Examining Joy of Use and Usability During Mobile Phone Interactions within a Multimodal Methods Approach

Katharina Lingelbach*
katharina.lingelbach@iao.fraunhofer.de
Fraunhofer Institute for Industrial Engineering IAO
Stuttgart, Germany

Nektaria Tagalidou
nektaria.tagalidou@iat.uni-stuttgart.de
Institute of Human Factors and Technology Management, University of Stuttgart
Stuttgart, Germany

Patrick S. Markey
patrick.markey@univie.ac.at
Huawei Technologies Duesseldorf GmbH, German Research Center
Munich, Germany,

Bettina Föll
bettina.foell@huawei.com
Huawei Technologies Duesseldorf GmbH, German Research Center
Munich, Germany,

Matthias Peissner
matthias.peissner@iao.fraunhofer.de
Fraunhofer Institute for Industrial Engineering IAO
Stuttgart, Germany

Mathias Vukelić
mathias.vukelic@iao.fraunhofer.de
Fraunhofer Institute for Industrial Engineering IAO
Stuttgart, Germany

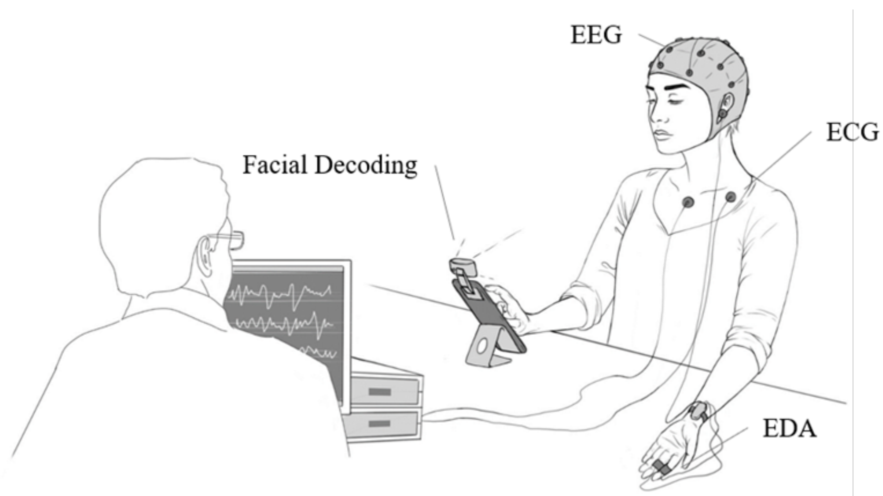

Figure 1: Overview of the experimental set-up and illustrative placement of the sensors.

ABSTRACT

Objective: We investigate experienced joy of use (JoU) and usability using a multimodal methods approach by systematically varying mobile phone interactions. **Methods**: We combined subjective and objective measures to investigate whether positive emotional experiences and moments of joy during the interaction can be distinguished from neutral and negative emotional experiences. In a study with 30 participants, electroencephalography (EEG), electrocardiography (ECG), electrodermal activity (EDA), facial emotion recognition, and questionnaires were used. **Results**: There were greater positive experiences in interactions designed to elicit JoU, even under bad usability. We did not observe a difference between the conditions in the EEG indices. However, a higher heart rate and components in the EDA phasic response as well as facial muscle activity associated with anger were linked to good usability combined with no JoU. **Conclusion**: The multimodal methods approach reveals great potential to investigate JoU and usability in naturalistic scenarios. **Application**: The developed framework provides a groundwork to evaluate and improve interactions with technology. Thereby, users and their emotional experiences are placed at the centre when designing user interfaces. By detecting moments of joy, this approach can support a better understanding of how technology can be purposefully designed for joyful experiences.

MuC '22, September 4–7, 2022, Darmstadt, Germany

ACM ISBN 978-1-4503-9690-5/22/09.
https://doi.org/10.1145/3543758.3543771

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; **User studies**; **Usability testing**; **Field studies**.

KEYWORDS

user experience, affective computing, multimodal measures, electroencephalography, electrocardiography, electrodermal activity, facial expression decoding, joy of use, human-machine interaction

ACM Reference Format:

Katharina Lingelbach, Nektaria Tagalidou, Patrick S. Markey, Bettina Föll, Matthias Peissner, and Mathias Vukelić. 2022. Examining Joy of Use and Usability During Mobile Phone Interactions within a Multimodal Methods Approach. In *Mensch und Computer 2022 (MuC '22), September 4–7, 2022, Darmstadt, Germany*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3543758.3543771

1 INTRODUCTION

When we interact with technology, it is the experienced emotions, such as joy or frustration that determine how we evaluate the interaction and the product itself [42, 46]. While former studies have focused primarily on usability factors to predict user judgments, recent research increasingly includes emotional dimensions for explaining user evaluations. Previous approaches assume that overall user experience is determined in a hierarchical way, with functionality and then usability acting as the most important factors, and other factors like pleasure or joy playing a less crucial role (e.g., [34]). This has been called into question, with authors claiming that hedonic elements contributing to the joy of use (JoU), such as fun or psychological need fulfilment, are just as important if not even more important than basic usability (e.g., [25]). Several studies and models aimed to describe the relationship between the two components usability and JoU for the overall user experience. Hassenzahl et al. [26] introduced two factors affecting users' evaluations: (1) pragmatic quality describing goal-oriented functions and perceived usability, and (2) hedonic quality describing rather facets of positive user experience and JoU such as psychological needs of novelty and augmentation [27, 30].

1.1 The usability-joy of use framework

Therefore, in order to adequately understand the psychology of human-technology interaction, we will need two complementary perspectives: (1) *usability (U)*, which is the "extent to which a product can be used (…) to achieve specified goals with effectiveness, efficiency, and satisfaction (…)"[1], and (2) *joy of use (JoU)* which is the positive hedonic assessment. Figure 2 provides an overview of the framework underlying this study.

JoU describes a holistic concept beyond pure task fulfilment. It is determined by (i) aesthetic properties [5, 11, 79], (ii) moments of »Wow« (i.e., positively perceived surprise, [74]), and (iii) fulfilment of individual needs [28, 56, 68]. Aesthetic properties can be either of narrative (e.g., deeper meaning) or perceptual nature [45]. They are associated with both a hedonic, beautiful is useful, as well as a pragmatic halo effect, useful is beautiful [49]. Both effects lead to higher quality perceptions, attractiveness, and intention to use a product [5]. Surprise and related concepts such as Wow, have to date received relatively sparse literature coverage and understanding.

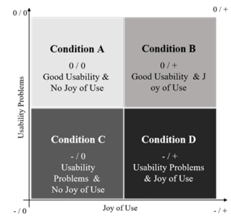

Figure 2: Theoretical framework of the study: Usability (y-axis) and Joy of Use (x-axis) are systematically varied.

Desmet and Fokkinga [14], Desmet et al. [15] measured the Wow experience of products and described Wow as an emotional experience comprising three factors (1) fascination, (2) desire, and (3) pleasant surprise. Väänänen-Vainio-Mattila et al. [74] examined Wow effects in the context of cloud applications with qualitative user interviews and suggested long-term positive feelings towards such products or services. However, they describe challenges like inappropriate scenarios (e.g., under stressful conditions) or potential cultural differences regarding norms and expectations for evoking Wow. Gross and Thüring [23] explored surprise and unexpected events with either desirable or undesirable consequences. Their results suggest that only positive surprises enhancing goal achievement without distracting or impeding the interaction and goal-directed behaviour were evaluated positively (see also [22]).

Several studies examined how fulfilment of psychological needs is associated with satisfaction and positive emotions [14, 60, 67]. Further research extracted needs with specific relevance for positive user experiences during the interaction with technology [18, 24]. For the current study, we selected five needs that were considered as particularly important in the specific context of mobile phone interaction. The consideration was done in an extensive workshop with professionals in the field of mobile phone interactions and researchers in needs-based design. Needs generally evoke positive user experience [68]. The following five needs were evaluated as particularly important to create a positive user experience while using a mobile phone:

- stimulation,
- relatedness,
- competence,
- self-expression/autonomy,
- and luxury.

1.2 Multimodal methods approach

A majority of studies on positive user experience use subjective evaluations like questionnaires (see [75] for a review), albeit these methods are prone to potential bias of post-hoc evaluations and socially desirable answers [51]. Objective physiological (e.g., electrodermal activity; EDA, and electrocardiography; ECG) and neurophysiological measures (e.g., electroencephalography, EEG) provide an unbiased alternative and allow to detect emotional evaluations without disturbing the interaction or task [75]. The use of a multimodal methods approach stems from the idea that insights from different sources can (1) complement each other and (2) compensate respective weaknesses [50, 57]. The combination of subjective and (neuro-)physiological approaches enables a comprehensive, holistic, and robust understanding of the phenomena investigated [57]. The circumplex model of affect [63] suggests that affective states are composed by two diametric neurophysiological systems, namely (a) valence (i.e., continuous evaluation of pleasant vs. unpleasant states) and, (b) arousal (i.e., continuous measure of physical activation; see [59], for a review). Joy and hence JoU, is strongly positively correlated with valence and moderately to strongly correlated with arousal [5, 59]. While physiological measures such as heart rate variability or electrodermal activity are well-suited to capture arousal [19, 43, 77], neurophysiological methods measuring brain activity (e.g., captured via EEG recordings) allow identifying differences in valence (e.g., [9]). The frontal alpha asymmetry (FAA) measured in EEG is well-suited as an index distinguishing positive and negative emotions [2, 36, 37, 73]. Stimuli rated as positive are associated with greater relative left-hemispheric alpha activation in frontal electrodes, while stimuli rated as negative reveal greater relative right-hemispheric frontal activation [2, 6, 9]. Ekman [16] introduced a further theory with discrete basic emotions cross-culturally shared and distinguishable via facial muscle activity [62]. Recent studies revealed the potential of facial expressions to predict consumer liking, preference, and intent to purchase a product [47, 64]. Especially camera-based facial expression recognition is a suitable, easy-to-use method to identify positive and negative user experience via computer algorithms [55].

In this study, we, therefore, explored a converging multimodal methods approach combining subjective and objective measures to investigate (1) whether positive and negative emotional experiences can be distinguished based on activation patterns of the central and peripheral nervous system and (2) whether we can detect moments of joy during the interaction with technology. We examined the specific relation and interdependence between usability and JoU in a naturalistic setting by using typical mobile phone interaction scenarios. (1) We hypothesize that high levels of JoU are associated with positive emotional experiences reflected in more positive subjective user ratings (valence, Wow, visual aesthetics, fulfilled needs), as well as greater left-hemispheric frontal alpha activation, a higher engagement index and lower workload index, and activation in facial muscles associated with positive emotions. Low levels of usability, however, should elicit negative emotions and lower valence ratings compared to positive emotions. Furthermore, we explored (2) which aspects of the user interaction design can evoke positive user experiences. For this purpose, we systematically varied the usability and JoU in the interaction scenarios.

2 METHODS

2.1 Participants

A total of 30 healthy, examined by self-reports in an online screening, participants took part in this study (19 female; age: 18 to 36 years with M = 25.50; SD = 4.93). Participants were recruited through the participant pool of the Fraunhofer IAO and received monetary reward to compensate for time and travel costs. All participants had normal or corrected-to-normal vision and were screened for their suitability to participate in an EEG study, to be right-handed (since influences of handedness on hemispheric asymmetry cannot be excluded; [52]), and to show adequate language skills. Prior to the experiment, they signed a written informed consent according to the recommendations of the declaration of Helsinki. The study was approved by the local ethics committee of the Medical Faculty of the University of Tuebingen, Germany (ID: 886/2019BO2). Four participants had to be excluded in the EEG analysis because no artefact-free EEG signal could be obtained. For the analysis of facial expressions, another set of four participants were excluded due to technical problems during the recording.

2.2 Materials

We measured EDA and ECG (arousal) as well as EEG and camera-based facial expression recognition (valence) during the interaction. Stimuli were presented on a Mate 20 Pro Huawei (screen size: 6.39 inches, resolution: 1440 x 3120 pixels). To avoid recognizing the smartphone brand, all signs were covered or removed in advance. The mobile phone was placed in a stand on a desk in front of the participant with a maximum distance of 30 cm. A web-camera (Logitech HD Pro Webcam C920) was positioned on top to extract the facial muscle activity (i.e., action units; AUs; [17]) frame by frame via the iMotions Biometric Research Platform [33] software. Frames were sampled at 102.4 Hz and aligned as a pre-processing step within iMotions. For the facial expression recognition, we explored the AFFDEX algorithm classifying seven basic emotions [70]. We recorded scalp EEG potentials according to the international 10-20 system with 32 electrodes (actiCAP, Brainproducts GmbH, Germany). The left mastoid was used as common reference and EEG was grounded to FCz. Impedance of electrodes was kept below 20 kΩ at the onset of each session. EEG data was digitized at 1 kHz, high-pass filtered with a time constant of 10 s, and stored for off-line analysis using the BrainVision Recorder Software (BrainProducts GmbH, Germany). ECG was recorded according to the Einthoven technique with electrodes placed on the left clavicle, at the sternum and the ground electrode at the left elbow (BrainAmp, Brainproducts GmbH, Germany). The data was digitized at 1 kHz using the BrainVision Recorder Software. EDA was recorded using the wearable Shimmer3 GSR + Unit and iMotions software with a digitization at 128 Hz. We mounted electrodes on the fingertips of the left index finger and middle finger. Figure 1 provides an overview of the experimental set-up and illustrative sensor placement. The unified collection of multimodal signals from the different recording and presentation systems were synchronized and stored for off-line data analysis using Lab Streaming Layer (LSL), a web-based JavaScript application (stimulus presentation) in combination with TCP protocols, a wireless access-point, and the iMotions platform. Subjective measurements were collected via LimeSurvey [71].

2.3 Stimuli and design

We designed specific user interface mock-ups for three mobile phone scenarios: (1) onboarding (i.e., set-up process at first use), (2) general settings (changing Wifi settings, background, and unlocking method), and (3) calendar (making an entry and sharing it). The interaction scenarios were developed based on the criteria of the JoU model (aesthetics, Wow, needs) and selected to represent everyday common use, ensuring familiarity, and intuitive handling. Each scenario was designed for four conditions (A) low JoU – high U, (B) high JoU – high U, (C) low JoU – low

Table 1: Overview of the number of stimuli in each scenario-condition combination.

	Calender	General Settings	Onboarding	N
low JoU - high U	17	47	50	114
high JoU – high U	17	39	26	82
low JoU – low U	15	43	47	105
high JoU – low U	12	30	19	61

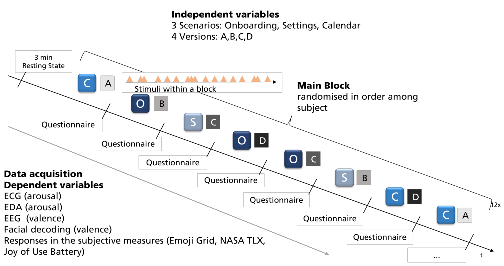

Figure 3: Overview of the experimental procedure with the four conditions (A) low JoU – high U, (B) high JoU – high U, (C) low JoU – low U, and (D) high JoU – low U and three interaction scenarios onboarding (O), general settings (S), and calendar (C).

U, and (D) high JoU – low U (see Figure 2). The 12 interactions (three scenarios, four conditions) were prototyped in B.V. [8] and exported as html-files via Anima [3]. The html-files as well as illustrative videos of the scenarios are online accessible under https://osf.io/g865s/?view_only=c99fe513ba934aca8d072e836754144a. Interactions lasted between 21 s to 12 min and 26 s. 1 provides an overview of the number of stimuli for each scenario-condition combination. We randomized the order across participants.

2.4 Procedure

Prior to the experiment, participants were informed about muscle and movement artefacts in all recordings and instructed to keep movements to a minimum. They were comfortably seated in a quiet room. An overview of the experimental procedure is provided in Figure 3. The experiment started with a three-minute EEG, ECG and EDA baseline recording during relaxation with eyes fixating a crosshair on a computer screen in front of the participant. All participants explored each scenario-condition combination, resulting in 12 blocks. They were instructed to explore the interactive scenarios and complete small pre-defined tasks (e.g., setting an appointment in the calendar).

After each block, a composite JoU questionnaire was provided to analyse subjective experiences of the interaction. Subscales comprised the following constructs: (1) Arousal and valence were examined on a continuous visual scale ranging from 0 to 9 via the Emoji Grid [72]. (2) The Nasa Task Load Index (TLX; [32]) examines usability factors on an adapted sliding scale ranging from 0 to 100. (3) The Extended User Experience Questionnaire (UEQ+; [65]) was used to assess visual aesthetics, as well as Wow, the latter via the subscale originality and stimulation, on a 7-point Likert scale. Psychological needs were measured on a 5-point Likert scale using the meCUE luxury subscale [48], and an adapted version of the need-based questionnaire for the items examining competence, relatedness, and autonomy [67]. In addition, we asked participants to rate their general satisfaction on a scale ranging from 1 to 6 and provide positive and negative feedback within two open-format items.

2.5 Analysis

All data-analyses were performed with custom written scripts in pythonTM, JASP 0.13.1, and IBM SPSS® Statistics 20. Means and standard deviations of the measures per conditions are provided in Table 1 and 2 in the Supplementary Material.

2.5.1 *Subjective measures.*

We used repeated measure analyses of variance (rmANOVA) with condition as a within-subject factor (A, B, C, D). If sphericity was violated, Greenhouse-Geisser correction was chosen. To examine the differences within conditions, we computed contrasts using t-test and a Bonferroni multiple-comparison correction (Table 3). Furthermore, effect sizes (partial eta squared) of the main effects are provided for parametric analyses with $\eta^2 p$ of .01 indicating a small, $\eta^2 p$ of .06 indicating a medium, and $\eta^2 p$ of .14 indicating a large effect [12, 61].

2.5.2 *Objective measures.*

In the EEG pre-processing, we de-trended, zero-padded, and re-referenced the data to a common average reference (CAR). We excluded seven EEG channels (FT9, FT10, Oz, TP9, TP10, T8, and T7) due to artefact contamination. Data was band-pass filtered between 1 to 40 Hz (first order zero-phase lag FIR filter). Afterwards, epochs of the length of 4 s were extracted starting at the stimulus-onset and grouped into the four experimental conditions. We rejected epochs containing a maximum deviation above 180 μV in any of the frontal and occipital EEG channels (Fp1, Fp2, F7, F8, O1, O2). To remove further cardiac-related and muscular artefacts as well as ocular movement, we performed an independent component analysis (ICA) on the epoched data using the extended infomax ICA algorithm [40] as implemented in the MNE-Python toolbox [21]. Artefact-contaminated components were manually selected after visual inspection of the topography, times course, and power spectral intensity [10, 31]. After pre-processing, we calculated EEG power spectra using a modified version of the Fast Fourier Transformation (FFT), the so-called Welch's method, for the individual electrodes in the frequency range 1 to 35 Hz as implemented in the MNE-Python toolbox [21]. Resting state data used for normalizing power spectra in each frequency band was pre-processed as described above using non-overlapping epochs of 2 s. From the calculated power spectra, we derived indices specifically suitable to examine user experience and usability comprising the (1) mental workload index (WL), (2) visuo-motoric engagement index (EI), and (3) frontal alpha asymmetry index (FAA). Frequencies of interest were the theta (4 – 7 Hz), alpha (8 – 13 Hz), and beta band (15 – 25 Hz). Before the EEG indices were calculated, data was log-transformed to dB scale and normalized to the mean and standard deviation of the resting state.
We calculated the WL-index [7, 20, 35, 54] by dividing the average theta band power in Fz by the average power in the alpha band of

Pz:

$$WL = \frac{Frontal\ Theta\ Band}{Parietal\ Alpha\ Band} \quad (1)$$

The EI-index associated with visuo-motoric information processing and coordination was calculated according to Pope et al. [58] and Dehais et al. [13]:

$$EI = \frac{Beta\ Band}{Theta\ Band\ +\ Alpha\ Band} \quad (2)$$

Hence, we divided the average beta by the sum of the average theta and alpha band power over six EEG central and parietal electrodes (C3, Cz, C4, P3, Pz, and P4). For the calculation of the FAA, we subtracted the average alpha band power in F3 (left hemisphere) from F4 (right hemisphere):

$$FAA = Right\ Alpha\ Band - Left\ Alpha\ Band \quad (3)$$

For the ECG pre-processing, we processed the data to acquire the inter-beat interval (IBI) which is the inverse of the heart rate (HR). The ECG signal was band-pass filtered between 5 and 15 Hz using a third order Butterworth filter and artefact-corrected according to Lipponen and Tarvainen [41]. Peaks were detected following Pan and Tompkins [53]. We transformed the IBI semi-time series into a timeseries by interpolating consecutive IBIs (quadratic spline interpolation), resampled at 512 Hz, and transformed the IBI to HR. ECG and EDA data analysis was performed with the neurokit2 toolbox in pythonTM [44]. The EDA signal was low-passed filtered with a cut-off frequency of 1 Hz followed by a moving average smoothing using a linear convolution with a filter kernel size of 0.75 * sampling rate and a boxzen window. Phasic and tonic components were extracted using a convex optimization model with the cvxEDA algorithm. For the statistical comparison of the ECG and EDA, we cut event-locked epochs of the length of 7 s after stimulus-onset. As a baseline correction, we subtracted the average signal of a baseline interval ranging from 400 ms prior to stimulus-onset. The data was normalized by z-scoring signals with the mean and standard deviation of the resting state. For the ECG statistical analysis, we extracted the mean, minimum, maximum value of the HR of each epoch separately for the four experimental conditions. For the EDA analysis, we extracted the mean, maximum value, peak onset and amplitude of the phasic response including and excluding the tonic component of the skin conductance responses (SCR) from the phasic component. For the facial expression recognition, we analyzed values of the seven basic emotions (1) joy, (2) fear, (3) disgust, (4) sadness, (5) anger, (6) surprise, and (7) contempt. We only used frames with a pitch value between $-10°$ to $+20°$ as well as roll and yaw values between $-10°$ and $+10°$, since values can be reliably estimated within this frame interval. A baseline of 200 ms was extracted before each stimulus-onset and subtracted from the raw signal to account for inter-subject-variability. We cut epochs of the length of 4 s after stimulus-onset and normalized the data by z-scoring it with the mean and standard deviation over all epochs. For the statistical analysis, we extracted the median of a non-overlapping sliding window of 500 ms for each epoch and averaged it over epochs for each condition, separately. For the comparison of conditions in the objective measures, we used non-parametric Friedman's Q tests to account for non-normally distributed data and outliers. To examine exact differences, nonparametric Wilcoxon signed ranks tests were conducted and corrected for multiple comparisons using the false discovery rate (FDR) with the Benjamini-Hochberg method.

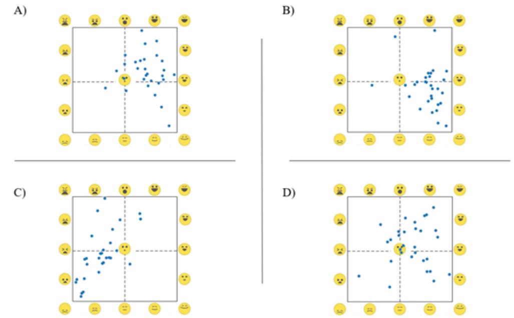

Figure 4: Ratings in the Emoji Grid for the four conditions (A) low JoU – high U, (B) high JoU – high U, (C) low JoU – low U, and (D) high JoU – low U. The x-axis represents the valence dimension, while the y-axis represents the arousal dimension.

3 RESULTS

3.1 Subjective measures

3.1.1 Arousal and Valence. Figure 4 shows ratings in the Emoji Grid examining subjectively perceived arousal and valence. Since the data was non-normally distributed a non-parametric statistical analysis was chosen. It revealed a significant difference between conditions (Table 2) for arousal and valence. B was rated significantly more arousing than A and D (Table 3). For the valence ratings, C was rated significantly more negative compared to A, B, and D; and B significantly more positive than D. Wow. Wow revealed a significant main effect for condition with B triggering Wow on a high level. There were significant differences between A compared to B and D as well as C compared to B and D.

3.1.2 Aesthetic pleasure. Furthermore, conditions differed in their aesthetic pleasure with higher ratings in B compared to A, C and D.

3.1.3 Individual needs. Regarding the experienced stimulation and luxury, the conditions were rated differently with B triggering stimulation on a significantly higher level compared to A, C, and D and was rated as more luxurious compared to A and C. However, there was no difference between B and D for the latter, indicating that luxury can be evoked even when bad usability is experienced. Regarding competence, we found no difference between the conditions. Furthermore, means were rather low indicating that competence could not be evoked successfully in all conditions. However, for relatedness, we observed a significant main effect for the conditions with significantly higher ratings in B compared to A and C. There was no significant difference in D indicating that relatedness could be evoked even under usability problems. Ratings of autonomy showed a significant main effect for condition with B differed to all other conditions.

3.1.4 Usability. The conditions differed significantly in effort with significantly lower ratings in B compared to C and D. However, there was no difference to A indicating that usability plays an

Table 2: Results of the repeated measure analyses of variance (rmANOVA) or non-parametric Friedman's Q tests with condition as within-subject factor (A, B, C, D).

Measure	*Model*	F / χ^2	p	$\eta^2 p$	N
Arousal	$\chi^2(3)$	12.476	.006**	-	29
Valence	$\chi^2(3)$	55.80	≤ .001**	-	29
Needs					
Wow	$F(2.12, 61.51)$	73.81	≤ .001**	.718	30
Aesthetic pleasure	$F(2.24, 64.82)$	82.49	≤ .001**	.740	30
Individual Needs					
Stimulation	$F(2.00, 58.12)$	39.74	≤ .001**	.578	30
Luxury	$F(2.44, 70.83)$	16.37	≤ .001**	.361	30
Competence	$F(2.45, 24.35)$	1.906	.152	.361	30
Relatedness	$F(2.42, 70.39)$	8.87	≤ .001**	.234	30
Autonomy	$F(2.20, 63.83)$	17.40	≤ .001**	.375	30
NASA Taskload					
Effort	$F(2.09, 60.46)$	59.69	≤ .001**	.673	30
Temporal Demand	$F(2.07, 60.14)$	51.79	≤ .001**	.641	30
Frustration	$F(2.82, 81.68)$	81.30	≤ .001**	.564	30
Fun	$F(2.11, 61.10)$	29.81	≤ .001**	.507	30

important role in perceived effort. Regarding the perceived temporal demand, there was a main effect for condition with A evoking less temporal demand compared to B, C, and D. The frustration scale revealed a significant main effect for condition with A and B evoking equally low frustration. A and B differ significantly to D and C. Regarding the experienced fun, participants rated the conditions significantly different with higher ratings in B compared to A, D, and C.

3.2 Objective measures

3.2.1 EEG results. The analysis reveals no difference between the conditions for the EEG indices EI, WL, and FAA (Table 4).

3.2.2 ECG results. Our results reveal no differences regarding the minimum and maximum HR. However, the mean HR differed significantly among the conditions with exceptionally high HR responses in A, indicating low JoU and high U, compared to all other conditions (Table 5).

3.2.3 EDA results. The EDA analysis reveals no differences in the mean phasic response; but for the maximum phasic response was significantly higher in A compared to D. Regarding the mean amplitude of the phasic response, a significant difference could be observed. However, the post-hoc comparisons were not significant after the FDR correction. The conditions did not differ in further EDA measures.

Table 3: Results of the post-hoc comparisons between the conditions for the subjective measures using t-test with a Bonferroni multiple-comparison correction. Post-hoc comparisons were only calculated in case of a significant rmANOVA or Friedman's Q tests. For the Emoji Grid the non-parametric Wilcoxon signed ranks test (Z) was chosen.

Measure	*Contrast*	*test*	p	N
Emoji				
Arousal	A – B	Z	.003**	29
	B – D	Z	.002**	29
Valence	A – C	Z	≤ .001**	29
	B – C	Z	.009**	29
	C – D	Z	≤ .001**	29
	B – D	Z	.002**	29
Needs				
Wow	A – B	t	≤ .001**	30
	A – D	t	≤ .001**	30
	B – C	t	≤ .001**	30
	C – D	t	≤ .001**	30
Aesthetic pleasure	A – B	t	≤ .001**	30
	B – C	t	≤ .001**	30
	B – D	t	.002**	30
Individual Needs				
Stimulation	A – B	t	≤ .001**	30
	B – C	t	≤ .001**	30
	B – D	t	≤ .001**	30
Luxury	A – B	t	.018*	30
	B – C	t	≤ .001**	30
Relatedness	A – B	t	.002**	30
	B – C	t	≤ .001**	30
Autonomy	A – B	t	.003**	30
	B – C	t	.001**	30
	B – D	t	.004**	30
NASA Taskload				
Effort	B – C	t	≤ .001**	30
	B – D	t	.015*	30
Temporal demand	A – B	t	≤ .001**	30
	A – C	t	≤ .001**	30
	A – D	t	≤ .001**	30
Frustration	A – C	t	≤ .001**	30
	A – D	t	≤ .001**	30
	B – C	t	≤ .001**	30
	B – D	t	≤ .001**	30
Fun	A – B	t	≤ .001**	30
	B – C	t	≤ .001**	30
	B – D	t	≤ .001**	30

3.2.4 Facial emotion recognition results. In the facial emotion recognition analysis, we did not observe differences between the conditions for the emotions contempt, disgust, fear, sadness, and surprise (Table 4). Surprisingly, there was a significant difference for anger with higher activity values in A compared to C. In addition, we

Table 4: Results of the non-parametric Friedman's Q tests of the objective measures with condition as factor (A, B, C, D).

Measure	*Model*	χ^2	p	N
EEG				
EI	3	0.831	.842	26
WL	3	1.523	.678	26
FAA	3	0.323	.956	26
ECG				
Mean HR	3	17.080	$\leq .001^{**}$	30
Minimum HR	3	2.440	.486	30
Maximum HR	3	6.520	.089	30
EDA				
Mean Phasic Response	3	4.360	.225	30
Max Phasic Response	3	10.600	$.014^{*}$	30
Mean Amplitude	3	8.00	$.046^{*}$	30
Max Amplitude	3	6.880	.076	30
Mean Amplitude incl. tonic component	3	5.320	.150	30
Max Amplitude incl. tonic component	3	4.880	.181	30
Facial emotion recognition				
Contempt	3	3.000	.392	25
Disgust	3	3.096	.377	25
Fear	3	1.56	.668	25
Sadness	3	3.528	.317	25
Surprise	3	0.840	.840	25
Anger	3	7.896	$.048^{*}$	25
Joy	3	7.704	.053	25

Table 5: Results of the post-hoc comparisons between the conditions for the objective measures using non-parametric Wilcoxon signed ranks tests and a Bonferroni-Hochberg FDR multiple-comparison correction.

Measure	*Contrast*	p	N
ECG - Mean HR	A - B	$.009^{**}$	30
	A - C	$.009^{**}$	30
	A - D	$.015^{*}$	30
EDA - Max Phasic Response	A - D	$.030^{*}$	30
Facial emotion recognition Anger	A - C	$.015^{*}$	25

found a slight non-significant trend for joy with higher activity in the respective muscles in C but a large standard deviation signalling to strong outliers.

4 DISCUSSION

4.1 Subjective measures

Our results of the subjective measures showed that participants had greater positive experiences in interactions designed to elicit JoU even under bad usability. These findings challenge the theory that usability is rather a necessary than sufficient hygiene factor for positive user experience [27, 30]. It seems that JoU compensates usability problems in some cases. However, usability plays an important role for specific individual needs such as autonomy and experienced effort to ensure control, execution of the intentions, and predictability of interaction patterns. When perceiving usability problems, participants were frustrated and experienced less fun due to obstacles to the task achievement. Interestingly, JoU alone could not compensate usability problems regarding the experienced fun. We observed that especially the combination of JoU and good usability elicits high arousal. But there was no difference to the condition with low JoU and low usability, indicating that arousal is less suited to distinguish between positive and negative emotional events.

4.2 Objective measures

We did not find differences between conditions in the EEG indices. This might be because the stimuli were not suitable to induce emotional experience in these measures, especially for workload and engagement since stimuli aimed to induce rather affective responses. Moreover, our approach of using large time windows of 4 sec epochs after stimulus onset may have been not sensitive enough to capture the direct neuronal activation pattern associated with experiencing JoU and usability. More sophisticated analysis like time-frequency and spatially resolved oscillatory EEG measures, such as event-related spectral perturbation and functional connectivity analysis, might reveal higher sensitivity and distinctiveness for detecting affective experiences during the interaction with more naturalistic stimuli [66, 76]. In the physiological measures, we observe that the average HR as well as maximum and mean amplitude of the phasic response in the EDA were elevated for the usability condition without JoU. Higher HR activity and phasic response in the EDA in the high usability condition could be related to motivational components due to an easy task achievement compared to a more relaxed, joyful, and less ambitious state in the other conditions [78]. Previous research has revealed that experiencing negative emotions is associated with increased HR, whereas positive emotions display rather ambivalent HR responses [17, 39]. Still, more research needs to be conducted to disentangle whether the increased HR and phasic EDA response during good usability is explained by motivational aspects or rather subtle negative emotions and stressful experiences. Contrary to our hypotheses, we observed flipped effects in the facial expression analysis with higher muscle activation associated with anger in the good usability condition without JoU and higher activation associated with joy during bad usability without JoU. We suspect the latter phenomenon to represent an exaggerated response in form of a disdainful smile than a signal of anger (e.g., a frown). In general, usage of facial expressions as indicator for experienced emotions is under debate as high context sensitivity and inter-individual variability make a valid and robust classification more difficult [4, 29, 38]. Nevertheless, the availability of camera sensors in mobile phones as well as future improvements in classification algorithms give reason to further explore the potential of the measurement method including other classification algorithms. For instance, a recent validation study comparing the

AFFDEX with the FACET algorithm revealed higher accuracy of emotion recognition for the latter; in particular when classifying anger and surprise [69]. Concluding, our results of the objective measures suggest that experiencing usability without JoU seem to elicit rather high physiological responses, arousal, and tension than mere joy and relaxation.

4.3 Outlook and Limitations

Future studies should continue to investigate the specific relation and interdependence between usability and JoU. The interaction is still not clear and requires replicational studies - ideally with a mulitmodal study design. Regarding the design of future technology, we implicate to concentrate more on the hedonic aspects rather than only on usability. However, it is still to be defined what aspects of JoU are the most important ones (the fulfilment of needs, the Wow aspects or the aesthetic pleasure). Limitations of the study comprise the less-standardized stimulus material and missing information about personality traits (e.g., emotional processing). We are planning to behaviorally validate the stimulus material regarding the effectiveness of the manipulation (i.e., high vs. low usability and JoU) by using videos of the interactions with the scenarios and an independent sample. We suggest an approach using time-point-to-time-point continuous ratings during the presentation of the videos that separately investigate usability and the JoU components. Further, a rather homogeneous sample participated in the here presented study. Participants were predominantly young, German, students, and well educated. Future studies should include more diverse groups (older age, non-students, and other cultures).

5 CONCLUSION

Our multimodal methods approach provides a holistic picture on user experience in a highly ecologically valid setting. It allows to detect subtle differences and, thus, grasp the relationship and interdependence between JoU and usability. Our findings contribute to the ongoing debate about the importance of hedonic aspects to overall user experience. Additionally, we laid the groundwork for systematically and effectively capturing (neuro-)physiological and behavioural indices related to the different elements determining user experience, most importantly joy and positive emotions. This groundwork can not only be used in a highly standardized lab setting but is viable for being implemented with naturalistic interactions with real consumer devices. The study contributes to the evaluation and improvement of future technology, placing users at the centre of designing human-machine interactions.

ACKNOWLEDGMENTS

The authors would like to thank Michael Bui for his support in the technical set-up, Vanessa Budde for her valuable contribution to the experimental design as well as Sabrina Gado for the constructive criticism of the manuscript.

REFERENCES

[1] ISO 9241-11: 2018. 2018. Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts.

[2] Geoffrey L Ahern and Gary E Schwartz. 1985. Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. *Neuropsychologia* 23, 6 (1985), 745–755.

[3] Anima. 2018. *Anima (Version 68).* Anima. https://www.animaapp.com

[4] David Bethge, Thomas Kosch, Tobias Grosse-Puppendahl, Lewis L Chuang, Mohamed Kari, Alexander Jagaciak, and Albrecht Schmidt. 2021. VEmotion: Using Driving Context for Indirect Emotion Prediction in Real-Time. In *The 34th Annual ACM Symposium on User Interface Software and Technology.* Association for Computing Machinery, New York, NY, United States, 638–651.

[5] Upasna Bhandari, Klarissa Chang, and Tillmann Neben. 2019. Understanding the impact of perceived visual aesthetics on user evaluations: An emotional perspective. *Information & management* 56, 1 (2019), 85–93.

[6] Benny B Briesemeister, Sascha Tamm, Angela Heine, Arthur M Jacobs, et al. 2013. Approach the good, withdraw from the bad—a review on frontal alpha asymmetry measures in applied psychological research. *Psychology* 4, 03 (2013), 261.

[7] Anne-Marie Brouwer, Maarten A Hogervorst, Jan BF Van Erp, Tobias Heffelaar, Patrick H Zimmerman, and Robert Oostenveld. 2012. Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of neural engineering* 9, 4 (2012), 045008.

[8] Sketch B.V. 2018. *Sketch (Version 68).* Sketch B.V. https://www.sketch.com

[9] Jing Chai, Yan Ge, Yanfang Liu, Wen Li, Lei Zhou, Lin Yao, and Xianghong Sun. 2014. Application of frontal EEG asymmetry to user experience research. In *International conference on engineering psychology and cognitive ergonomics.* Springer, Springer, Berlin/Heidelberg, Germany, 234–243.

[10] Maximilien Chaumon, Dorothy VM Bishop, and Niko A Busch. 2015. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of neuroscience methods* 250 (2015), 47–63.

[11] Junho H Choi and Hye-Jin Lee. 2012. Facets of simplicity for the smartphone interface: A structural model. *International Journal of Human-Computer Studies* 70, 2 (2012), 129–142.

[12] J Cohen. 1969. *Statistical Power Analysis for the Behavioral Sciences.* Academic Press, New York, NY.

[13] Frédéric Dehais, Raphaëlle N Roy, Gautier Durantin, Thibault Gateau, and Daniel Callan. 2017. EEG-engagement index and auditory alarm misperception: an inattentional deafness study in actual flight condition. In *International Conference on Applied Human Factors and Ergonomics.* Springer, Springer, Berlin/Heidelberg, Germany, 227–234.

[14] Pieter Desmet and Steven Fokkinga. 2020. Beyond Maslow's pyramid: introducing a typology of thirteen fundamental needs for human-centered design. *Multimodal technologies and interaction* 4, 3 (2020), 38.

[15] PMA Desmet, Rick Porcelijn, and MB Van Dijk. 2005. HOW to design WOW? Introducing a layered-emotional approach. In *Proceedings of The International Conference on Designing Pleasurable Products and Interfaces.* Association for Computing Machinery, New York, NY, United States, 24–27.

[16] Paul Ekman. 1992. An argument for basic emotions. *Cognition & emotion* 6, 3-4 (1992), 169–200.

[17] Paul Ekman, Robert W Levenson, and Wallace V Friesen. 1983. Autonomic nervous system activity distinguishes among emotions. *science* 221, 4616 (1983), 1208–1210.

[18] Nora Fronemann and Matthias Peissner. 2014. User experience concept exploration: user needs as a source for innovation. In *Proceedings of the 8th nordic conference on human-computer interaction: Fun, fast, foundational.* Association for Computing Machinery, New York, NY, United States, 727–736.

[19] Yan Ge, Yanan Chen, Yanfang Liu, Wen Li, and Xianghong Sun. 2014. Electrophysiological measures applied in user experience studies. *Advances in Psychological Science* 22, 6 (2014), 959.

[20] Alan Gevins, Michael E Smith, Linda McEvoy, and Daphne Yu. 1997. High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral cortex (New York, NY: 1991)* 7, 4 (1997), 374–385.

[21] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A Engemann, Daniel Strohmeier, Christian Brodbeck, Lauri Parkkonen, and Matti S Hämäläinen. 2014. MNE software for processing MEG and EEG data. *Neuroimage* 86 (2014), 446–460.

[22] Alice Gross and Johanna Silvennoinen. 2014. Surprise as a design strategy in goal-oriented mobile applications. In *Advances in Human Factors and Ergonomics 2014.* The Applied Human Factors and Ergonomics Conference, CRC Press, Boca Raton, Florida, FL, 4716–4726.

[23] Alice Gross and Manfred Thüring. 2013. Encountering the Unexpected: Influencing User Experience through Surprise.. In *UMAP Workshops.* Springer, Berlin/Heidelberg, Germany, 1–7.

[24] Marc Hassenzahl. 2007. The hedonic/pragmatic model of user experience. *Towards a UX manifesto* 10 (2007), 10–14.

[25] Marc Hassenzahl. 2018. A personal journey through user experience. *Journal of Usability Studies* 13, 4 (2018), 168–176.

[26] Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality. In *Mensch & Computer*, Vol. 57. Springer, Springer, Berlin/Heidelberg, Germany, 187–196.

[27] Marc Hassenzahl, Sarah Diefenbach, and Anja Göritz. 2010. Needs, affect, and interactive products–Facets of user experience. *Interacting with computers* 22, 5 (2010), 353–362.

[28] Marc Hassenzahl, Annika Wiklund-Engblom, Anette Bengs, Susanne Hägglund, and Sarah Diefenbach. 2015. Experience-oriented and product-oriented evaluation: psychological need fulfillment, positive affect, and product perception. *International journal of human-computer interaction* 31, 8 (2015), 530–544.
[29] Douglas Heaven. 2020. Why faces don't always tell the truth about feelings. *Nature* 578, 7796 (2020), 502–505.
[30] Frederick Herzberg, Bernard Mausner, and Barbara B Snyderman. 1959. *The Motivation to Work*. Vol. 195. John Wiley & Sons Inc., New York, NY, United States.
[31] Joerg F Hipp and Markus Siegel. 2013. Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. *Frontiers in human neuroscience* -, - (2013), 338.
[32] Peter Hoonakker, Pascale Carayon, Ayse P Gurses, Roger Brown, Adjhaporn Khunlertkit, Kerry McGuire, and James M Walker. 2011. Measuring workload of ICU nurses with a questionnaire survey: the NASA Task Load Index (TLX). *IIE transactions on healthcare systems engineering* 1, 2 (2011), 131–143.
[33] iMotions Biometric Research Platform. 2018. *iMotions A/S (Version 8.1)*. iMotions Biometric Research Platform. https://imotions.com
[34] Patrick W Jordan. 2000. *Designing pleasurable products: An introduction to the new human factors*. CRC press, Boca Raton, Florida, FL.
[35] Altyngul T Kamzanova, Almira M Kustubayeva, and Gerald Matthews. 2014. Use of EEG workload indices for diagnostic monitoring of vigilance decrement. *Human factors* 56, 6 (2014), 1136–1149.
[36] Min-Ki Kim, Miyoung Kim, Eunmi Oh, and Sung-Phil Kim. 2013. A review on the computational methods for emotional state estimation from the human EEG. *Computational and mathematical methods in medicine* 2013 (2013), 1–13.
[37] Sang Hee Kim and Stephan Hamann. 2007. Neural correlates of positive and negative emotion regulation. *Journal of cognitive neuroscience* 19, 5 (2007), 776–798.
[38] Thomas Kosch, Mariam Hassib, Robin Reutter, and Florian Alt. 2020. Emotions on the go: Mobile emotion assessment in real-time using facial expressions. In *Proceedings of the International Conference on Advanced Visual Interfaces*. Association for Computing Machinery, New York, NY, United States, 1–9.
[39] Sylvia D Kreibig. 2010. Autonomic nervous system activity in emotion: A review. *Biological psychology* 84, 3 (2010), 394–421.
[40] Te-Won Lee, Mark Girolami, and Terrence J Sejnowski. 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural computation* 11, 2 (1999), 417–441.
[41] Jukka A Lipponen and Mika P Tarvainen. 2019. A robust algorithm for heart rate variability time series artefact correction using novel beat classification. *Journal of medical engineering & technology* 43, 3 (2019), 173–181.
[42] Sascha Mahlke and Gitte Lindgaard. 2007. Emotional experiences and quality perceptions of interactive products. In *International Conference on Human-Computer Interaction*. Springer, Springer, Berlin/Heidelberg, Germany, 164–173.
[43] Marco Maier, Daniel Elsner, Chadly Marouane, Meike Zehnle, and Christoph Fuchs. 2019. DeepFlow: Detecting Optimal User Experience From Physiological Data Using Deep Neural Networks. In *AAMAS*. International Joint Conferences on Artificial Intelligence, California, CA, United States, 2108–2110.
[44] Dominique Makowski, Tam Pham, Zen J Lau, Jan C Brammer, François Lespinasse, Hung Pham, Christopher Schölzel, and SH Chen. 2021. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior research methods* 53, 4 (2021), 1689–1696.
[45] Slobodan Marković. 2012. Components of aesthetic experience: aesthetic fascination, aesthetic appraisal, and aesthetic emotion. *i-Perception* 3, 1 (2012), 1–17.
[46] Deana McDonagh, Anne Bruseberg, and Cheryl Haslam. 2002. Visual product evaluation: exploring users' emotional relationships with products. *Applied Ergonomics* 33, 3 (2002), 231–240.
[47] Daniel McDuff, Abdelrahman Mahmoud, Mohammad Mavadati, May Amr, Jay Turcot, and Rana el Kaliouby. 2016. AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. Association for Computing Machinery, New York, NY, United States, 3723–3726.
[48] Michael Minge, Laura Riedel, and Manfred Thüring. 2014. Modulare Evaluation interaktiver Technik. Entwicklung und Validierung des meCUE Fragebogens zur Messung der User Experience. *Grundlagen und Anwendungen der Mensch-Technik-Interaktion* 10 (2014), 28–36.
[49] Michael Minge and Manfred Thüring. 2018. Hedonic and pragmatic halo effects at early stages of user experience. *International Journal of Human-Computer Studies* 109 (2018), 13–25.
[50] John Mingers. 2001. Combining IS research methods: towards a pluralist methodology. *Information systems research* 12, 3 (2001), 240–259.
[51] Richard E Nisbett and Timothy D Wilson. 1977. Telling more than we can know: Verbal reports on mental processes. *Psychological review* 84, 3 (1977), 231.
[52] Sebastian Ocklenburg, Patrick Friedrich, Judith Schmitz, Caroline Schlüter, Erhan Genc, Onur Güntürkün, Jutta Peterburs, and Gina Grimshaw. 2019. Beyond frontal alpha: investigating hemispheric asymmetries over the EEG frequency spectrum as a function of sex and handedness. *Laterality: Asymmetries of Body, Brain and Cognition* 24, 5 (2019), 505–524.
[53] Jiapu Pan and Willis J Tompkins. 1985. A real-time QRS detection algorithm. *IEEE transactions on biomedical engineering* 32, 3 (1985), 230–236.
[54] Mirka Pesonen, Heikki Hämäläinen, and Christina M Krause. 2007. Brain oscillatory 4–30 Hz responses during a visual n-back memory task with varying memory load. *Brain research* 1138 (2007), 171–177.
[55] Rosalind W Picard. 2011. Measuring affect in the wild. In *International conference on affective computing and intelligent interaction*. Springer, Springer, Berlin/Heidelberg, Germany, 3–3.
[56] Anna E Pohlmeyer, Martin Hecht, and Lucienne Blessing. 2009. User experience lifecycle model ContinUE [continuous user experience]. *Der Mensch im Mittepunkt technischer Systeme. Fortschritt-Berichte VDI Reihe* 22 (2009), 314–317.
[57] Kathrin Pollmann, Oilver Stefani, Amelie Bengsch, Matthias Peissner, and Mathias Vukelić. 2019. How to work in the car of the future? A neuroergonomical study assessing concentration, performance and workload based on subjective, behavioral and neurophysiological insights. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, United States, 1–14.
[58] Alan T Pope, Edward H Bogart, and Debbie S Bartolome. 1995. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology* 40, 1-2 (1995), 187–195.
[59] Jonathan Posner, James A Russell, and Bradley S Peterson. 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology* 17, 3 (2005), 715–734.
[60] Steven Reiss. 2004. Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. *Review of general psychology* 8, 3 (2004), 179–193.
[61] John TE Richardson. 2011. Eta squared and partial eta squared as measures of effect size in educational research. *Educational research review* 6, 2 (2011), 135–147.
[62] Erika L Rosenberg and Paul Ekman. 2020. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, Oxford, United Kingdom.
[63] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.
[64] Shilpa S Samant, Matthew J Chapko, and Han-Seok Seo. 2017. Predicting consumer liking and preference based on emotional responses and sensory perception: A study with basic taste solutions. *Food research international* 100 (2017), 325–334.
[65] Martin Schrepp and Jörg Thomaschewski. 2019. Eine modulare Erweiterung des User Experience Questionnaire. In *Mensch und Computer 2019 - Usability Professionals*, Holger Fischer and Steffen Hess (Eds.). Gesellschaft für Informatik e.V. Und German UPA e.V., Bonn, 148–156. https://doi.org/10.18420/muc2019-up-0108
[66] Dennis JLG Schutter and Gennady G Knyazev. 2012. Cross-frequency coupling of brain oscillations in studying motivation and emotion. *Motivation and emotion* 36, 1 (2012), 46–54.
[67] Kennon M Sheldon, Andrew J Elliot, Youngmee Kim, and Tim Kasser. 2001. What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of personality and social psychology* 80, 2 (2001), 325.
[68] Andreas Sonnleitner, Marvin Pawlowski, Timm Kässer, and Matthias Peissner. 2013. Experimentally manipulating positive user experience based on the fulfilment of user needs. In *IFIP Conference on Human-Computer Interaction*. Springer, Springer, Berlin/Heidelberg, Germany, 555–562.
[69] S Stöckli, M Schulte-Mecklenbeck, S Borer, and AC Samson. 2017. Validation of iMotions emotional facial expression analysis with the AFFDEX and FACET module. *Behavior Research Methods* 50, 4 (2017), 1446–1460.
[70] Sabrina Stöckli, Michael Schulte-Mecklenbeck, Stefan Borer, and Andrea C Samson. 2018. Facial expression analysis with AFFDEX and FACET: A validation study. *Behavior research methods* 50, 4 (2018), 1446–1460.
[71] LimeSurvey Project Team and Carsten Schmitz. 2012. LimeSurvey: An open source survey tool. *LimeSurvey Project Hamburg, Germany* 1, 1 (2012), 1. http://www.limesurvey.org
[72] Alexander Toet and Jan BF van Erp. 2019. The EmojiGrid as a tool to assess experienced and perceived emotions. *Psych* 1, 1 (2019), 469–481.
[73] Andrew J Tomarken, Richard J Davidson, and Jeffrey B Henriques. 1990. Resting frontal brain asymmetry predicts affective responses to films. *Journal of personality and social psychology* 59, 4 (1990), 791.
[74] Kaisa Väänänen-Vainio-Mattila, Jarmo Palviainen, Santtu Pakarinen, Else Lagerstam, and Eeva Kangas. 2011. User perceptions of Wow experiences and design implications for Cloud services. In *Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces*. Association for Computing Machinery, New York, NY, United States, 1–8.
[75] Arnold POS Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. 2010. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries*. Association for Computing Machinery, New York, NY, United States, 521–530.

[76] Mathias Vukelić, Katharina Lingelbach, Kathrin Pollmann, and Matthias Peissner. 2020. Oscillatory EEG Signatures of Affective Processes during Interaction with Adaptive Computer Systems. *Brain Sciences* 11, 1 (2020), 35.
[77] Gillian M Wilson and M Angela Sasse. 2004. From doing to being: getting closer to the user experience. *Interacting with Computers* 16, 4 (2004), 697–705.
[78] Yan Wu, Ruolei Gu, Qiwei Yang, and Yue-jia Luo. 2019. How do amusement, anger and fear influence heart rate and heart rate variability? *Frontiers in neuroscience* 1, 1 (2019), 1131.
[79] Haotian Zhou and Xiaolan Fu. 2007. Understanding, measuring, and designing user experience: The causal relationship between the aesthetic quality of products and user affect. In *International Conference on Human-Computer Interaction*. Springer, Springer, Berlin/Heidelberg, Germany, 340–349.