UNIVERSITY OF TWENTE.
Faculty of Engineering Technology

Industrial Design Engineering
Collected publications
programme homepage: utwente.nl/ide

Design approach to predict and optimise the User Experience: Ebike charging system case study.

Peter Hepburn

Emerging Technology Design

Graduation date: 06,09, 2023

Designing a good user experience (UX) is crucial for the success of products and is typically evaluated through user testing after a physical prototype has been produced. Within literature there is a lack of tools to predict and optimise UX during the early stages of design. This research proposes a new predictive design approach to quantify and compare UX components. By focusing on quantifying the mental and physical workloads different ideated procedures could be compared and the best one selected to achieve the optimal UX. The study involved the redesign of an Electric bicycle (Ebike) charging system and assessment of whether there had been a significant improvement in the UX. This was evaluated using different measurable components of UX; Usability Metric for UX (UMUX) and emo-cards. As a result, the newly designed Ebike charging system resulted in a significant improvement of the overall UX with a much smaller variance.

User Experience, Physical workload, Mental workload, Emo-cards, Design research, Industrial Design

1. Introduction

Ebikes are a fast-growing segment in the bicycle market and are gradually superseding the traditional bicycle. Traditionally Ebikes were designed to make cycling accessible for less-abled people or for the older population, however Ebikes are becoming widely adopted by different users of all ages [1]. This is due to several advantages, including increased comfort, power, speed and convenience [2]. All these factors can improve the riding experience, which is why the Ebike market will continue to expand. The design of better-looking Ebikes with greater autonomy, higher performance, and more convenience has benefited from recent developments including thinner batteries with increasing capacity, system integration and better user interfaces [3]. This has allowed more design freedom. However, the main focus on developing battery performance has resulted in the overall system integration and User eXperience (UX) lagging behind. One unavoidable activity which is commonly overlooked is the charging experience. This recurring procedure is rarely experienced as pleasant and is most frequently seen as a burden.

To address this the UX should be improved, and the charging system redesigned. However, UX is primarily used as a reflective tool to monitor the success of a product through user testing, KPI's or Usability metrics [4, 5]. All of which involves a product, an app or a prototype to evaluate the experience as, or after, it happens. There are currently no tools within literature that apply a methodology to anticipate and optimise the UX of a product in the early stages of the design process, before there is a prototype or concept that can be user tested. This research aims to address this research gap by proposing a new predictive design approach that creates a quantifiable component of UX to compare and perform concept selection to improve UX of a new product.

2.0 Understanding UX:

To determine how to predict UX, the components must first be understood. UX can be quantified in multiple ways, but is generally understood as a multi-dimensional construct [4]. Consisting of different aspects such as learnability, aesthetics, and efficiency [4]. UX is defined by ISO 9241-210 as *"A person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service"* [5],

Within literature there are several different ways UX can be classified; these include as a holistic view [4], extension of usability [6] or with a primary focus on emotion [6]. In these cases, the holistic view provides a broad overview, whereas a primary focus on emotion explores how the user's emotion and experience are correlated [6]. Since the UX is very broad topic a study by Sauer, (2020), provides a higher-level concept called user interaction [6]. This involves the incorporation of usability and accessibility within the UX. Traditionally the term usability is used for everyday products, accessibility in housing environments and UX is used within a software development context [6]. For the redesign of a charging system the incorporation of software, physical interfaces as well as the environmental context all play an important role. The incorporation of usability and accessibility as part of the UX is appropriate to include, providing more depth on ways to improve the UX [6]. To visualise the relationship between UX, usability, and accessibility a diagram based on Sauer, (2020) is proposed [6]. This model has been tailored so it can be used as a bases for evaluating the UX of the charging system.

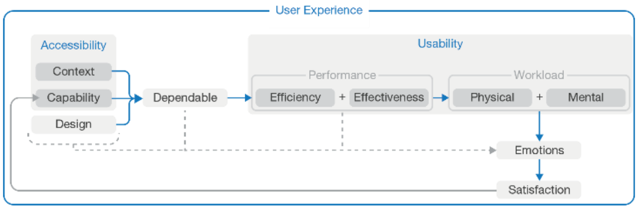

Figure 1: Relationship between UX, usability, and accessibility inspired from Sauer, (2020) is proposed [7].

UX is a continuous loop that starts with accessibility [7]. This involves the context, capability, and design. In the case of this study the context involves where the charger will be used and whether the battery is removed or remained within the bike. The capability involves the user's knowledge, how experienced they are with Ebike charging systems and whether they have used a system before. The design entails how the product looks and how it functions [8]. This has large implications on how the product is interpreted [8]. Dependability is a result of the accessibility components. If the user's capability changes so will the products predictability, reliability, and trustworthiness; all of which are components of dependability [6].

Usability is the core to UX defined by the International standard ISO FDIS 9241-210 as the: 'Extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use' [5]. Compared to UX, usability does not consider the anticipated use or pre-usage phase and only focuses on the direct interaction of the user. The performance component of usability addresses how easy the product is to interpret, which can be achieved with good efficiency and effectiveness. Efficiency involves how long it takes to perform the task (to charge the Ebike) [6]. Effectiveness involves how helpful the feedback is to inform the user of success. Ideally the user should understand when success has occurred and receive feedback on the status of the system. The workload component of usability entails how much effort is required to charge the Ebike, both physically and mentally. By reducing the complexity of decisions that are required to use the product, as well as reducing the amount a user must move to complete a task, will reduce the mental and physical workload respectively.

Emotions concern how the user is feeling during and after the product use. Their emotions will be generated during every use (highlighted by dashed lines in figure 1) of the product and may change after the task has been performed [9]. Finally, once the task is performed the combination of all aspects will lead to a certain level of satisfaction. This satisfaction gives an overview of the UX, however diving deeper into the categories mentioned above will help determine the causation to achieve a high level of satisfaction. After the UX has occurred there is a feedback loop [10]. As a result of a second time usage the user's capability (knowledge) will have changed and they will tackle the procedure in a different manner, resulting in a slightly different emotion and/or level of confidence.

3.0 Measuring UX:

To evaluate the new design and see whether it provides any significant improvement the components of UX need to be quantified. While not every component within UX is easily measurable. three significant components are and were used as the basis for evaluating the UX. The context, which is a component of accessibly, was the independent variable and the user was asked to perform the task of charging an Ebike in two different scenarios: charging the bike and charging the battery. For this study changing the contexts, such as location and lighting levels was not performed but is a method which could be explored in the future. As the user performed the tasks observations were made, while these are not quantifiable, they acted as a basis to see whether the user's self-reflection was consistent. Efficiency can be measured by timing how long each task takes [6]. However, for this user study, timing the user was not performed as this could have introduced anxiety within the testing, or prevented the user from performing the task naturally. To measure the effectiveness observations were made to see whether the user understood when a success was obtained, how they interpreted feedback and whether any errors were made such as keep pushing the charging plug into the socket even though it has been pushed in far enough. For the workload component the amount of movement and effort was documented through observations and discussions with the user.

Measuring emotion is a complex construct to study [10]. There are two main ways to measure the user's emotion: verbal and nonverbal. Verbal measurements tend to be lengthy [10]. Since emotions occurs instantaneously, getting the user to verbally portray how they feel may distort the result. A self-report approach is widely used; however this works best when the user is participating in the task passively [10]. The other challenge in assessing through verbal communication is language barriers. To determine what emotions were felt during each usability test the user was asked to reflect on what emotion they felt using emo-cards developed by Desmet,(2001), [9]. This is a fast and effective way for the user to indicate how they felt during the process: the user predicted what emotion they would have and then after each of the four tasks their level of pleasantness and arousal was documented. To evaluate the results the average level of arousal and average level of pleasantness was calculated treating the emo-cards as if they were on am x,y axis (see figure 3 for example).

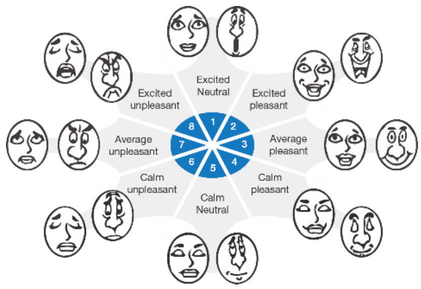

Figure 2: Emocards developed by Desmet, (2001) [9]

To measure the overall satisfaction and usability, the Usability Metric for UX (UMUX) was used. This metric is a four-item Likert scale (from 1-7) to evaluate the perceived usability [11]. It is designed to obtain similar results to the 10-item System Usability Scale (SUS), however fewer questions align best with the ISO 9241-210 [12,11]. This satisfaction/reflection allows for a gut reaction from the user to see if their experience aligns with the documented observations. Each of the four statements from the UMUX give a determination of the effectiveness, satisfaction, overall UX and efficiency.

Jsability Component	Candidate UMUX
1. Effectiveness	The charging system capabilities meet my requirements
2. Satisfaction	Using the charging system is a frustrating experience
3. Overall	The charging system is easy to use
4. Efficiency	I must spend too much time correcting things with the charging system

Table 1: UMUX statements

Based on the response the score is calculated by [score -1] for statements 1 and 3 and [7-score] for statements 2 and 4. This is to remove the positive/negative keying of the items and allows for a minimum score of zero. From this the maximum score is out of 24.

$$Umux\ score = \frac{((\mathrm{EN}) - 1) + (7 - (S)) + ((\mathrm{O}) - 1) + (7 - (EN))}{24} \times 100$$

This numerical result puts the UMUX score on the same scale as the SUS which can be used to determine the system's usability and act as a goal setting reference [12]. It created data to determine which elements of the UX needed addressing. These tests methods were used to evaluate a current Ebike to determine the overall usability and create a set of guidelines to provide focus for the charging redesign. The outcome from this test, observations, UMUX score and emo-card results, highlighted that the area for focus was to improve the efficiency and effectiveness, with three Guidelines:

G1. Reduce the physical and mental workload when removing the battery.
G2. Reduce the number of steps to charge the bike/battery.
G3. Provide fast feedback.

Using these guidelines as a basis the design approach could begin. In redesigning the charging system the focus was on Guideline 1 "*reducing the mental and physical workload"* because guideline 2 and 3 are both elements which contribute to the mental and physical workload.

An analogy is often made between mental and physical load to stress (mental load i.e. task demands) and strain (physical load i.e the impact on the human) [13]. Stress (comprised of multiple features, such as time pressure and task complexity) and strain show variations depending on the type of job the user must complete. To be able to evaluate the concepts and select the best one to develop the design the mental and physical workload needs to be quantified.

4.0 Physical workload Value (PWV):

To measure the physical workload tests are typically conducted in real life using different devices such as posture monitoring instruments, heart rate measurements, blood pressure, and triaxle accelerometers (which measure the joint angle, range of motion, angular velocity, and angular acceleration) [14]. This data can be combined to determine the overall physical workload [14]. However, within literature, no research has been conducted to predict the physical workload and evaluate design concepts before a physical prototype is produced. Therefore, an evaluation method has been created to help determine the 'ideal' concept direction through. An assessment criterion for the physical workload based on scenarios. Based on the research by van der Beek and Westgaard [15, 16]. This approach focuses on the External exposure and is based on the environment (context) rather than the Internal exposure (individual's capability) since the external exposure can be more easily controlled [15, 16].

Working situation: This incorporates the demands and level of decision freedom with the opportunity for the user to develop and improve their situation [16]. This element is based primarily on the user and therefore will not be used in the assessment analysis.

Working method (T): This is the number of tasks (n) the user must perform. Between concepts the distinction of the tasks will be consistent. For some scenarios/concepts the number of tasks will be higher compared to others.

Level of movement (M): This is how much the user would likely have to move to perform each task and their working height. This will be evaluated on a score from 1-10. where 1 = No movement to 10 = High level of movement.

Physical exertion (P): This is primarily the weight of the objects the user must move, so how much muscle strength and endurance is required to carry out the task. This will be evaluated on a score from 1-10. Where 1 = No physical excursion to 10 = High level of physical excursion.

Duration of task (D): This is a prediction of how long the task will take to be performed. It is measured in seconds, while it is difficult to determine the exact time since, it provides a comparison so consistency between design concepts is important. For example, if in multiple concepts the user must transport their battery to the house this task should be considered to take the same time.

Measuring the Overall Physical workload: To measure the physical workload a new value is proposed. This is referred to as the physical workload value (PWV). The following formula is derived based on the elements which together create the external exposure:

$$PWV = \sum_{n=1}^{T} (M_n \cdot P_n \cdot \frac{D_n}{10})$$

4. Mental Workload value (MWV):

Mental workload or cognitive load represents how much of the user's working memory is occupied [17]. Common ways to measure mental workloads is by using methods involving eye tracking, blink rate, heart rate, speech activity, brain activity and Electroencephalogram (EOG) to name a few [17]. Looking at the mental workload against the user's performance. If the workload is not optimal the user's performance will be reduced [17] . An optimal workload is not under or overstimulated [13] . However, the user interaction with Ebike chargers is for a very short period and during user testing on a current Ebike, under stimulation occurred, with a low level of arousal (gathered from the Emo-cards [9]). For the charging system under stimulation is acceptable, due to this short usage timeframe, but overstimulation must not occur [13].

To quantify this mental workload so it can be reduced to prevent overstimulation, involves physical testing which cannot be used to analyse concepts. That said, by looking at the elements of mental workload there is a possibility that these can be measured and quantified for concept comparison. According to cognitive load theory (CLT) there are two types of cognitive load: Intrinsic load and Extraneous load [18]. Intrinsic load refers to the user's working memory. The working memory resources are allocated to process information intended on learning a task [18]. The task complexity and amount of learning that is required based on the user's expertise (capability) which must be used simultaneously results in a greater intrinsic load [19]. Extraneous load involves the way in which the task is presented, external factors such as the physical environment, and/or internal factors such as the emotional state of the user [19]. Based on the intrinsic and extraneous loads the user's capability is independent on the concept, however the task complexity/decision is dependent which mean can be quantified. Additionally, as seen the (mental) workload present during each task therefore to quantify workload four variables will be measured.

Working method (T): This is the number of decisions (n) the user must perform (usually the same number as the number of tasks since during all task decisions need to be made).

Decision Complexity (C): This is how complex the decision is the user must make, based on the number of elements the user must think of and tasks they must perform. This will be evaluated on a score from 1-10. Where 1 = No complexity to 10 = High level of complexity.

Decision Severity (S): Within the extraneous load internal factors are the user's emotions therefore if the decision is going to result in a more severe outcome, then this may result in an increase in negative emotions such as stress or anxiety. The severity of the decision will contribute to the overall working. This will be evaluated on a score from 1-10. Where 1 = no severity to 10 = High level of severity.

Decision Feedback (F): Based on the evaluation the user makes to determine whether they have completed their task/the response of their decision has been confirmed. This involves the effectiveness. If the feedback is clear the user will know they have completed the task correctly and their working memory can be reduced as they are certain they can progress to the next task. This is evaluated on a score from 1-10. Where 1 = Certain the task is complete to 10 = High level of uncertainty that the task is complete.

Measuring the Overall Mental workload: To calculate the mental workload the following formula is derived based on the elements which together contribute to both the Intrinsic and Extraneous loads:

$$MWV = \sum_{n=1}^{T} (C_n \cdot S_n \cdot F_n)$$

Both the MWV and PWV contain several tasks. For each of the designs the number of tasks will vary as well as the scores for the other components. Smaller tasks may result in a lower PWV or MWV. For the current system the MWV and PWV was derived by assigning a value of 1-10 for each of the tasks which were defined that the user had to perform. While the MWV & PWV are subjective both values are comparative.

The MWV's and PWV's were applied in the ideation phase. For every idea a flow diagram was generated predicting each task the user would have to perform such as insert key, rotate key, pull lever to release battery etc and the flow diagram was assigned a MVW & PWV. Through refining these flow charts, by changing, removing or adding tasks the 'ideal' scenario was generated with the lowest MWV & PWV.

Procedure	Original		Optimised	
Workload Type	MWV	PWV	MWV	PWV
Charging Bike	88	9.3	76	4.5
Charging Battery (remove from frame)	113	171	83	130
Un-charging Bike	11	7.8	3	3
Un-charging Battery (insert into frame)	40	128	30	110

Table 2: Difference in the PWV and MWV between the original and optimised design

The final concept was developed to ensure that the flow diagram with the optimised MWV and PWV was fulfilled. Prototypes of a new charger, plug, battery, and how the battery was integrated within the frame, were produced. Using observations, Emo-cards and the UMUX the UX of the original design and the new design was user tested. Nine randomly selected participants with varying knowledge of Ebikes were used to evaluate the experience of the original charging system and the new charging system.

Usability component	Original	Optimised
1. Effectiveness	3.8	5.3
2. Satisfaction	3.3	5.4
3. Overall	3	5.2
4. Efficiency	3.1	5.9
Total/100	*56.67*	*91.33*
Standard Deviation	*19.32*	*6.38*

Table 3: Difference in the average UMUX score of original design and newly optimised design

Comparing the UMUX values the average score of the new system had a significantly higher UMUX value compared to the original system. The new charging system score also has a much smaller standard deviation suggesting that there is more consistency in the experiences between the users. Based on these results and the observations made the users found the new design approach 'simpler and more intuitive', resonating with the reduction in mental and physical workload.

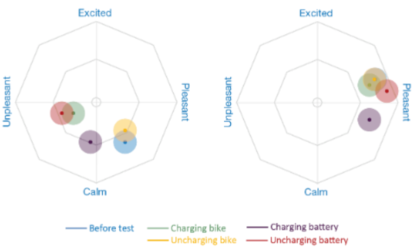

Figure 3.; Emotion response of the original (left) Vs the optimised design (right)

When evaluating the user's emotion for each of the different tests, the new designed system met the user's expectation; the different tasks were perceived as a more pleasant experience, with a medium level of arousal when comparing to the current system. This level of arousal was stated by the responds as 'excitement because they were trying something new'. With a few minor alterations to the design, and if the optimised system was manufactured within a factory environment the tolerance levels would be far less and therefore should create a "better" product and improve the UX further.

5 Conclusion:

A novel comparison score of the mental and physical workload was used as a comparative assessment tool to compare different concepts and their associated procedure (flow) diagrams. This approach led to a final concept with a reduced workload. The results led to an improvement in overall UX across all components: Effectiveness, Satisfaction, Overall and Efficiency, with a much smaller variation in the experience. A similar approach could be suitable for the design of products that involve a series of steps, where performing the task both efficiently and effectively is of high priority; these could include medical applications, kitchen appliances or within the construction industry. It is important to note that reducing the physical and mental workload for certain products is not desirable and under stimulation can occur leading to a reduction in the overall UX. This tool does not address every element of UX but encompasses key components, to ensure that the overall UX is improved, attention must also be given to the context, and capability of the user.

References

[1] de Haas, M., et al., E-bike user groups and substitution effects: evidence from longitudinal travel data in the Netherlands. Transportation, 2022. **49**(3): p. 815-840.

[2] Stilo, L., et al., Electric bicycles, next generation low carbon transport systems: A survey. Transportation Research Interdisciplinary Perspectives, 2021. **10**: p. 100347.

[3] Bayer, C., E-MTB Battery Revolution? A realistic Assessment of the Situation. 2017.

[4] Hinderks, A., et al., Developing a UX KPI based on the user experience questionnaire. Computer Standards & Interfaces, 2019. **65**: p. 38-44.

[5] Fdis, I.S.I., Ergonomics of human–system interaction. 2022.

[6] Sauer, J., A. Sonderegger, and S. Schmutz, Usability, user experience and accessibility: towards an integrative model. Ergonomics, 2020. **63**(10): p. 1207-1220.

[7] Allam, A. and H.M. Dahlan, User experience: challenges and opportunities. Journal of Information Systems Research and Innovation, 2013. **3**(1): p. 28-36.

[8] Hu, H., et al., A quantitative aesthetic measurement method for product appearance design. Advanced Engineering Informatics, 2022. **53**: p. 101644.

[9] Desmet, P., K. Overbeeke, and S. Tax, Designing Products with Added Emotional Value: Development and Appllcation of an Approach for Research Through Design. The Design Journal, 2001. **4**: p. 32-47.

[10] Agarwal, A. and A. Meyer, Beyond usability: evaluating emotional response as an integral part of the user experience, in CHI'09 Extended Abstracts on Human Factors in Computing Systems. 2009. p. 2919-2930.

[11] Finstad, K., The usability metric for user experience. Interacting with Computers, 2010. **22**(5): p. 323.

[12] Lewis, J.R. and J. Sauro. The factor structure of the system usability scale. in International conference on human centered design. 2009. Springer.

[13] Young, M.S., et al., State of science: mental workload in ergonomics. Ergonomics, 2015. **58**(1): p. 1-17.

[14] Lee, W., Methods for measuring physical workload among commercial cleaners: A scoping review. International Journal of Industrial Ergonomics, 2022. **90**: p. 103319.

[15] Winkel, J. and S.E. Mathiassen, Assessment of physical work load in epidemiologic studies: concepts, issues and operational considerations. Ergonomics, 1994. **37**(6): p. 979-988.

[16] van der Beek, A.J. and M.H. Frings-Dresen, Assessment of mechanical exposure in ergonomic epidemiology. Occupational and Environmental Medicine, 1998. **55**(5): p. 291.

[17] Miller, S., Workload Measures. 2001, University of Iowa p. 65.

[18] Plass, J.L., R. Moreno, and R. Brünken, Cognitive load theory. 2010.

[19] Naismith, L.M., et al., Chapter 10 - Using Cognitive Load Theory to Optimize Simulation Design, in Clinical Simulation (Second Edition), G. Chiniara, Editor. 2019, Academic Press. p. 129-141.