Cross-Modality Matching for Evaluating User Experience of Emerging Mobile EEG Technology

Thea Radüntz and Beate Meffert, *Senior Member, IEEE*

***Abstract*—Emerging technology for brain-state monitoring offers the possibility to conduct measurements outside the laboratory. However, user-experience research is lacking. In this article, we present and test an approach for determining the development of user experience in the course of time using the so-called cross-modality matching (CMM). We conducted experiments with 24 subjects and evaluated seven mobile electroencephalography (EEG) devices. Using the CMM method, we registered the headset pressure of the EEG devices and subject's mood. We are able to identify a correlation between headset pressure and mood and to observe time trends. Subjects rated the heaviest, pin-based device as less comfortable in the course of time. The gel-based EEG cap is the most comfortable device regarding its long-time properties. The CMM approach for user-experience evaluation of new EEG technologies is direct, rapid, and easy to perform. This fact creates new opportunities for future studies in the field of user experience and human factors.**

***Index Terms*—Dry sensors, electroencephalography (EEG), psychophysical methods, usability testing and evaluation, wearable devices.**

I. Introduction

Registration of brain activity by means of electroencephalography (EEG) outside the lab is of increasing interest but also coupled with various challenges. The lack of research about user acceptance regarding the measuring technique is one of them. Meanwhile, mobile and easier to use EEG devices are emerging. They make use of wireless signal transmission and allow the subject to move more freely. Additionally, gel-free sensors enable a quick and easy application of the electrodes. The wearing comfort of the new devices is still unknown, as well as whether user acceptance is improved relative to traditional EEG acquisition. For the use of the devices in future studies, it is of major importance that they do not cause head pressure, discomfort issues, or alter subject's mood state. This is particularly important if the subjects are asked to wear the device for a longer period of time. Knowledge about the wearing time, which is free of complaints or inconvenience can be especially important for many investigations.

Usability studies with more than ten subjects and a within-subject design for the comparison of more than three mobile, consumer-grade EEG devices are rare. Little is known about the evolution of comfort and the influence of the device on subject's mood in the course of time [1]. The few studies involving user-experience research concentrated on one device and used the traditional method of questionnaires to register subjective ratings ([2]–[5]). In the study of Ekandem *et al.* [6] participants were asked to evaluate two devices by completing a post-experiment comfort survey after 15 min of wearing the device. Three different EEG headsets were tested by Nijboer *et al.* [7]. The 13 subjects participating wore every device for approximately an hour during three sessions. At the end of each session, they answered questions regarding the usability of the headset by means of questionnaires. The study by Izdebski *et al.* [1] consisted of two experiments. During the first experiment, four devices were tested by four subjects while during the second experiment three devices were tested by nine subjects. Duration of the sessions varied between one and three hours and the usability was assessed at the end of each session by a questionnaire. Hairston *et al.* [8] conducted a usability research experiment with a wearing time of 60 min and three wireless devices. At the end of the session, participants provided comfort ratings by means of a Likert scale and overall preference ratings based on an ordinal scale. At the end of their article, the authors stated that future studies should include the evolution of the ratings over time and not only the subjective ratings conducted after the experiment. This was a major goal of our study.

For this, we employed the method of cross-modality matching (CMM). The CMM method can be traced back to psychophysical research that aims to describe the relationship between changes in the amplitude of a physical stimulus and the subjective perception of these variations.

To recap, an important psychophysical question is the quantitative relation between a stimulus S and its subjective perception. First relations were found experimentally by Weber in 1864. They were characterized by the so called just-noticeable difference JND that described the smallest change ΔS that could be perceived between two stimuli. In this context, Weber noticed that the greater the initial stimulus S, the larger the difference ΔS needed to distinguish between a first and a second stimulus and that the relation was a constant k (Weber's law)

$$k \sim \frac{\Delta S}{S}. \tag{1}$$

Manuscript received October 29, 2019; revised February 12, 2020; accepted April 5, 2020. Date of publication June 1, 2020; date of current version July 14, 2020. This work was supported by the Federal Institute for Occupational Safety and Health under Project F 2402. This article was recommended by Associate Editor D. Wu. *(Corresponding author: Thea Radüntz.)*

Thea Radüntz is with the Unit Mental Health and Cognitive Capacity, Federal Institute for Occupational Safety and Health, 10317 Berlin, Germany (e-mail: raduentz.thea@baua.bund.de).

Beate Meffert is with the Department of Computer Science, Humboldt-Universität zu Berlin, 12489 Berlin, Germany (e-mail: meffert@informatik.hu-berlin.de).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/THMS.2020.2989380

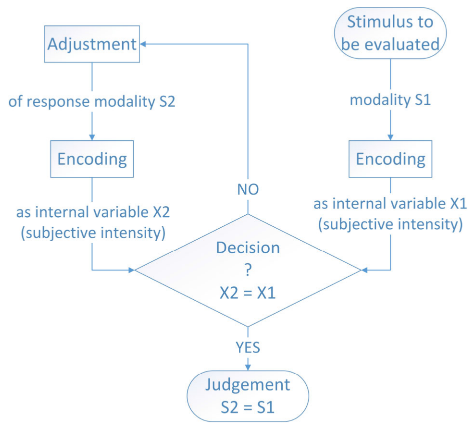

Fig. 1. Principle of CMM according to Sydow and Petzold [31].

Fechner, a scholar of Weber, found that the relation between stimulus S and perception P was logarithmic. In 1957, Stevens introduced an extension ([9]) that showed that sensation magnitude was a power function of stimulus intensity (Stevens' power law)

$$P = b \cdot S^m. \tag{2}$$

Both parameters (the constant b and Stevens' exponent m) are specific for each modality and were already determined for many different ones (e.g., brightness, loudness, apparent length).

Based on the fact, that perceptions of different modalities could be compared to each other, it could be concluded that a stimulus intensity S_1 of one modality could be described by a stimulus intensity S_2 of another modality

$$b_1 \cdot S_1^{m_1} = b_2 \cdot S_2^{m_2} \tag{3}$$

$$\log_{10} S_1 = \frac{m_2}{m_1} \cdot \log_{10} S_2 + \frac{\log_{10} b_2 - \log_{10} b_1}{m_1}. \tag{4}$$

The principle of CMM relies on the idea of perception equalization between different modalities (see Fig. 1). This way, a not measurable modality (e.g., discomfort) can be expressed by a measurable physical modality.

Currently, the method of CMM is gaining again more attention in the scientific community. Researchers show increasing interest to explore [10]–[12] and use the CMM method in order to study human factors and usability aspects [13], [14]. The CMM method can be applied in several situations and research studies. It also provides a good option for conducting ratings from children. The basic idea is similar to a standard procedure used by pediatricians to assess children's pain. They ask the child to press their hand as strong as the pain is. This way, a not measurable modality (e.g., pain, discomfort) can be expressed by a measurable physical modality (e.g., grip force). Application of CMM in user-experience research appears quite appropriate due to the fact that the method can give estimates of sensation's magnitude and, hence, of subjective perception. CMM can be conducted in the course of time and provides real-time measurements. Pepermans and Corlett [15] stated that CMM was well applicable in ergonomics for the evaluation of perceived environmental conditions, which could be difficult to measure subjectively and for the investigation of pain, discomfort, or well-being of a person [16]. At the same time, Pepermans and Corlett [15] conceded that the CMM method has not experienced an extended use in ergonomics. There exist a number of articles employing CMM in order to study somatosensory perception ([17]–[23]), pleasure and pain ([24]–[28]), or discomfort ([29], [30]). The research of Forta *et al.* [30], as one of the latest published articles focusing on practical ergonomics, is the most relevant item to our study. Similar to our aim to assess user experience of several EEG devices, they used CMM for obtaining subject's subjective comfort regarding whole-body vibrations while sitting.

To the best of our knowledge, the CMM method has rarely been used in the context of user-experience research and has never been used for evaluating emerging mobile EEG technology. According to the International Organization for Standardization, user experience is defined as user's perceptions during the use of a product. Thereby, "users' perceptions and responses include the users' emotions, beliefs, preferences, perceptions, comfort, behaviors, and accomplishments that occur before, during, and after use" ([32], Section 3.15). In this study, we focused on two factors of user experience: comfort and mood. We employed the hand-grip force as a modality for CMM ratings. By this, we registered the experienced head pressure caused by the EEG headsets, subject's general mood state, and their change in the course of time. The employment of hand-grip CMM for assessing mood and head pressure evolvement as connected to emerging EEG technology is totally new.

In general, we expected that the wearing time of the devices would have an impact on the comfort. Most of the few studies related to wearing comfort of dry-EEG devices did not explicitly report on the influence of wearing time of the devices (e.g., [3], [33]). The ones that did, reported a duration in the range between 15 and 60 min ([2], [6], [8]). We assumed that subjects' perception of headset pressure would increase after half an hour while their current mood would become worse as long as the headset was worn.

We also addressed the relation between discomfort and mood that had its roots in the research area of embodied cognition, particularly embodied emotion [34]. In this context, the physical condition of a human has a direct influence on the mental state. Hence, we assumed that subject's mood was positively correlated to the wearing comfort of the devices as assessed by the individual perception of head pressure caused by the headset. For the case of a positive correlation, we further assumed that head pressure mediated the relation between device properties and mood. Taken together, we formulated the following hypotheses.

1) During wearing time of the devices, the headset pressure of all EEG devices will increase and subject's mood will get worse, regardless of model, or electrode type.
2) There is a significant positive correlation between current head pressure from the EEG headset and subject's current mood.
   a) The number of electrodes has an effect on the head pressure and, thus, influences subject's mood.
   b) Device's weight has an effect on the head pressure and thus, influences subject's mood.

TABLE I
EEG DEVICES TESTED

Device (manufacturer)	Electrode type	Number of electrodes	Weight
MindCap (mindTec)	Dry sensors	1	119 g
4S Jellyfish (Mindo)	Dry: foam-based	4	95 g
BR8+ (BRI)	Dry: spring-loaded and foam-based	8	269 g
EPOC (Emotive)	Saline based wet sensors	14	116 g
g.SAHARA (g.tec)	Dry pin sensors	16	233 g
g.LADYbird (g.tec)	Gel based	16	165 g
32 Trilobite (Mindo)	Dry: spring-loaded and foam-based	32	524 g

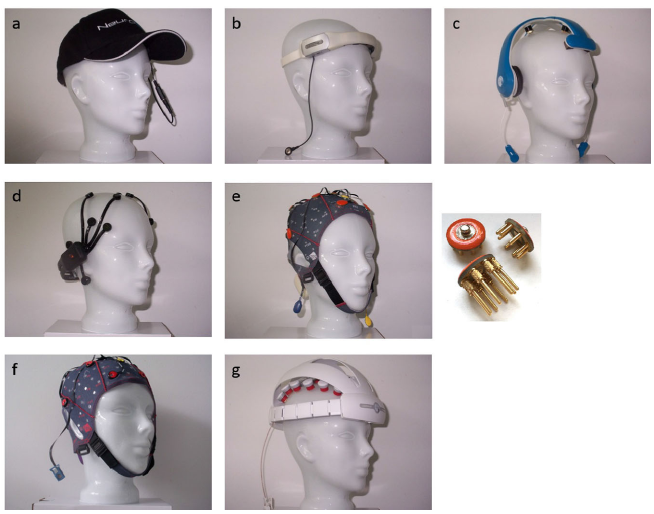

Fig. 2. EEG devices used: (a) MindCap; (b) 4S Jellyfish; (c) BR8+; (d) EPOC; (e) g.SAHARA with dry electrodes; (f) g.LADYbird; and (g) 32 Trilobite.

II. Materials and Methods

A. EEG Systems

We conducted market research and chose EEG devices that had left the research-prototype state. These were expected to be suitable for field studies, i.e., quickly and easily applicable without limiting subject's movement while sitting. Seven mobile EEG devices with different characteristics were purchased (see Table I, Fig. 2). In total, six of them were equipped with gel-free electrodes. We also included g.tec's g.LADYbird/g. Nautilus system as a standard gel-based device well suited for mobile use due to its wireless signal transmission and the use of active electrodes.

B. Procedure and Subjects

Our study took place in an office where only the subject and supervisor were present. In total, 24 subjects participating (11 females and 13 males, 26–66 years of age, with a mean age of 42.8) completed in the course of eight consecutive workdays a total of eight sessions with duration of about 90 min each. The first session was aimed at familiarizing the subjects with the method of CMM and the computer tasks they had to perform while wearing the EEG devices. We instructed the subjects that we will not evaluate their performance because the main goal of our study was the evaluation of the devices. During the following sessions, one device per day was selected in

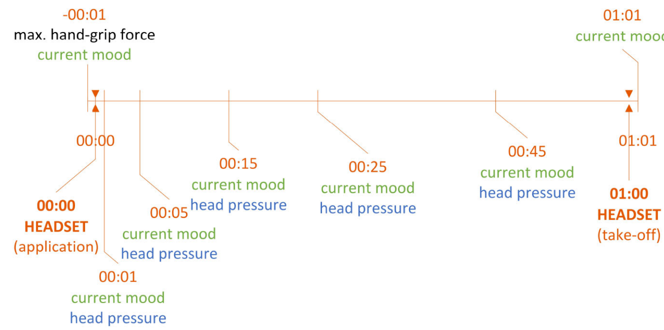

Fig. 3. Timeline of daily sessions for the CMM registration.

Fig. 4. LabQuest2 interface for data display. (a) Used by the experimenter and hand dynamometer. (b) Used by the subject [36].

random order and tested independently of the others. The study was task independent with strong focus on devices' comfort evaluation. In order to control for side effects related to the tasks, we kept task sequence identical for each device. For more information, we want to draw readers' attention to our paper about the signal-quality evaluation of the devices ([35]). The timeline of the daily session is presented in Fig. 3. The subjects were wearing each device for approximately 60 min. All of the investigations acquired were approved by the local review board of our institution and complied with the tenets of the Declaration of Helsinki. All procedures were carried out with the adequate understanding and written consent of the subjects.

During the following sessions, CMM measurements aimed at evaluating our hypotheses related to the EEG devices. At clearly defined registration time points (see Fig. 3), subjects used the hand-grip force device to answer questions regarding the current head pressure caused by the EEG device and their current mood. They were instructed to apply a greater hand-grip force, the more negative their current mood was. Immediately thereafter, the experienced headset pressure was registered similarly. For doing so, subjects were instructed to grip stronger, the bigger the experienced head pressure of the device was.

Hand-grip force was registered with a strain-gauge-based hand dynamometer by Vernier Company (see Fig. 4). Measurement of hand-grip force was performed with subject's dominant hand and all subjects were right handed. All values were related to subject's maximal grip force that was measured at the beginning of every experimental day.

III. Results

For all subsequent calculations, we used the logarithms of the relative grip-force values and proceeded as described in the

TABLE II
RESULTS OF THE ANOVAS FOR HEADSET PRESSURE AND CURRENT MOOD ASSESSED BY CMM ACROSS REGISTRATION POINTS AND DEVICES

		F	p	η^2
Registration points	Headset pressure	10.083^a	**.001**	.305
	Current mood	8.882^a	**.001**	.279
Device	Headset pressure	6.101	**.001**	.210
	Current mood	1.887^a	.122	.076
Registration points and device	Headset pressure	2.789^a	**.005**	.108
	Current mood	2.150^a	**.039**	.086

Note: Values of .001 are actually $p \leq .001$.
[a] Indicates Mauchly's test of sphericity was significant ($p < .05$) and a Greenhouse–Geisser correction was made to degrees of freedom.

following. All statistical calculations were carried out by means of the SPSS software.

A. Evolvement of Headset Pressure and Mood in the Course of Time

Our first hypothesis assumed that headset pressure would increase in the course of time for all devices and subject's current mood would become worse. We carried out two analyses of variance (ANOVAs) in order to find out if there were significant differences between the registration points, devices, and if there was an interaction between both. The dependent variable was either the head pressure or current mood assessed by CMM. For each ANOVA, we utilized a repeated-measures design with two within-subject factors (seven levels for the device factor and five or seven levels for the registration-points factor for head pressure and current mood, respectively). Results are summarized in Table II. General differences between the levels were examined and tested with posthoc tests (Bonferroni corrected).

CMM values for headset pressure revealed a significant main effect for time, device, and an interaction between both. Fig. 5 (top) shows the headset pressure averaged over the registration points and subjects for each device. Bonferroni corrected posthoc tests showed significant differences between the Trilobite and Jellyfish ($p = .014$), EPOC ($p = .043$), and g.LADYbird ($p = .003$) as well as between the g.LADYbird and BR8+ devices ($p = .014$). The head pressure was increased for the Trilobite device and lowest for the g.LADYbird. Regarding subjects' mood CMM values indicated a significant main effect for time and a weekly significant effect for the interaction between time and device. No significant main effect could be found for the device factor (Fig. 5, bottom). Results for the posthoc tests regarding the registration points and the nature of the interaction between the two factors are shown in Fig. 6 (bottom right) for headset pressure and in Fig. 7 (bottom right) for current mood.

In general, headset pressure decreased five minutes after application of the devices, gradually increased thereafter, and reached its maximum value in 45 min (see Fig. 6). This held true for almost all devices except for the g.SAHARA and g.LADYbird that revealed a flat temporal evolvement. For testing the differences between the registration points for each device separately, we used seven one-factorial, repeated measures ANOVAs with head pressure as dependent variable. Significant differences were obtained only for the Trilobite device (F(2.44; 56.22) $= 17.97$, $p < .001$, $\eta^2 = .439$). Bonferroni corrected posthoc tests revealed a significant difference between the registration point immediately after application of the device and the first 5 min. Thereby, the pressure decreased. Significant changes could also be obtained between the means of the 5th min and the 25th min and between the means of the 45th min and all other registration points before. In these cases, the headset pressure increased significantly in the course of time.

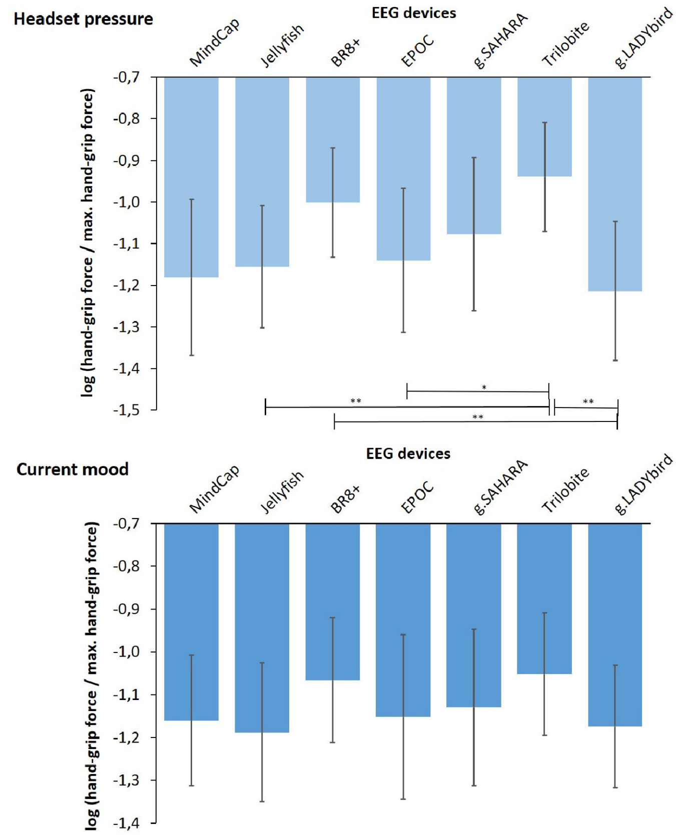

Fig. 5. Headset pressure and current mood as measured by the hand-grip force and averaged over the registration points and subjects for each device (the stronger the grip, the bigger the experienced headset pressure of the device and the more negative the current mood; calculation of analysis of variance with repeated measures design and Bonferonni-corrected posthoc tests: $***$: $p \leq .001$; $**$: $.001 < p \leq .01$; $*$: $.01 < p \leq .05$; error bars indicating the 95% confidence interval).

Descriptive evaluation of subjects' mood revealed that 5 min after device wearing the mood got better for most devices. Thereafter subjects' mood got worse in the course of time and became better after take off of the device. An exception was the g.LADYbird device that did not show the same tendency. In order to statistically evaluate differences between registration points for each device, we computed one-factorial, repeated measures ANOVAs with subjects' current mood as dependent variable. We found significant differences for the devices MindCap (F(3.92; 90.24) $= 2.06$, $p = .09$, $\eta^2 = .082$), BR8+ (F(2.74; 63.07) $= 7.75$, $p < .001$, $\eta^2 = .252$), and Trilobite (F(2.93; 67.31) $= 7.24$, $p < .001$, $\eta^2 = .239$). For the MindCap device Bonferroni corrected posthoc tests showed a significant difference between the registration point after take off of the

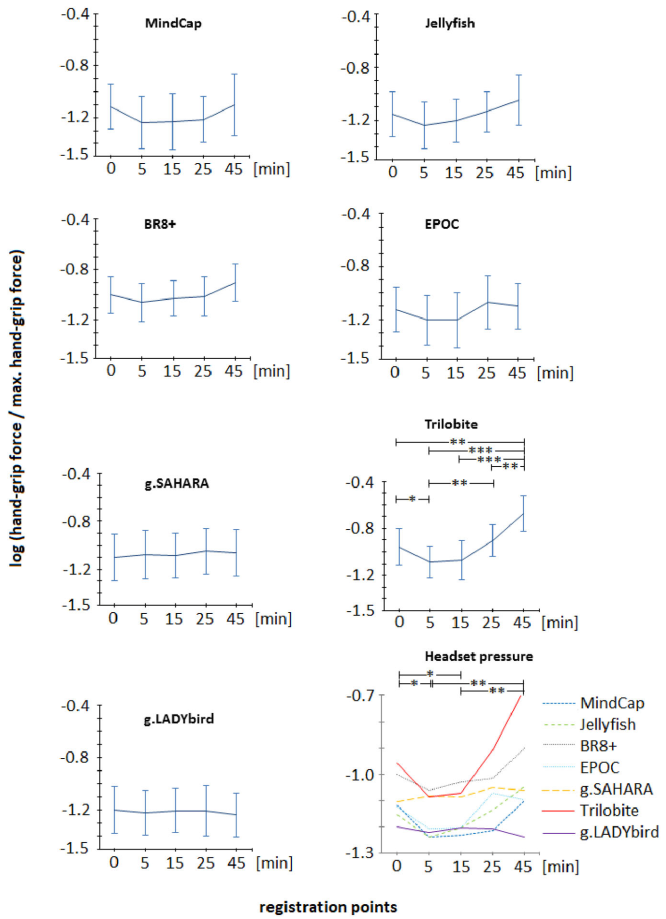

Fig. 6. Headset pressure: Development of the headset pressure measured by the hand-grip force in the course of time and averaged over the subjects for each device and registration point (the stronger the grip, the bigger the experienced headset pressure of the device; calculation of analysis of variance with repeated measures design and Bonferonni-corrected post-hoc tests: $***: p \leq .001$; $**: .001 < p \leq .01$; $*: .01 < p \leq .05$; error bars indicating the 95% confidence interval).

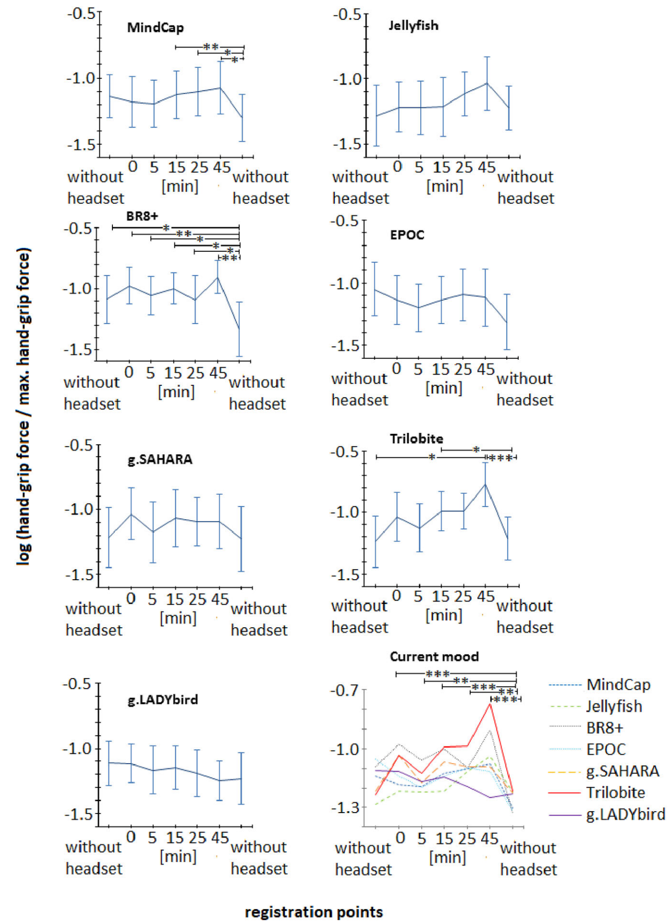

Fig. 7. Current mood: Development of the current mood measured by the hand-grip in the course of time and averaged over the subjects for each device and registration point (the greater the hand-grip force value, the more negative the current mood; calculation of analysis of variance with repeated measures design and Bonferonni-corrected posthoc tests: $***: p \leq .001$; $**: .001 < p \leq .01$; $*: .01 < p \leq .05$; error bars indicating the 95% confidence interval).

headset and the mood in the 15th, 25th, and 45th min. Similar significant differences indicating that the mood got better after take off of the headset were found for the BR8+ device. We observed significant changes in the means between the registration point without the device at the end and all other registration points. Posthoc tests for the Trilobite device revealed that subjects' mood decreased significantly between the registration point before the application of the device and the 45th minute of wearing. Moreover, subjects' mood became significantly better after removal of the device compared to the registration points of the 15th and 45th min, respectively.

B. Correlation Between Headset Pressure and Subject's Mood

We proceeded with the investigation of the relation between subject's mood and experienced head pressure caused by the EEG headsets. To recap, subjects were instructed to apply a greater hand-grip force, the more negative their current mood was. Similarly, they were asked to grip stronger, the bigger the experienced head pressure of the device was. Hence, with

TABLE III
CORRELATIONS BETWEEN HEADSET PRESSURE AND SUBJECT'S CURRENT MOOD FOR EACH DEVICE ($N = 24$, **: $p \leq .01$)

	Pearson's correlation coefficient r
MindCap	0.828**
Jellyfish	0.884**
BR8+	0.964**
EPOC	0.772**
g.SAHARA	0.924**
g.LADYbird	0.838**
Trilobite	0.910**

increasing hand-grip values for the headset pressure, we expected higher grip-force values for the current mood as well.

We computed the means for both, the current mood and the headset pressure over subject's single values from the five registration points. This was done for each device separately in order to have an overall value of head pressure and mood from the whole session for each subject and device. In the following, we calculated the correlations between headset pressure and mood for the devices. The results were highly significant, as shown in Table III. All of the obtained effect sizes for the correlation

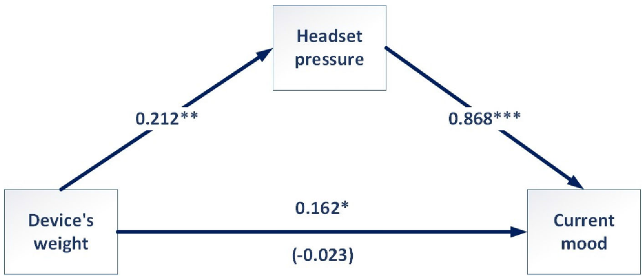

Fig. 8. Standardized regression coefficients for the relationship between device's weight and subject's current mood as mediated by headset pressure. The standardized regression coefficient between device's weight and current mood, controlling for headset pressure, is in parentheses ($***$: $p \leq .001$; $**$: $.001 < p \leq .01$; $*$: $.01 < p \leq .05$; $N = (7 \text{ devices} \times 24 \text{ subjects}) = 168$).

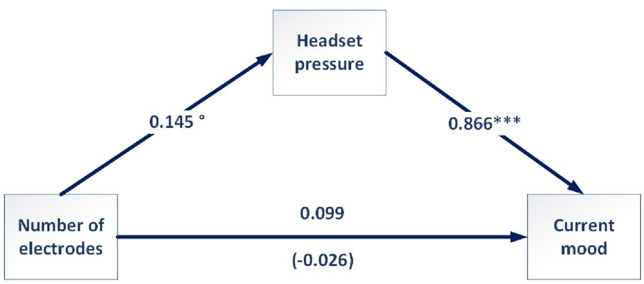

Fig. 9. Standardized regression coefficients for the relationship between number of electrodes and subject's current mood as mediated by headset pressure. The standardized regression coefficient between number of electrodes and current mood, controlling for headset pressure, is in parentheses ($***$: $p \leq .001$; $\circ$: $p = .06$; $N = (7 \text{ devices} \times 24 \text{ subjects}) = 168$).

coefficients of the devices could be interpreted as large according to the guidelines of Cohen denoted for r ([37]).

In the following, we wanted to know if head pressure mediated the relation between device properties and mood. As postulated in our two subhypotheses, the number of electrodes or device's weight could have an effect on the head pressure and, thus, influence subject's mood.

The relationship between device's weight and subject's current mood was mediated by head pressure. As Fig. 8 illustrates, the standardized regression coefficient between device's weight and head pressure was statistically significant, as was the standardized regression coefficient between head pressure and subject's mood. The standardized indirect effect was $0.21 \times 0.87 = 0.18$. Standardized indirect effects were computed for each of 5000 bootstrapped samples, and the 95% confidence interval was computed by determining the indirect effects at the 2.5th and 97.5th percentiles. The results indicated the indirect coefficient was significant ($b = 0.18$, SE $= 0.06$, 95% CI $= [0.07, 0.29]$). Device's weight was no longer a significant predictor of mood after controlling for the mediator head pressure ($b = -0.02$, SE $= 0.04$, $p = .58$). That is consistent with full mediation.

Results of the investigation of the relationship between number of electrodes and subject's current mood as mediated by head pressure are shown in Fig. 9. The standardized regression coefficient between number of electrodes and head pressure was not significant but near significance level ($p = .06$), while the standardized regression coefficient between head pressure and subject's mood was highly significant. The standardized indirect effect was $0.15 \times 0.87 = 0.13$. Standardized indirect effects were computed for each of 5000 bootstrapped samples, and the 95% confidence interval was computed by determining the indirect effects at the 2.5th and 97.5th percentiles. The results indicated the indirect coefficient was significant ($b = 0.13$, SE $= 0.06$, 95% CI $= [0.01, 0.25]$). There was no significant total effect between number of electrodes and mood ($b = 0.1$, SE=0.08, $p = 0.2$), i.e., number of electrodes did not directly predict subject's mood but only indirectly through head pressure.

IV. Discussion

The main aim of our study was to use the rarely-used method of CMM in order to investigate user-experience issues of emerging EEG technology in the course of time. We employed 24 subjects, tested seven different mobile EEG devices, and conducted ratings of headset pressure and subject's mood by means of CMM.

A. *Evolvement of Headset Pressure and Mood in the Course of Time*

We expected that in the course of time all EEG headsets would be perceived as burdensome, regardless of model, or electrode type. We also expected that subject's mood would become worse over the headset's wearing time. The hypothesis could not be confirmed for all devices. For headset pressure, we obtained a significant main effect not only regarding registration points but also regarding the devices. Additionally, there was an interaction effect between both. Surprisingly, after 5 min of wearing, headset pressure decreased for almost all devices and became significant for the Trilobite device. Here, we assumed that subjects were familiarized with the new device on their head and the initial discomfort decreased. We have to note that in our study, subjects did not have any previous experience with mobile, dry-electrode EEG devices. Thus, the extent to which the found relations depend on headset experience remains an interesting topic for future research.

In the course of time, the headset pressure generally increased until the 45th minute. This increase was particularly prominent for the Trilobite, our heaviest device. This fits well to our results from the mediator analysis. Furthermore, the Trilobite device showed significant differences regarding head pressure to the devices with soft electrodes (i.e., to the Jellyfish with foam-based electrodes, EPOC with felt-pad electrodes, and g.LADYbird with gel electrodes). For the sake of correctness, we have to mention that these were also the lightest devices. The g.LADYbird device seemed to be the most comfortable device with significant differences to the BR+ and Trilobite devices. It revealed no head-pressure evolvement in the course of time and no significant differences between the registration points. This could be responsible for the highly-significant interaction effect of registration points and device.

Regarding subjects' mood no differences between devices could be obtained although there was a significant interaction between device and registration points. In general, we observed that after the headsets were removed from subject's head the mood became obviously better for all devices with significant differences to all previous measurements. An exception was observed for the g.LADYbird device where subjects' mood remained almost constant not only in the course of time but also after the removal of the headset. This might be a reason for the weakly-significant interaction effect.

B. Correlation Between Headset Pressure and Subjects' Mood

In our second hypothesis, we suggested a positive significant correlation between headset's comfort and subject's current mood. Correlation analysis yielded large, positive, and significant correlation coefficients for all devices. We concluded that subject's mood was highly correlated to the headset pressure and, thus, to the wearing comfort of the devices.

Furthermore, we investigated if headset pressure mediated the effect of device properties on subject's mood. Results indicated that device's weight was a significant predictor of head pressure and that head pressure was a significant predictor of subject's mood. This supported the mediational hypothesis. After controlling for the mediator, the significant relationship of device's weight and mood became insignificant, indicating a full mediation. This result seemed reasonable because a greater weight could contribute to a greater head pressure and lead to a worse mood.

In contrary, the number of electrodes had only a significant indirect effect through head pressure but no significant total effect on mood. We assumed that there might be other factors apart from the number of electrodes affecting both head pressure and mood (as was the case with device's weight). These might confound the head pressure-mood relationship of our second model. The predictor (in this case the number of electrodes) might be only a part of a more complex model. For instance, the type of electrode could have a greater involvement in subject's current mood than the number of electrodes. This impact could been even amplified by, e.g., device's weight or wearing duration.

Finally, we must be aware that other factors might exist influencing subject's current mood during the sessions. These might be related to the environmental conditions, time on task, or the interaction with the investigator. The randomized testing of the devices across subjects tried to account for some of them. Although, our sample size was relatively large for this kind of study, it was fairly small for elaborate inferential statistics. As a further limitation, we have to mention that during this study new devices appeared on the market, e.g., the actiCAP Xpress Twist/LiveAmp and the saltwater-based electrode system R-Net both by BrainProducts or the new highly innovative approach using in-ear EEG technology ([38], [39]). For evaluating these and further emerging EEG technology, our study design and the proposed CMM method could easily be used. Taken the CMM results as a benchmark to make across-group comparisons [27] would allow for an integration of the test results from new devices into the findings already in existence. This would make it possible to compare emerging EEG devices. Future studies could also evaluate possible effects of specific tasks on user experience of EEG devices as well as further aspects like appealing design, emotions, and pleasure by means of CMM.

V. Conclusion

To sum up, subject's mood and headset pressure were related to each other and changed over the wearing time. This alternation was particularly prominent for the Trilobite device where the changes in the course of time became significant. In contrast, the g.LADYbird device seemed to be the most comfortable. We also found that head pressure was a mediator between device properties and subject's mood, with device's weight as significant predictor. We conclude that developers should attach importance to the weight of the headset for assuring comfort and well-being caused by their devices. In this respect they should be aware of possible interaction effects between the weight, electrode type, and the number of electrodes.

For our investigation, we made use of the method of CMM that is gaining again more attention in the scientific community [40]–[43]. We presented and tested this psycho-physiological approach for evaluating user experience. By this, we compared seven mobile EEG devices and gained reasonable results in the course of time. Although our results might not be surprising, they provide evidence about the feasibility and quality of the CMM ratings. Compared to traditional methods for subjective ratings the CMM approach is direct, rapid, and easy to perform. These facts create new opportunities for future studies in the field of user experience, experimental psychology, and human factors research. Furthermore, our results provide scientific feedback regarding the comfort claims of manufacturers of emerging EEG technology. They are of particular interest for researchers that want to use the new wearable devices for their studies.

In general, CMM offers good possibilities to overcome linguistic or reading barriers or to assess ratings from cognitively impaired subjects. Furthermore, subjects are not limited to a preset scaling or limited number of answers and, thus, less prone to social desirability restrictions caused by predefined answers that could be interpreted as right or wrong. We hope that our article contributes not only to the user-experience evaluation of emerging EEG devices in the course of time but offers also a new example with positive results regarding the applicability of the CMM method.

Acknowledgment

The authors would like to thank G. Menzel and R. Blüthner for their technical and overall support, our student assistant F. Schröder for conducting the experiments, and M. Freyer and the student assistants Y. Cao and I. Pritschke for daily operational support and graphic editing. Furthermore, we would like to thank G. Freude and U. Rose for their general project support. More information about the project where our data were acquired can be found online.[1]

Author contributions: T. Radüntz initiated the project and was responsible for the overall conception of the investigation. Data analysis was performed by T. Radüntz. Data interpretation was performed by T. Radüntz and B. Meffert. The manuscript was written by T. Radüntz. Final critical editing was performed by B. Meffert.

[1][Online]. Available: http://www.baua.de/DE/Aufgaben/Forschung/Forschungsprojekte/f2402.html

Ethics statement: All of the investigations acquired were approved by the local review board of the Federal Institute for Occupational Safety and Health and complied with the tenets of the Declaration of Helsinki. All procedures were carried out with the adequate understanding and written consent of the subjects.

References

[1] K. Izdebski *et al.*, "Usability of EEG systems: User experience study," in *Proc. 9th ACM Int. Conf. Pervasive Technologies Related Assistive Environ.*, 2016, pp. 34-1–34-4. [Online]. Available: http://doi.acm.org/10.1145/2910674.2910714

[2] V. V. Nikulin, J. Kegeles, and G. Curio, "Miniaturized electroencephalographic scalp electrode for optimal wearing comfort," *Clin. Neurophysiology*, vol. 121, pp. 1007–1014, 2010.

[3] C. Grozea, C. Voinescu, and S. Fazli, "Bristle-sensors—low-cost flexible passive dry eeg electrodes for neurofeedback and BCI applications," *J. Neural Eng.*, vol. 8, no. 2, 2011, Art. no. 025008. [Online]. Available: http://stacks.iop.org/1741-2552/8/i=2/a=025008

[4] E. Mathe and E. Spyrou, "Assessment of user experience with a commercial BCI device," in *Proc. Hellenic Conf. Elect. Comput. Eng. Students*, 2015.

[5] T. O. Zander *et al.*, "Evaluation of a dry EEG system for application of passive brain-computer interfaces in autonomous driving," *Frontiers Human Neuroscience*, vol. 11, p. 78, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnhum.2017.00078

[6] J. Ekandem, T. Davis, I. Alvarez, M. James, and J. Gilbert, "Evaluating the ergonomics of BCI devices for research and experimentation," *Ergonomics*, vol. 55, no. 5, pp. 592–598, 2012. [Online]. Available: http://dx.doi.org/10.1080/00140139.2012.662527

[7] F. Nijboer, B. van de Laar, S. Gerritsen, A. Nijholt, and M. Poel, "Usability of three electroencephalogram headsets for brain-computer interfaces: A within subject comparison," *Interacting Comput.*, vol. 27, no. 5, pp. 500–511, 2015. [Online]. Available: https://doi.org/10.1093/iwc/iwv023

[8] W. Hairston *et al.*, "Usability of four commercially-oriented EEG systems," *J. Neural Eng.*, vol. 11, no. 4, 2014, Art. no. 046018. [Online]. Available: http://stacks.iop.org/1741-2552/11/i=4/a=046018

[9] S. Stevens, "On the psychophysical law," *Psychological Rev.*, vol. 64, no. 3, pp. 153–181, 1957.

[10] T. M. Benz and V. Nitsch, "Is cross-modal matching necessary? A bayesian analysis of individual reference cues," in *Haptics: Science, Technology, and Applications*, D. Prattichizzo, H. Shinoda, H. Z. Tan, E. Ruffaldi, and A. Frisoli, Eds. Berlin, Germany: Springer, 2018, pp. 15–26.

[11] K. M. Gomes and S. L. Riggs, "The effect of age on crossmodal matching using auditory frequency," in *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, vol. 61, no. 1, pp. 1552–1556, 2017. [Online]. Available: https://doi.org/10.1177/1541931213601752

[12] K. Gomes and S. L. Riggs, "Crossmodal matching: A comparison of two methods," in *Proc. Human Factors Ergonom. Soc. Annu. Meeting*, vol. 60, no. 1, pp. 1595–1599, 2016. [Online]. Available: https://doi.org/10.1177/1541931213601368

[13] M. C. Lee and J. Park, "There is no perfect evaluator: An investigation based on prospect theory," *Human Factors Ergonom. Manuf. Service Industries*, vol. 28, no. 6, pp. 383–392, Jun. 2018.

[14] S. L. Riggs and N. Sarter, "Tactile, visual, and crossmodal visual-tactile change blindness: The effect of transient type and task demands," *Human Factors*, vol. 61, no. 1, pp. 5–24, 2019. [Online]. Available: https://doi.org/10.1177/0018720818818028

[15] R. G. Pepermans and E. N. Corlett, "Cross-modality matching as a subjective assessment technique," *Appl. Ergonom.*, vol. 14, no. 3, pp. 169–176, 1983. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0003687083900789

[16] A. Campbell, "Subjective measures of well-being," *Amer. Psychologist*, vol. 31, pp. 117–124, 1976.

[17] L. Walker and P. Walker, "Cross-sensory mapping of feature values in the size-brightness correspondence can be more relative than absolute," *J. Exp. Psychol., Human Perception Perform.*, vol. 42, no. 1, pp. 138–150, 2016.

[18] T. N. Coates, "Source localization in cross modality matching of brightness and loudness of young adults," Communication Disorders, Master's Thesis, David O. McKay School of Education, Provo, UT, USA, 2015.

[19] J. Hayes, "Measuring sensory perception in relation to consumer behavior," in *Rapid Sensory Profiling Techniques*. New York, NY, USA: Elsevier, 2015, pp. 53–69.

[20] S. Simonetti, J. Kim, and C. Davis, "Cross-modality matching of linguistic and emotional prosody," in *Proc. Interspeech*, 2015, pp. 56–59.

[21] J. Venrooij *et al.*,"Perception-based motion cueing: Validation in driving simulation," in *Proc. Eur.: Driving Simul. Conf. Exhib.*, 2015, pp. 153–161.

[22] M. W. Heft and M. E. Robinson, "Age differences in suprathreshold sensory function," *AGE*, vol. 36, no. 1, pp. 1–8, Apr. 2013.

[23] K. Bronner, K. Frieler, H. Bruhn, R. Hirt, and D. Piper, "What is the sound of citrus?," in *Proc. 12th Int. Conf. Music Perception Cognition*, 2012, pp. 142–148.

[24] M. W. Heft and M. E. Robinson, "The challenge of measuring pain in humans," in *Biological Measures of Human Experience Across the Lifespan*. Berlin, Germany: Springer, 2016, pp. 105–115.

[25] C. Morewedge, *Utility: Anticipated, Experienced, Remembered*, 2nd ed. Hoboken, NJ, USA: Wiley, 2016, ch. 10.

[26] G. Lathwal, I. K. Pandit, N. Gugnani, and M. Gupta, "Efficacy of different precooling agents and topical anesthetics on the pain perception during intraoral injection: A comparative clinical study," *Int. J. Clin. Pediatric Dentistry*, vol. 8, pp. 119–122, 2015.

[27] L. Bartoshuk, "The measurement of pleasure and pain," *Perspectives Psychological Sci.*, vol. 9, no. 1, pp. 91–93, Jan 2014.

[28] O. C. G. Colhado *et al.*,"Evaluation of low back pain: Comparative study between psychophysical methods," *Pain Medicine*, vol. 14, no. 9, pp. 1307–1315, Sep. 2013.

[29] Y. Huang and M. J. Griffin, "The discomfort produced by noise and whole-body vertical vibration presented separately and in combination," *Ergonomics*, vol. 57, no. 11, pp. 1724–1738, Aug. 2014.

[30] N. Forta, M. Schust, P. von Löwis, H. Kaiser, and A. Kreisel, "Subjective comfort of fore-and-aft whole-body vibration measured with the cross-modality matching method," in *Proc. 47th U.K. Conf. Human Response Vibration*, Southampton, England, 2012.

[31] H. Sydow and P. Petzold, *Mathematische Psychologie: Mathematische Modellierung und Skalierung in der Psychologie*. New York, NY, USA: Springer, 1982.

[32] EN ISO-9241-210:2019, *Ergonomics of Human-System Interaction—Part 210: Human-Centred Design for Interactive Systems*, Beuth-Verlag, International Organization for Standardization, Geneva, Switzerland, 2019.

[33] A. Pinegger, S. C. Wriessnegger, J. Faller, and G. R. Müller-Putz, "Evaluation of different EEG acquisition systems concerning their suitability for building a brain–computer interface: Case studies," *Frontiers Neuroscience*, vol. 10, p. 441, 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2016.00441

[34] T. F. Price and E. Harmon-Jones, "Embodied emotion: The influence of manipulated facial and bodily states on emotive responses," *Wiley Interdisciplinary Rev., Cogn. Sci.*, vol. 6, no. 6, pp. 461–473, Sep. 2015.

[35] T. Radüntz, "Signal quality evaluation of emerging EEG devices," *Frontiers Physiol.*, vol. 9, p. 98, 2018. [Online]. Available: https://www.frontiersin.org/article/10.3389/fphys.2018.00098

[36] Vernier. (2017) Software and technology. Accessed: Jul. 2, 2017. [Online]. Available: https://www.verniers.com

[37] J. Cohen, "A power primer," *Psychological Bull.*, vol. 112, no. 1, pp. 155–159, 1992.

[38] V. Goverdovsky *et al.*, "Hearables: Multimodal physiological in-ear sensing," *Scientific Rep.*, vol. 7, no. 1, Jul. 2017, doi : 10.1038/s41598-017-06925-2.

[39] D. Looney *et al.*,"The in-the-ear recording concept: User-centered and wearable brain monitoring," *IEEE Pulse*, vol. 3, no. 6, pp. 32–42, Nov. 2012.

[40] S. L. Riggs and N. Sarter, "Crossmodal matching: The case for developing and employing a valid and feasible approach to equate perceived stimulus intensities in multimodal research," *Human Factors*, vol. 61, no. 1, pp. 29–31, 2019. [Online]. Available: https://doi.org/10.1177/0018720818816439

[41] G. Hamilton-Fletcher, K. Pisanski, D. Reby, M. Stefańczyk, J. Ward, and A. Sorokowska, "The role of visual experience in the emergence of cross-modal correspondences," *Cognition*, vol. 175, pp. 114–121, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010027718300593

[42] B. Pitts, S. L. Riggs, and N. Sarter, "Crossmodal matching: A critical but neglected step in multimodal research," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 3, pp. 445–450, Jun. 2016.

[43] R. S. Schaefer, L. J. Beijer, W. Seuskens, T. C. M. Rietveld, and M. Sadakata, "Intuitive visualizations of pitch and loudness in speech," *Psychonomic Bull. Rev.*, vol. 23, no. 2, pp. 548–555, Apr. 2016. [Online]. Available: https://doi.org/10.3758/s13423-015-0934-0