-design to gather personal preferences through users’ physiological data for customising the presentation of digital content. This will pave a way for the development of customised digital applications/platforms in a variety of domains which improve the interactivity of digital devices and their content (Bergstrom et al. 2014); in the case of this position paper, applied to a clinical guidelines mobile app.

Utilising user-centred design (UCD) methodologies, the Bedside Clinical Guidelines (BCG) app has been developed to deliver 157 clinical guidelines to acute care clinicians via a mobile device (Mitchell et al. 2020). During the development of the BCG application, methods such as focus groups, think-aloud during basic clinical scenarios (lab-based), and system usability questionnaires were used to elicit feedback from clinical users. The findings from these mostly qualitative UCD
methods informed the production of a set of ‘general’ design guidelines for producing such applications and presenting information to users (Mitchell et al. 2021). This research demonstrated that the use of UCD methods can be useful in developing and producing generalised design guidelines for all users but also pointed to the fact that different types of clinicians may want guideline information displayed to them in a more personalised manner, dependent on their level of experience.

As access to clinical environments is often restricted in nature and the fact that clinicians also work in high-stress environments and frequently face interruptions in workflow, standard UCD methodologies are often limited and restricted to predefined scenarios or unrealistic settings (Mitchell et al. 2020). Methods such as these then can fail to gather representative data in terms of experience in a more realistic setting and do not gather appropriate types of data to inform adaption and personalisation of information presentation. Therefore, more implicit/passive and quantitative methods of gathering information that more accurately reflect the overall Quality of Experience (QoE) are required.

To access web based digital content, similar measures have been developed and used such as usefulness, desirability, accessibility, findability and credibility, (Richardson et al. 2021) through binary questionnaires. Furthermore, some UX tools have also been developed for in-use assessment for web products such as event-tracking, click-tracking, A/B split testing, visual feedback and Wireframing (Haije 2017). However, these tools are technology focused and are helpful in accessing the usefulness of the content provided but fail to address the customisation of the product or a service. Hence we need a merger of the two methodologies i.e. qualitative as well as quantitative analysis of user’s preferences. This paper then proposes utilising other sources of data such as physiological signals to measure QoE, in order to determine potential customisation and presentation of content in the BCG app.

2. USER CENTRED VS. TECHNOLOGY-CENTRED ASSESSMENT

As (Nielsen 2001) has pointed out “Don’t listen to your user, watch them doing it”. UCD related methods therefore quite often suffer from being too subjective, a lack of personalisation, differences in interpretation and not eliciting tacit knowledge. Analysis is often based on collective responses from multiple users and therefore, can serve a population, but not an individual’s preferences. Therefore, it is important to combine user centred research with technology-centred assessment of users’ experiences. QoE aims at solving this by considering all factors that contribute to a user’s perceived quality of a system or service, including system, human and contextual factors (Reiter et al. 2014). In the case of the the BCG app mentioned in the previous section, its development has so far been reliant on traditional user-centred methods and pre-determined scenarios. By considering QoE however, a more complete measurement of the users expectations feelings, perceptions, cognition and satisfaction can be captured (Laghari et al. 2011), creating not only a more meaningful assessment of user interaction but also more quantitative methods to utilise in-user modelling, adaptation and personalisation of content presentation.

3. ASSESSING MEANINGFUL USER INTERACTION

(Pine and Gilmore 2011) have mapped a user’s engagement experience to active participation and absorption during an interaction. The experience is then categorised into four realms as shown in Figure 1: Entertainment, education, escapist, and aesthetics. The spectator of a magic show, a football game, or a music concert shows an entertainment level engagement, while a wine taster or a cookery class participant experiences an educational engagement. Similarly, tourism, paragliding, or any unusual activity gives escapist engagement, and lastly hair cutting, drinking coffee in a coffee shop reflect an aesthetic experience. This basic user engagement can be mapped to a basic valence and arousal model also shown in Figure 1 and therefore, one model can be used to point out the other model’s state. For example, a person’s interaction with a clinical guideline can be ranked
based on his/her engagement level, and based on active engagement and absorption measures the user can be identified as the one gaining more knowledge. This can also be reflected by identifying the emotional status of the user from the arousal and valence measure and high valence and arousal values also depict the excitement and happiness of the user in turn reflecting an information gain. This gives enough base to combine the two models (Emotional-Model and Experience Realm) together to develop meaningful, attractive, and engaging interfaces for human-computer interaction.

Using a valence-arousal model we can classify a human’s emotional response from their physiological data. Physiological data include EEG, ECG, heart rate, Galvanic Skin Response, blood pressure, and pupil response (Dzedzickis et al. 2020; Rafique et al. 2021a). Sensors are usually required to be physically attached to the human body to get these physiological responses which make people uncomfortable and hence cannot be used on a large scale. But some of these responses can be collected in an unobtrusive manner such as pupil’s response using a camera. Using pupillometry for classifying human emotion is an emerging research domain Bergstrom et al. (2014); Ren et al. (2011); Rafique et al. (2021b) with significant interest from industry 1 2.

4. PUPILLOMETRY AS A METRIC FOR QOFE

Pupillometry is the study of the reactivity of the pupil in the human eye, this reactivity is measured by capturing the change in pupil size. A pupil constantly constricts or dilates due to the neurological activity in the human brain also described as cognition. High level cognition reflects mental effort, whereas neutral or low-level cognition shows satisfaction or uninterested behaviour. The relationship between eye movements and pupil responses can be used to interpret brain functions as shown in Table 1. Pupil responses are not discrete and therefore, temporal data analysis is required to assess the response accurately. Therefore, machine learning and deep learning methods can be employed to capture the constriction and dilation patterns (Rafique et al. 2021a).

This paper proposes that the pupil responses can also be used to assess the QoE from the technology’s perspective. Figure 2 shows low-cost but comfortable equipment for capturing pupil responses and gaze detection. Pupil size variations, fixation and eye tracking can be combined together to measure a QoE against individual functionalities of the technology under study without affecting normal activities.

\[
QoE_t(x, y) = \text{Emotion}_t(\text{valence}, \text{Arousal})
\]

where \(t\) represents the time

Relating this back to the BCG app, this method would be particularly useful in this scenario to make assessments in real-time at real locations i.e. in a hospital. Eye-tracking can be used to detect the function in use while pupil response can be used to understand the cognitive load. A proposed framework for assessing the technology-based QoE would include capturing a video recording of pupil size variation as well as the visual context through forward-facing camera. Furthermore, eye gaze will also be captured and recorded by the equipment shown in figure 2. This data can then be fed into a deep learning model as described in (Rafique et al. 2021a) to classify the user’s behaviour in terms of high-arousal, high-valence for specific functionality to be the most convenient to be used. A simple mathematical equation representing QoE is shown in Equation 1, where \(x, y\) represents the gaze location on the video frame showing visual content. Analysis of valance and arousal data for specific visual content needs to be done for a time interval defined by gaze detection module.

The proposed method is not suitable and hence not recommended for requirement analysis or getting feedback during development phase. It should be

| Reference | Study | Table 1: Interpretation of Brain Activity using Pupillometry |
|-----------|-------|----------------------------------------------------------|
| Hess (1965) | Interest in food and music | |
| Hahnenmann and Beatty (1967) | Decision Making | |
| Ellermeier and Wesphal (1995) | Pain estimation | |
| Ramsay et al. (2012) | Visual Processing | |
| Nowack et al. (2013) | Online information processing | |
| Wykowski et al. (2013) | Action-planning task | |
| Xiao and Brisson (2014) | Data Processing | |
| Zekveld and Kramer (2014) | Language processing | |
| Strohl et al. (2016) | User Experience | |
| Rafique et al. (2021a,b) | Emotions classification | |

1https://pupil-labs.com/
2https://www.tobii.com/
used after the product/service passes other software testing processes to avoid any technology-based bias in user’s behaviour.

5. CONCLUSION

In this position paper, we have proposed the use of QoE measurement, using pupillometry as a metric, to provide a more meaningful assessment of the user experience, by enabling in-context use, measured in real-time. This not only has the advantage of being more realistic, compared to the use of scenario-driven research completed in previous studies, but also the ability to collect quantitative data that can be used to model user behaviour. In areas such as the design of mobile apps for use in hospitals by clinicians, this offers a unique method for not only studying real-world usage but also how digital content can be adapted and personalised.

REFERENCES

Poonam Madan. Digital marketing: a review. V Paradigm shifts in management practices in the era of industry, 4: 64–71, 2021.

Jennifer Romano Bergstrom, Sabrina Duda, David Hawkins, and Mike McGill. Physiological response measurements. In Eye tracking in user experience design, pages 81–108. Elsevier, 2014.

James Mitchell, Ed de Quincey, Charles Pantin, and Naveed Mustafa. The development of a point of care clinical guidelines mobile application following a user-centred design approach. In International Conference on Human-Computer Interaction, pages 294–313. Springer, 2020.

James Mitchell, EJ De Quincey, Charles Pantin, Naveed Mustafa, et al. 15 usability recommendations for delivering clinical guidelines on mobile devices. In 34th British HCI Conference, pages 82–93. BCS Learning & Development, 2021.

Brianna Richardson, Marsha Campbell-Yeo, and Michael Smit. Mobile application user experience checklist: A tool to assess attention to core UX principles. International Journal of Human–Computer Interaction, 37(13):1283–1290, 2021.

Erin Gilliam Haije, 2017. URL https://mopinion.com/user-experience-ux-tools-tech-companies/. Last accessed: March 31, 2022.

Jakob Nielsen, 2001. URL https://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/. Last Accessed: March 31, 2022.

Ulrich Reiter, Kjell Brunström, Katrien De Moor, Mohamed-Chaker Larabi, Manuela Pereira, Antonio Pinheiro, Jyunyong You, and Andrej Zgank. Factors Influencing Quality of Experience, pages 55–72. Springer International Publishing, Cham, 2014. doi: 10.1007/978-3-319-02681-7_4.

Khalil ur Rehman Laghari, Noel Crespi, B. Molina, and C.E. Palau. Qoe aware service delivery in distributed environment. In 2011 IEEE Workshops of International Conference on Advanced Information Networking and Applications, pages 837–842, 2011. doi: 10.1109/WAINA.2011.58.

B Joseph Pine and James H Gilmore. The experience economy. Harvard Business Press, 2011.

Andrius Dziedzickis, Artūras Kaklauskas, and Vytautas Bucinskas. Human emotion recognition: Review of sensors and methods. Sensors, 20(3):592, 2020.

Sidra Rafique, Nadia Kanwal, Mohammad Samar Ansari, Mamooona Asghar, and Zuhair Akhtar. Deep learning based emotion classification with temporal pupillometry sequences. In 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), pages 1–6, 2021a. doi: 10.1109/ICECET52533.2021.9698663.

Peng Ren, Armando Barreto, Ying Gao, and Malek Adjoudi. Affective assessment of computer users based on processing the pupil diameter signal. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 2594–2597. IEEE, 2011.

Sidra Rafique, Nadia Kanwal, Irfan Karamat, Mamoon A Asghar, and Martin Fleury. Towards estimation of emotions from eye pupillometry with low-cost devices. IEEE Access, 9:5354–5370, 2021b. doi: 10.1109/ACCESS.2020.3048311.

Eckhard H Hess. Attitude and pupil size. Scientific american, 212(4):46–55, 1965.

Daniel Hahnemann and Jackson Beatty. Pupillary responses in a pitch-discrimination task. Perception & Psychophysics, 2(3):101–105, 1967.

Wolfgang Ellermeier and Wolfgang Westphal. Gender differences in pain ratings and pupil reactions to painful pressure stimuli. Pain, 61(3):435–439, 1995.

Thomas Z Ramsøy, Morten Fris-Oliviarius, Catrine Jacobsen, Simon B Jensen, and Martin Skov. Effects of perceptual uncertainty on arousal and preference across different visual domains. Journal of Neuroscience, Psychology, and Economics, 5(4):212, 2012.

Kati Nowack, Taciano L Milfont, and Elke van der Meer. Future versus present: Time perspective and pupillary response in a relatedness judgment task investigating temporal event knowledge. International Journal of Psychophysiology, 87(2):179–182, 2013.

Agnieszka Wykowska, Christine Anderl, Anna Schubö, and Bernhard Hommel. Motivation modulates visual attention: evidence from pupillometry. Frontiers in psychology, 4:59, 2013.
Using physiological signals to measure the Quality-of-Experience of Health Care Professionals when interacting with a clinical guideline mobile app

Mitchell • Kanwal • de Quincey

Sylvain Sirois and Julie Brisson. Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6):679–692, 2014.

Adriana A Zekveld and Sophia E Kramer. Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, 51(3):277–284, 2014.

Jonathan Strohl, Joseph Luchman, James Khun, Edward Pierce, and Kyle Andrews. Exploring the relationship between eye movements and pupillary response from formative user experience research. In *International Conference on Universal Access in Human-Computer Interaction*, pages 472–480. Springer, 2016.

Kassner Moritz and Will Patera. 2014. URL [https://pupil-labs.com/products/invisible/](https://pupil-labs.com/products/invisible/). Last Accessed: March 31, 2022.