Testing the Convergent Validity of Continuous Self-Perceived Measurement Systems: An Exploratory Study

Sébastien Lourties[(⊠)], Pierre-Majorique Léger, Sylvain Sénécal,
Marc Fredette, and Shang Lin Chen

HEC Montreal, Tech3Lab, 3000 Chemin de la Côte-Sainte-Catherine, Montreal,
QC H3T 2A7, Canada
sebastien.lourties@hec.ca

Abstract. This paper explores the convergent validity of instruments that can provide higher temporal resolution when measuring user experience: the continuous self-perceived measurement system and psychophysiological measures. Specifically, we explore the extent to which primacy and recency effects may have an impact on the convergent validity of two constructs: valence and arousal. Using a Wilcoxon signed rank test, results suggest that users self-evaluate their valence more accurately at the end of each of the sequences than at the beginning while they evaluate their arousal more accurately at the beginning of each of the sequences. This suggests that the recency effect has more impact on valence and the primacy effect has more impact on arousal. These findings contribute to human-computer interaction research by providing more information about the psychophysiological measures that cause recency and primacy.

Keywords: Continuous self-perceived measurement system
Physiological measurement · Primacy effect · Recency effect

1 Introduction

Research is calling for a multimethod approach in human computer interactions [26, 48, 53]. Multiple measurement approaches offer a richer perspective on user experience context and can enable UX researchers to gain better insight on what the user really experiences in a given task [11, 21].

Amongst the new methods proposed in the literature, new instruments offer better temporal resolution on user experience, i.e. measures that can provide a continuous measure of the experience over time. In contrast, non-continuous measure such as psychometric scales only provide a measure at a precise moment in time. Continuous measures can be very useful to designers as they can help to identify the timing of non-optimal experiences (in other words, pain points in an interactive experience). This article focuses on two of those measures: continuous self-perceived measurement systems (CSP) [12, 22, 23, 33, 41] and continuous psychophysiological measures [2, 18, 25, 32]. CSP are retrospective measurements that "let the observer track the

emotional content of a stimulus as they perceive it over time, allowing the emotional dynamics of speech episodes to be examined" [12] (p. 1). Psychophysiological measures are "an unobtrusive and implicit way to determine the user's affective or cognitive state on the basis of mind-body relations" [14] (p. 1362).

It is of high importance to UX researchers to understand the extent to which these new instruments converge in measuring the same constructs. As those signals can evolve over time, they may not to coevolve in the same manner over time. Thus, the objective of this paper is to explore the convergent validity of two important constructs in UX research: valence and arousal. Specifically, we explore the extent to which primacy and recency effect may have an effect on the convergent validity of these constructs.

To answer this research question, we have conducted laboratory experiment with 13 participants performing a series of utilitarian tasks on an insurance company website for 15 min. Our results suggest that users self-evaluate their valence more accurately at the end of each of the sequences than at the beginning and evaluate their arousal more accurately at the beginning of each of the sequence. This could suggest that valence has more impact on recency effect and arousal has more impact on primacy effect.

This paper is organized as follows. First of all, in the literature review, we will introduce emotion in user experience, then we will talk about psychophysiological measures of emotion, then its self-perceived evaluation. After that, we will develop our hypothesis and we will explain our research methodology. Finally, we will present the results and discuss the implications.

2 Literature Review: Measuring UX with High Temporal Resolution Instruments

This article focuses on two types of UX measurements that provide high temporal resolution: Neurophysiological measure and CSP measure. High temporal measures offer precision with respect to time during an experience. High temporal resolution measures are characterized by a sampling rate that defines the number of measures per minutes.

2.1 Continuous Self-Perceived Measures of User Experience

Continuous self-perceived measurement systems have been proposed as a novel way to enable a user to dynamically report on its experience. The simplest tools to use are composed of only one dimension of the emotion (e.g. emotional valence). The CARMA software and the Emotion Slider [22, 33] were developed with one dimension in order to facilitate the report of basic emotion (negative vs positive), so participants just have to push up or down to report their affective state. For instance, Girard and Wright [23] propose a measurement system in which users operate a joystick to indicate their reactions to a stimulus on two dimensions (e.g., emotional arousal and valence). Before that, other systems have been proposed to measure in a such way emotion. Feel-Trace [12] and EmuJoy [41] were the first software packages to propose a CSP on two dimensions. Software with two dimensions were only tested in a hedonic

context: FeelTrace, Emu-joy and DARMA were tested in the music or commercial ad context that had extreme (negative or positive) arousal and valence. Furthermore, even if the authors suggest the use of the tools for retrospection in a utilitarian context, they were mainly tested directly to evaluate music or commercials.

Although these systems provide richer information than self-reported scales due to their continuous nature, to our knowledge, no research has investigated their validity. Since research suggests that any self-reported measure is subject to biases, e.g., retrospective bias, efforts to validate these continuous measurement systems will inform researchers and practitioners using such systems to evaluate user experience.

Then, we could wonder why some researchers have focus their activity on CSP. It comes from the limitations of the traditional self-reported measurement system [3]. However, as Lallemand and Grenier [31] report, the limited number of emotions reported in some tools could make it difficult for the participant to clearly express what they feel. Finally, many tools provide a scale and a score that doesn't explain what really happened during the interaction with the product/interface even if much progress has been made by researchers with the succession of tools put at our disposal. All of the self-reported tools provide the overall grade that the participant experience without explaining positive and negative emotions selected during their interaction [7, 8]. Moreover, this last limitation motivated researchers to explore CSP measurement systems. Also, it should be noted that, depending on the tool we use, important biases such as primacy and recency effects, could occur during the retrospection of the participant.

2.2   Continuous Neurophysiological Measures of User Experience

In the last decade, there has been a wide range of studies that employed tools and theories from neurosciences to inform the design of information systems in a user-computer interaction context [42, 48]. More specifically, measures of valence and arousal have been very informative [34]. Valence and arousal measurements are part of psychophysiological measures that are already defined above as "an unobtrusive and implicit way to determine the user's affective or cognitive state on the basis of mind-body relations" [14] (p. 1362). Over the years, psychophysiological measures have received a lot of attention as they provide a large quantity of information to pose an accurate diagnostic on an emotional reaction caused by a stimuli [19]. Researcher such as Plutchik, [44] proposed that there are from two to twenty different emotions. In this article, we will focus our intention on the circumplex model that was created to represent emotions with arousal and valence in two dimensions [45, 49], because it has been reused for the construction of the continuous self-perceived system DARMA [23].

Arousal is the reaction to an emotional stimuli characterized by neurophysiological level of vigilance or a state of attention [29]. The arousal construct is used to contrast states of low arousal (e.g., calm) and high arousal (e.g., surprise) [2] and is a useful construct in the human computer literature. For example, Ravaja et al. [46], in a video game context, find that the nature of the opponent (computer, friend, stranger) varied the impact on the arousal of the participant. To measure arousal, electrodermal activity (EDA) has been proposed as an accurate way to collect data using the variation of the

electrical properties of the skin in response to the secretion of sweat from the palm of the hands in our case [2, 25].

Valence usually refers to "the 'positive' and 'negative' character of an emotion and/or of its aspects such as behavior, affect, evaluation, faces, adaptive value, etc." [9]. The valence construct is used to contrast states of pleasure (e.g., happy) and displeasure (e.g., angry) [2]. For example, in the IT field, this construct was used as a determinant factor to create an intelligent tutoring system (AutoTutor) that measures the frustration of the student learning the content and interacting with it [39]. Detecting this negative valence makes it possible to improve the learning curve of the student. Sas et al. [50], conducted an experiment to test most memorable moments and the attachment to the Facebook website and positive valence common denominators for all the participants. One way to measure this valence, is to use facial expression [15] because most individuals express their emotion with micro-movements of the facial muscles. Automatic facial analysis tools such as Facereader (Noldus, Wageningen, Netherlands) permit to determine with relative precision, valence change in an experience and this, in real time with a high temporal precision (6 times per second) [52].

There are several advantages to taking psychophysiological measures into consideration when studying emotion. First, you can capture the unconscious or automatic reactions of a participant without interrupting the experiment to get the most natural reaction because it is non-intrusive [42]. Then, you can have data on the user reaction in real time which is a real advantage to avoid the common bias of memory [26]. Finally, you can combine different measures with neurophysiological tools to limits the common method bias [34].

In this article, we are measuring the neurophysiological states of the user using valence and arousal dimensions of emotion because we are then able to use these two dimensions that are well accepted by research [3], characterized by a two-dimensional cartesian plane. The arousal dimension goes from calm to excited whereas valence ranges from negative to positive.

3 Hypothesis Development

While remembering an experience could be hard for someone, there exist some biases that consciously or unconsciously make this step harder.

This article aims at exploring the convergent validity of neurophysiological measures and CSP measures. Specifically, we explore to what extent the convergent validity is affected by primacy and recency bias. Primacy effect refers to the increased memory and over-weighted influence of the first moment of an experience [40, 51, 54] and recency effect correspond to the important influence of the last moment of an experience [28]. These two biases have the same root and are part of the larger topic of memory bias. To be concerned by the primacy and the recency effect you have to exercise a retrospection which is the capacity to remember the context of an experience and to explain the intention [10]. Several factors can influence the ability to remember an experience; valence and arousal are important ones.

Fredrickson and Kahneman [20] have suggested that the duration of an experience has little effect on the retrospective evaluation. They call it the "duration neglect". They conducted an experiment with video clips, but also with physical experiences [28] suggesting the same results. Overall, in a free recall context, primacy and recency effect play a U-shape curve described in the serial position curve of Craik and Lockhart [13]. After the experience, items at the beginning and at the end of a series are better reported than those in the middle.

We Posit H1: Primacy and Recency Effect Have the Same Influence on the Accuracy of Self-reported Evaluation

Fredrickson and Kahneman's results [28] suggested that adding a positive experience at the end of the task improve the retrospective evaluation. Furthermore, researchers have found that an experience with a strong emotional relevance is more likely to be remembered than a more neutral experience [4, 24]. Valence and arousal contribute in different ways to the formation of memory, but studies have shown that arousal is a much more important factor when it is time to re-member an experience [6]. First, the neural processes of memory are different for valence and arousal [30]. When remembering an experience with high arousal, different sub-stances such as glucose are released into the bloodstream [38] and the memory using emotional arousal leads to peripheral and central nervous systems activation and intensively involves the amygdala [37] whereas the prefrontal cortex hippocampal network is used for valence [30].

Moreover, recency and primacy effects are part of the selectivity bias. As Mather and Sutherland [36] explained, we are surrounded by information every day and there's a "battle for a share of our limited attention and memory, with the brain selecting the winners and discarding the losers" (p. 114). Indeed, arousal is enhancing the memory for specific details but is removing other collateral details [5]. Selectivity bias is common, our brain cannot recall each information that we have absorbed during the day, this is why recency and primacy effect have been found as one of the answer of this selection. Recency effect has been studied much more than the primacy effect due to the peak-end rule, characterized by the most intense moment at the end of an experience [28]. Different stimuli were used to test the recency effect such as medical pain [47] they found a strong relation between the intensity of pain recorded during the last 3 min of the treatment and the self-perceived evaluation.

We Posit H2: Arousal Has More Influence than Valence on the Recency Effect

4   Research Method

To test our hypothesis, we conducted a laboratory experiment with 13 subjects (7 males and 6 females), between the age of 21 and 48 (mean of 32). We provide a 100$ compensation to each participant upon completion of the experiment. Participants had normal or corrected-to-normal vision and were pre-screened for glasses, laser eye surgery, astigmatism, epilepsy, neurological and psychiatric diagnoses. This project was approved by the Ethics Committee of our institution.

4.1 Stimuli and Procedure

First, a scenario about an incident or a situation was presented to participants. Then, they had to perform a series of utilitarian tasks on website for 15 min in order to prepare a claim about an incident (e.g., provide details of the incident, make an appointment). Finally, they had to view their recorded interaction with the website and use a joystick to indicate their level of self-reported emotional arousal and valence continuously during the recording (Fig. 1). Every participant had the same task to perform and the same instruction to use the joystick. Moreover, we trained the participants to be sure that they understood the use of the joystick and the use of the dial. To make sure that the experiment was going well, we conducted two pre-tests to make sure the tools were working and recording the right data. Also, we wanted to be sure that the tasks performed as well as the instructions were understood by the participants. Minor changes were then made to finalize the protocol.
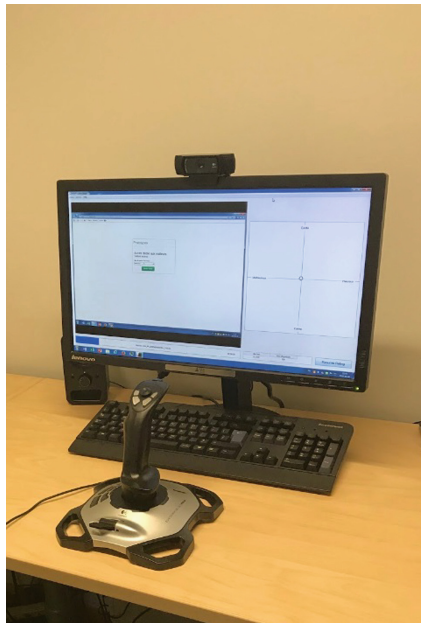

Fig. 1. The set up during the self-perceived measure using DARMA software [23]

4.2 Instruments and Measures

According to Ortiz de Guinea et al. [43] physiological measures and self-perceived evaluations interact together. Thus, to test the validity of CSP measurement systems, we assess them with neurophysiological inferences of the same constructs (emotional valence and activation) using physiological activation and automatic facial analysis [42].

We first of all measure emotion with physiological tools. Users' emotional and cognitive states can be measured with physiological signals such as electrodermal activity, heart rate, eye-tracking and facial expressions [42]. It allows researchers and practitioners the possibility of collecting real time information on what the user is experiencing through the interaction. Regarding electrodermal activity (EDA), it has been used to measure physiological arousal [2, 25]. Emotional valence represents the direction of an emotional response negative vs. positive [32]. In our case, facial expression and automatic facial analysis provide interesting insights for measuring human emotion [1, 16].

Participants physiological activation (i.e., electrodermal activity) during their interaction with the website was measured with the Acknowledge software mp150 sampled at 500 Hz (BIOPAC, Goleta, USA). The FaceReader software (Noldus, Wageningen, Netherlands) was used to measure emotional valence, calculated using the value of the positive emotion minus the strongest of the negative emotions which results in a valence score from −1 to +1 [17]. Media Recorder (Noldus, Wageningen, Netherlands) was used to record participants' interaction with the website. Observer XT (Noldus, Wageningen, Pays-Bas) was used to synchronize the signals of these three recording devices as per guidelines provided by Léger et al. [35]. Finally, participants self-reported continuous emotional reactions (arousal and valence) were measured using the DARMA software [23] (p. 1), which "synchronizes media playback and the continuous recording of two-dimensional measurements through the manipulation of a computer joystick to indicate changes in their emotional state" (see Fig. 2). The output of DARMA is an Excel file in which each line begins with a time code, then gives the valence and activation evaluation coordinates of the joystick at that time. All of the measures and instruments are summarized in Table 1.

Fig. 2. The two-dimension self-report window during media playback [23].

Table 1. Measures used for the experiment

	Self-reported (retrospective)	Physiological (during interaction)
Valence	Participant's joystick position on the valence axis at time t (Darma)	Facial emotional valence at time t (Facereader)
Arousal	Participant's joystick position on the arousal axis at time t (Darma)	Electrodermal activity for time t (Acqknowledge)

As you can notice, the software was set at 30 Hz, a bin size of 0.25 s and an axis magnitude of 100. These measures are summarized in the Table 1.

5   Results

To test the convergent validity of CSP measurement, we compared the retrospection performance of the first and last time interval between self-reported and physiological measures.

The validity of CSP emotion, i.e. the degree to which it accurately measures the emotions it was designed to measure, was assessed by recording sessions, broken down into short sequences. Each participant had to report their emotional valence and arousal for 14 short sequences for a total of 182 sequences for the overall sample. Sixty percent of tasks lasted less than 30 s. To test the convergent validity of CSP measurement, we compared the retrospection performance of the first and last time interval of each one of the 14 sequences between self-reported and physiological measures. Thus, we calculate the distance between physiological valence and CSP valence in the first and last time interval and the same calculation was done with the distance between physiological arousal and CSP arousal (respectively dis_valence_first vs dis_valence_last and dis_arousal_first vs dis_arousal_last in Table 2). When the movement of the joystick was less than .001, we considered it as static and these data were excluded from the analysis. The 2-tailed p-value from the Wilcoxon signed rank test indicates whether the difference is significant or not.

Table 2. Results for 1 s interval between the beginning and the end of each one of the sequences

Pair of variables (first and last 1 s)		Mean at beginning	Mean at end	P-value
distance_valence_first	distance_valence_last	.155	.1215	.0173
distance_arousal_first	distance_arousal_last	.0191	.0377	.0635

With an interval of 1 s, we find that participants report more accurately their valence at the end of the task compared to the beginning with a mean difference of .0335 (.155–.1215) and p-value = .0173. Also, participants report with more precision their arousal at the beginning of the task compared to the end with a mean difference of

.0186 (.0377–.0191) and p-value = .0635. Because sequences were short, the data analysis was performed using one-second time periods (e.g., comparing the first second of self-reported valence vs. physiological valence), but also with two-second periods.

Significant results were found for both time periods, the results for the one and two seconds time periods are reported in Tables 2 and 3.

Table 3. Result for 2 s interval between the beginning and the end of each one of the sequences

Pair of variables (first and last 2 s)		Mean at beginning	Mean at end	P-value
distance_valence_first	distance_valence_last	.153	.1235	.0143
distance_arousal_first	distance_arousal_last	.0133	.0216	.0436

Below you can find the name of the variables and their meaning:

- distance _valence_first: the reported valence minus the experienced valence at the first moment of each sequence
- distance _valence_last: the reported valence minus the experienced valence at the last moment of each sequence
- distance _arousal_first: the reported arousal minus the experienced arousal at the first moment of each sequence
- distance _arousal_last: the reported arousal minus the experienced arousal at the last moment of each sequence

6   Discussion

The objective of this first exploratory study was to test the convergent validity of CSP measurement systems with physiological emotional measures. First of all, results suggest that users self-evaluate their valence more accurately at the end of each of the sequences than at the beginning and more accurately their arousal at the beginning of each sequences. A second study will be conducted in the spring to further validate the results.

This paper confirms that continuous measurement allows for a richer self-perceived evaluation of emotion than traditional methods. Also, we show that there is a link between psychophysiological measurements such as facial expression analysis or electrodermal activity with the self-perceived evaluation of the participant. Regarding the primacy and recency effects our results confirm that they have the same influence, biasing the retrospection of participants [10, 13]. However, participants were more accurate at the beginning and at the end of their interaction, when reporting their emotion. Moreover, arousal has been proposed as the most influential factor regarding the self-evaluation of recency effect [28, 47]. With our experimental design, using continuous self-perceived measurement, we found that that users self-evaluate their valence more accurately at the end of each of the sequences than at the beginning, so our results do not converge with Cahill and McGaugh [6], proposing arousal as more

important factor than valence for retrospection in this specific context. This implies that primacy and recency effect have an influence in the way participants reported their emotion. It is important for researchers to have this result in mind when using tools such as DARMA software to explore the self-perceived evaluation because it could lead to the overestimation or underestimation of the results of an experiment.

Our findings can also be useful for the marketing or entertainment industry. With many tools to evaluate subjective emotion, our paper allows a better understanding of all the instruments that could be available for practitioners to conduct (continuous) self-reported measurement. Our results could also be interesting to have in mind when practitioners design their product/service. Indeed, if the valence is best reported at the end of the interaction with the website, you should probably focus on minimizing negative emotion and maximizing positive emotion during this sequence of time, and all the more so knowing that negative information have stronger influence on memory [27]. Regarding the arousal result that is more accurate at the beginning of the session, it depends on your goal, maximizing or minimizing arousal in accordance with the area of activity. Overall, designers could influence the user's memory of an experience when focusing on the beginning and the end of the task. Finally, during user experience testing, designers should first of all pay attention to these two periods of time.

Our experiment faced different limitations. Starting with technical limitations, when participants did not touch the joystick, it would automatically go back by default to the center which represents neutrality in terms of emotions. It is possible that some participants were influenced by this default position.

Regarding the experimental design, this initial study was performed in a specific utilitarian context, future research could use a different context in order to validate the results in a broader range of contexts.

For the results, with our data and tools that we used, we cannot know if participant under or overestimated their emotion when reporting. During our second data collection for this research project, we used IAPS pictures in order to calibrate emotional reaction with extreme stimuli in order to gain knowledge on the under/overestimated part. At this moment, we are still analyzing the results from this second phase.

7   Conclusion

Emotions and the way we measure them are destined to endless debate [9]. As Cockburn et al. [8] explained, an experience is internally composed of several sequences and is influenced by the most intense moment. Such an understanding accords more importance to psychophysiological measures and continuous self-perceived measures, as traditional post-task self-perceived measurements are more subject to bias [31]. In this paper, results suggest that an experience lived by a participant is not exactly the same as it is reported. Different biases may influence this misalignment of the "story". Primacy and recency through the influence of valence and arousal play an important role is the experience that is reported. Researchers in user experience have much to gain through a deeper understanding this topic.

References

1. Bartlett, M.S., Littlewort, G., Lainscsek, C., Fasel, I., Movellan, J.: Machine learning methods for fully automatic recognition of facial expressions and facial actions. In: IEEE International Conference Systems, Man Cybernetics, pp. 592–597 (2004)
2. Boucsein, W.: Electrodermal Activity. Springer, Heidelberg (2012). https://doi.org/10.1007/978-1-4614-1126-0
3. Bradley, M., Lang, P.J.: Measuring emotion: the self-assessment semantic differential manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry **25**(1), 49–59 (1994)
4. Bradley, M.M., Greenwald, M.K., Petry, M.C., Lang, P.J.: Remembering pictures: pleasure and arousal in memory. J. Exp. Psychol. Learn. Mem. Cogn. **18**(2), 379–390 (1992)
5. Burke, A., Heuer, F., Reisberg, D.: Remembering emotional events. Memory Cogn. **20**(3), 277–290 (1992)
6. Cahill, L., McGaugh, J.L.: A novel demonstration of enhance memory associated with emotional arousal. Conscious. Cogn. **4**, 410–421 (1995)
7. Cockburn, A., Quinn, P., Gutwin, C.: Examining the peak-end effects of subjective experience. In: CHI 2015, pp. 357–366 (2015)
8. Cockburn, A., Quinn, P., Gutwin, C.: The effects of interaction sequencing on user experience and preference. Int. J. Hum.-Comput. Stud. **108**, 89–104 (2017)
9. Colombetti, G.: Appraising valence. J. Conscious. Stud. **12**(24), 103–126 (2005)
10. Conway, M.A.: Cognitive Models of Memory. MIT Press, Cambridge (1997)
11. Courtemanche, F., Léger, P.-M., Dufresne, A., Fredette, M., Labonté-LeMoyne, É., Sénécal, S: Physiological heatmaps: a tool for visualizing users' emotional reactions. Multimed Tools Appl. **77**, 1–28 (2017)
12. Cowie, R., Douglas-Cowie, E., Savvidou, S., Mcmahon, E., Sawey, M., Schröder, M.: "Feeltrace": an instrument for recording perceived emotion in real time. In: ISCA ITRW on Speech and Emotion, pp. 19–24 (2000)
13. Craik, F.I.M., Lockhart, R.S.: Levels of processing: a framework for memory research. J. Verbal Learn. Verbal Behav. **11**(6), 671–684 (1972)
14. Dirican, A.C., Göktürk, M.: Psychophysiological measures of human cognitive states applied in Human Computer Interaction. Procedia Comput. Sci. **3**, 1361–1367 (2011)
15. Ekman, P.: Facial expression and emotion. Am. Psychol. **48**(4), 384–392 (1993)
16. Ekman, P., Friesen, W.: Constants across cultures in the face and emotion. J. Pers. Soc. Psychol. **17**(2), 124–129 (1972)
17. Ekman, P., Friesen, W.V.: Unmasking the Face. Prentice Hall, Cambridge (2003)
18. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure, 2nd edn. Morgan Kaufmann, Amsterdam (2004)
19. Fredette, M., Labonté-LeMoyne, É., Léger, P.-M., Courtemanche, F., Sénécal, S.: Research directions for methodological improvement of the statistical analysis of electroencephalography data collected in NeuroIS. In: Davis, F.D., Riedl, R., vom Brocke, J., Léger, P.-M. (eds.) Information Systems and Neuroscience. LNISO, vol. 10, pp. 201–206. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-18702-0_27
20. Fredrickson, B.L., Kahneman, D.: Duration neglect in retrospective evaluations of affective episodes. J. Pers. Soc. Psychol. **65**(1), 45–55 (1993)

21. Georges, V., Courtemanche, F., Sénécal, S., Léger, P.-M., Nacke, L., Pourchon, R.: The adoption of physiological measures as an evaluation tool in UX. In: Nah, F.F.-H. (eds.) HCIBGO 2017. LNCS, vol. 10293, pp. 90–98. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58481-2_8

22. Girard, J.M.: CARMA: software for continuous affect rating and media annotation. J. Open Res. Softw. **2**, 1–11 (2017)

23. Girard, J.M., Aidan, A.G.: DARMA: software for dual axis rating and media annotation. Behav. Res. Methods 1–8 (2017)

24. Hamann, S.: Cognitive and neural mechanisms of emotional memory. Trends Cogn. Sci. **5**(9), 394–400 (2001)

25. Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. Behav. Inf. Technol. **25**(2), 91–97 (2006)

26. Ortiz de Guinea, A., Webster, J.: An investigation of information systems use patterns: technological events as triggers, the effect of time, and consequences for performance. MIS Q. **37**(4), 1165–1188 (2013)

27. Ito, T.A., Larsen, J.T., Smith, N.K., Cacioppo, J.T.: Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. J. Pers. Soc. Psychol. **75**(4), 887–900 (1998)

28. Kahneman, D., Fredrickson, B.L., Schreiber, C.A., Redelmeier, D.: When more pain is preferred to less: adding a better end. Psychol. Sci. **4**(6), 401–405 (1993)

29. Kandel, E.R., Schwartz, J.H., Jessell, T.M.: Principles of Neural Science, 4th edn. McGraw-Hill, Health Professions Division, New York (2000)

30. Kensinger, E.A., Corkin, S.: Two routes to emotional memory: distinct neural processes for valence and arousal. Proc. Natl. Acad. Sci. U.S.A. **101**(9), 3310–3315 (2004)

31. Lallemand, C., Gronier, G.: Méthodes de design UX: 30 méthodes fondamentales pour concevoir et évaluer les systèmes interactifs. Eyrolles (2015)

32. Lane, R.D., Chua, P.M.-L., Dolan, R.D.: Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. Neuropsychologia **37**(9), 989–997 (1999)

33. Laurans, G., Desmet, P.M.A., Hekkert, P.: The emotion slider: a self-report device for the continuous measurement of emotion. In: Affective Computing and Intelligent Interaction and Workshops, ACII 2009, 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, pp. 1–6 (2009)

34. Léger, P.-M., Davis, F.D., Cronan, T.P., Perret, J.: Neurophysiological correlates of cognitive absorption in an enactive training context. Comput. Hum. Behav. **34**, 273–283 (2012)

35. Léger, P.-M., Titah, R., Sénecal, S., Fredette, M., Courtemanche, F., Labonté-LeMoyne, É., De Guinea, A.O.: Precision is in the eye of the beholder: application of eye fixation-related potentials to information systems research. J. Assoc. Inf. Syst. **15**(1), 651–678 (2014). Suppl. Special Issue on Methods, Tools, and Measurement

36. Mather, M., Sutherland, M.R.: Arousal-biased competition in perception and memory. Perspect. Psychol. Sci. **6**(2), 114–133 (2011)

37. McGaugh, J.L.: The amygdala modulates the consolidation of memories of emotionally arousing experiences. Annu. Rev. Neurosci. **27**, 1–28 (2004)

38. McGaugh, J.L., Cahill, L., Roozendaal, B.: Involvement of the amygdala in memory storage: interaction with other brain systems. Proc. Natl. Acad. Sci. U.S.A. **93**(24), 13508–13514 (1996)

39. Mello, S.K.D., Craig, S.D., Gholson, B., Franklin, S., Picard, R., Graesser, A.C.: Integrating affect sensors in an intelligent tutoring system. In: Affective Interactions: The Computer in the Affective Loop Workshop at 2005 International conference on Intelligent User Interfaces, pp. 7–13. ACM Press, New York (2005)
40. Murdoch, B.B., Lissner, E., Marvin, C.: The serial position effect of free recall. J. Exp. Psychol. **64**(5), 482–488 (1962)
41. Nagel, F., Kopiez, R., Grewe, O., Altenmuller, E.: EMuJoy: software for continuous measurement. Behav. Res. Methods **39**(2), 283–290 (2007)
42. Ortiz de Guinea, A., Titah, R., Leger, P.-M.: Explicit and implicit antecedents of users' information systems behavioral beliefs: a neuropsychological investigation. J. Manag. Inf. Syst. **30**(4), 179–210 (2014)
43. Ortiz de Guinea, A., Titah, R., Léger, P.-M.: Measure for measure: a two study multi-trait multi-method investigation of construct validity in IS research. Comput. Hum. Behav. **29**(3), 833–844 (2013)
44. Plutchik, R.: A general psychoevolutionary theory of emotion. In: Emotion: Theory, Research, and Experience, pp. 189–217. Academic Press, New York (1980)
45. Posner, J., Russell, J.A., Peterson, B.S.: The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev. Psychopathol. **17**(3), 715–734 (2008)
46. Ravaja, N., Saari, T., Turpeinen, M., Laarni, J., Salminen, M., Kivikangas, M.: Spatial presence and emotions during video game playing: does it matter with whom you play? Presence Teleoper. Virtual Environ. **15**(4), 381–392 (2006)
47. Redelmeier, D.A., Kahneman, D.: Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. Pain **66**(1), 3–8 (1996)
48. Riedl, R., Léger, P.-M.: Fundamentals of NeuroIS. Studies in Neuroscience, Psychology and Behavioral Economics, pp. 1–115. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-45091-8
49. Russel, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**(6), 1161–1178 (1980)
50. Sas, C., Dix, A., Hart, J., Ronghui, S.: Emotional experience on Facebook site. In: CHI 2009 Extended Abstracts on Human Factors in Computing Systems, pp. 4345–4350 (2009)
51. Shteingart, H., Neiman, T., Loewenstein, Y.: The role of first impression in operant learning. J. Exp. Psychol. Gen. **142**(2), 476–488 (2013)
52. Den Uyl, M.J., Van Kuilenburg, H.: The FaceReader: online facial expression recognition. In: Proceedings of Measuring Behavior, pp. 589–590 (2005)
53. Vermeeren, A.P.O.S., Law, E.L., Roto, V.: User experience evaluation methods: current state and development needs. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, pp. 521–530 (2010)
54. Zauberman, G., Diehl, K., Ariely, D.: Hedonic versus informational evaluations: Task dependent preferences for sequences of outcomes. J. Behav. Decis. Mak. **19**(3), 191–211 (2006)