

HUMBOLDT-UNIVERSITÄT ZU BERLIN
MATHEMATISCHE-NATURWISSENSCHAFTLICHE FAKULTÄT
INSTITUT FÜR INFORMATIK
HUMAN-COMPUTER INTERACTION

LikertShift - An Input Device for Recording Cycling Subjective Experiences

Bachelorarbeit

zur Erlangung des akademischen Grades
Bachelor of Science (B. Sc.)

eingereicht von: Max Schlecht
geboren am: 08.06.2000
geboren in: Berlin

Gutachter/innen: Prof. Dr. Thomas Kosch
Prof. Dr. Andrii Matviienko

eingereicht am: 25.01.2025 verteidigt am: 13.08.2025

Abstract

This work explores creating an improved method for recording real-time ground truth data on cyclists' subjective experiences (CSE). It presents the design process of a physical prototype, resembling a standard bicycle twist gearshift, which enables cyclists to provide ratings on a Likert scale without interrupting their cycling or compromising their safety. A field study was conducted to evaluate the prototype against other state-of-the-art methods. Results showed that participants strongly preferred using our prototype, highlighting its ease of use, intuitive handling, and lack of concerns regarding social acceptance or additional time requirements compared to other methods. Our evaluation of recorded travel satisfaction data showed little evidence for quantitative differences between the compared methods. This research paves the way for larger field studies by providing a better method to either directly record CSE or the necessary ground truth data to infer them from additional contextual and sensory data.

Contents

1 Motivation	1
2 Related Work	3
2.1 Methods for Recording CSE	3
2.2 Digital Control Devices for Bicycles	5
2.3 Research Questions	6
3 Prototype Design	7
3.1 Design Requirements	7
3.2 Mechanical Design	7
3.3 Electronics	11
3.4 Frontend Design	13
4 Study Design	16
4.1 Recording Methods	16
4.2 Route Selection	17
4.3 Data Collection	18
4.4 Procedure	19
5 Results	21
5.1 Participants	21
5.2 Route Data	21
5.3 Questionnaires and Interviews	22
6 Discussion	24
6.1 Research Questions	24
6.2 Future Work	24
7 Conclusion	26
References	i

1 Motivation

Recent works in the HCI space have been exploring detecting drivers' emotions while operating vehicles and understanding their effects. Many of them specifically focus on negative emotional extremes, such as anger and sadness [39], as these not only lead to an increased number of driving errors [15] but can also encourage more risky behaviors like speeding [14, 27].

Consequently, there is also a growing interest in developing affective user interfaces that can utilize said emotions to enhance user satisfaction and traffic safety. Examples of such affective user interfaces include route planners that do not only offer the fastest routes but also ones that evoke more positive emotions in drivers [4], voice assistants that adapt their tone empathically [1, 5], or even systems that prevent operation altogether if the driver is deemed incapable of operating the vehicle [5]. *Zepf et al.* [38] also investigated what situations during driving trigger certain emotions and categorized them, recommending measures to mitigate those that negatively impact driving performance.

Many of these studies, particularly real-world ones, rely on real-time audio-based self-reporting to obtain their results [3, 27, 38]. In an automotive context, this method is practical, as microphones can easily be mounted inside vehicles, which provide an environment with a generally well-predictable noise floor. Automatic voice processing can then be used to obtain the desired data [3].

While most of this research has focused on travel by car, the ongoing shift toward sustainable transportation, especially in urban areas, has generated significant interest in expanding it to other modes of individual transportation, most notably cycling.

Similar to automotive contexts, data on CSE can be used to improve traffic safety, infrastructure planning [29, 37], as well as the overall experience and satisfaction of cyclists, thereby engaging more people in a more sustainable and healthy mode of transportation [7, 22, 30].

Personalized route planning, for example, gains even more significance in the context of cycling, as there is not only a larger variance in road conditions and types of bicycle paths but also in cyclists' preferences concerning them. For instance, some cyclists may prefer riding only on well-maintained, flat asphalt roads, while others might actually enjoy more rugged, off-road paths. To properly detect and differentiate between triggers (e.g., weather, road condition, traffic) affecting the subjective experience of cyclists while riding, subjective real-time ratings must be consolidated with accurate time and location data, as well as external contextual data [40].

Multiple recent large reviews have found that existing studies predominantly rely on retrospective surveys and interviews as methods to evaluate CSE [17, 20, 40]. *Zhang et al.* [40] also claim that recently, there has been a growing interest in utilizing more mobile methods capable of capturing real-time data during cycling.

Methods that try to predict cyclists' emotions from contextual information such as driving behavior, weather, or traffic could also provide such real-time data. *Bethge et al.* [3] demonstrated that emotion recognition models based on contextual data can even outperform traditional techniques like facial expression recognition. However, creating

mathematical models or training machine learning models like this still requires obtaining ground-truth data from study participants in the first place, further underlining the need for developing better self-reporting methods [3, 7, 12, 36].

We argue that data obtained from retrospective methods often lacks the required detail [40], while live recording methods, namely mobile audio self-reporting, have been directly adopted from automotive research on the subject but are unsuitable for cycling due to social and environmental factors. Thus, to advance this field and enable researchers to collect real-time ground-truth data on CSE more effectively, we want to develop a new method using an electromechanical device that allows cyclists to self-report values on a simple one-dimensional scale.

We start by reviewing related work on existing real-time methods for recording cyclists' subjective experiences and digital control devices for cycling. From this, we derive a set of design requirements that guide the development of our prototype. Finally, we conduct a field study comparing our prototype to existing real-time recording methods, evaluate its performance, and discuss our results and how they could affect future research in the field.

2 Related Work

2.1 Methods for Recording CSE

In this section we provide an overview of methods used to record CSE. We focus on methods that allow fine-grained recording of CSE on subsections of a longer route, so either live-recording methods or retrospective methods that utilize a high-level of detail. Research using such methods largely focuses on measuring risk/safety, especially during passing events and general stress/comfort during cycling. There is a great focus on inferring these metrics from biosensory data and individual driving behavior.

2.1.1 Audio Recording

McIlroy et al. [25] and *Jones* [16] took a similar approach by utilizing audio diaries, letting participants freely reflect their thoughts and emotions by “thinking out loud” during their ride. *McIlroy et al.* utilized this method to not only record data of cyclists but also car drivers and motorcyclists, making all of them navigate the same, predefined route, to gain an understanding of affective differences between the groups. Through thorough coding of the recorded data and an elaborate analysis, they found that cyclists’ exhibited increased emotional arousal, compared to the other groups, which was highly influenced by the road surface quality. *Jones*, in contrast, took a more natural approach, asking participants to record their diaries during their commute, while also recording GPS locations. *Jones* then used this data to compare cyclists’ sensory load depending on the space they traversed.

Other works using speech-based methods, include ones by *Yamanaka et al.* [36] and *Rybarczyk et al.* [32], who both used mobile audio recordings as a baseline to develop models that infer cyclists’ comfort levels from sensory data. Both works used sensors attached to the bicycle to record speed, acceleration, and vibration data to develop classical regression models. *Yamanaka et al.* also recorded braking and steering behavior, while *Rybarczyk et al.* relied on additional biosensory heart rate data. The two of them also performed similar evaluations, subdividing routes into smaller segments to link discomfort with specific sections, like slopes, intersections or unclear pathways. *Yamanaka et al.* required participants to perform five numerical ratings, on different measures, per segment, while *Rybarczyk et al.* only instructed participants to state when they feel particularly uncomfortable.

2.1.2 Interceptive Interviews

Graystone et al. [10] performed a number of interceptive surveys, stopping cyclists’ in traffic, most of them while commuting and asking them to fill in a short questionnaire. They used their gathered data to investigate the effect of self-identified gender on cyclists’ safety perceptions and found that cyclists who identify as women reported significantly higher concerns regarding verbal abuse from car drivers.

Using a similar methodology, *Kazemzadeh et al.* [18] compared in-traffic interceptive surveys with retrospective, online ones. Their conclusion was that while both methods

produced similar results on a spatially broader level, cyclists' perceptions towards the specific spots they were surveyed in varied drastically, with online surveys being answered more positively and optimistically. They speculate that this could be because interceptive in-traffic questionnaires apply a certain pressure on cyclists, highlighting the effects of different data acquisition methods on CSE.

2.1.3 Mapping

Manton et al. [23] used a retrospective mental mapping approach to compare cyclists' subjective risk assessments to an objectively performed risk analysis, performed by evaluating the transport infrastructure. Participants had to color code segments of a previously driven, mentally mapped cycling route on a paper map, according to their perceived risk. When comparing this data against the objective one, *Manton et al.* found, only partial alignment between the two datasets, highlighting the difference between perception and reality. They conclude, that their mapping method is highly useful for obtaining risk assessment data, highlighting its potential for participation by the general public.

Meenar et al. [26] went a step further and used a similar method to identify and map arbitrary emotions of cyclists. In contrast to the work by *Manton et al.*, they did not provide participants with a printed map, but instead relied on them sketching out their route with added commentary on specific section. These sketches were then combined to create a color coded, emotional map, summarizing the experiences of multiple participants and evaluated by transportation planners, who confirmed their usefulness. This work highlights that recorded data on cyclists' emotions, with a high degree of spatial detail, can be used to improve transportation planning and overall traffic safety.

2.1.4 Physical Devices

Works by *Dozza et al.* [9] and *Beck et al.* [2] used a digital push-button, mounted to the handlebars to enable cyclists to record critical incidents during their ride, providing more detailed information on them at a later point in time. Both studies were performed in a very naturalistic setting, as they let participants use their modified bicycles for a period of up to two weeks, only analyzing the recorded incidents afterward. *Dozza et al.* used this data to compare incidents of traditional bicycles and electric bicycles and their causes. *Beck et al.* on the other hand specifically focused on analyzing cyclists' perceived risk induced by passing vehicles and the effect the passing distance has on it. Various works by *López et al.* [21, 31] also investigated CSE during overtaking maneuvers, but instead of using a single button to signal high-risk situations, they built a custom device with five buttons, that allow cyclists' to submit a risk-rating on a five point scale. They then used this data in combination with measured passing distances to construct a regression model and discuss its usefulness for automated driving applications or improved safety systems.

2.2 Digital Control Devices for Bicycles

Next, we look at previous works on the development of digital control devices for bicycles, to gather key takeaways on our physical prototype design to optimize it for safety and usability.

2.2.1 Physical Control Devices

To explore new paradigms for smartphone control during cycling, *Hochleitner et al.* [13] conducted a study, letting participants explore movements for smartphone control, without restrictions, on a bicycle trainer. They observed that participants highly preferred keeping their hands on the handlebars at all times, while only executing small movements with their fingers. Based on these results, they conducted a second field-study, comparing different three smartphone interaction methods: using physical buttons on the handlebars, a wristband that detects quick twists of the wrist and directly interacting with the smartphones' touchscreen. This confirmed their previous findings, as participants exhibited significantly lower physical demand and frustration when using the physical buttons, in comparison to the other methods.

These findings match with observations by *Woźniak et al.* [35], who developed an evaluated two different physical control devices for smartphone control. Both were attached to the handlebars, next to the grip-area, but one presented users with a simple three button interface, while the other one largely relied on rotation control, featuring only a single button. They evaluated their two prototypes against each other in a field-study and found that the rotation based device allowed participants to maintain significantly more lateral control of their bicycles, compared to the one using buttons, as it did not require them to change their grip position while operating.

Further, *Matvienko et al.* [24] investigated text input methods that could be used during cycling, comparing one-handed smartphone use with midair gestures and a set of ten physical buttons below the handlebars in a simulated, indoor environment. They noted that while the physical buttons, exerted a higher mental load on participants, they allowed participants to keep their hands on the handlebars, making the physical buttons overall less distracting in comparison to the other methods.

2.2.2 Gesture Based Controls

Finally, we briefly looked into control methods that rely solely on gestures. An analysis conducted by *Caon et al.* [6], similar to the previously discussed work by *Hochleitner et al.*, investigated finding gestures that could be performed during cycling, suitable for smartphone control. They focused on comparing the frequency of fingers used for said gestures and categorized different kinds movements to conclude that the most convenient gestures seem to be pressing motions with the thumb and index finger. It should be noted that this study was conducted using road bikes though, which require a very different grip position from ordinary city bikes, so its unclear whether their findings are applicable to general bicycle control.

In contrast to this theoretical approach, only observing possible gestures without actually implementing an interface that utilizes them, *Tan et al.* [34] built a wearable, glove like prototype, able to detect such gestures. While using their prototype resulted in a similar response time, compared to using physical input devices, they only observed a 92% accuracy of gesture recognition, which could lead to increased frustration.

2.3 Research Questions

While researchers have utilized many methods to record and classify CSE, little to no effort has been made to compare said methods against each other. Furthermore, the effects that these methods have on the very experience they are meant to record, have largely been disregarded. We believe that employing, intuitive, easy-to-use, physical devices to record CSE, would minimize these effects. Thus, to enable further research into this topic, we want to design an electromechanical device, which mimics the functionality of previous devices (such as the one by *López et al.*) while optimizing its safety and usability by employing findings from the field of Digital Control Devices for Bicycles.

In this thesis we describe our design process creating this device and evaluate it in a field-study to answer the following research questions:

- Q1: How do different methods to record cyclists' subjective experiences affect their rating behavior?
- Q2: What are cyclists' preferences regarding methods to record their subjective experiences?

3 Prototype Design

This section describes our process designing the electromechanical prototype. We outline a set of design requirements to adhere to, then describe our design decisions and implementation. We also provide an overview of our electrical design and required firmware, as well as the frontend we created for it.

3.1 Design Requirements

To streamline the design of our prototype and to ensure its usefulness to other researchers in future work, we established the following design requirements:

DRQ1 Intuitive

Controlling the prototype device should be intuitive to the user.

DRQ2 Robust

Using the prototype device should require minimal user intervention and maintenance. Users should not be able to easily break the device, and it should withstand external influences such as humidity and temperature variations.

DRQ3 Safe

The prototype device should not compromise the user's safety.

DRQ4 Affordable

The total cost of all required components should not exceed €25.00.

DRQ5 Easy to Reproduce

The prototype device should be able to be manufactured and assembled using standard tools commonly available in most research laboratories.

3.2 Mechanical Design

3.2.1 Initial Design Decisions

Mounting

Previous research on input interfaces that can be used while cycling showed that participants strongly preferred solutions that allowed them to keep both hands on the handlebars [6, 13, 24]. Studies on smartphone usage while cycling also confirmed that even removing one hand from the handlebars increases risk, especially when performing actions that require increased mental effort [8].

Thus, we determined that the device should be mounted directly to the handlebars, close to the natural grip area of the bicycle.

Off-the-shelf or Custom Built

The primary advantage of using off-the-shelf electromechanical components for the mechanism of the device is reliability - most components specify a life expectancy in their respective datasheets - which supports DRQ2. However, this conflicts with DRQ4 and DRQ5, as standard versions of the required components (potentiometers, multipole switches or digital encoders) are challenging to fit into the available space and more specialized variants of these components are often difficult and expensive to source. As an example, the center-space potentiometer used by Woźniak *et al.* [35] in their “Brotate” control device would not only take up $\frac{1}{4}$ of our budget set in DRQ4, but has also been marked as “Not Recommended for New Designs”¹ by the manufacturer, making its future availability uncertain.

To ensure building and using our device is both affordable and accessible, we decided to rely solely on widely available standard mechanical components (e.g., screws, nuts, bearing balls, springs) and to use Fused Deposition Modeling (FDM) 3D printing to create large parts of the mechanism. FDM 3D printing is the most widely used additive manufacturing technique and enables us to quickly iterate on our designs while producing low-cost parts that can be easily reproduced [28].

Note: We also briefly explored utilizing existing commercially available interfaces, such as electronic gearshifts or electronic throttles for e-bikes, but ultimately decided against using such an off-the-shelf solution as they do not meet our design requirements.

Continuous or Discrete Scale

As proposed in Section 1, our prototype should allow cyclists to communicate their subjective experiences by providing a rating on a simple one-dimensional scale. Theoretically, a continuous scale could capture more precise data than a discrete one. However, since our use case is recording subjective ratings, recorded values are likely to vary significantly between users, so to obtain meaningful results, any recorded data must be averaged across a reasonably large pool of users. Therefore, a continuous input scale would not provide a substantial benefit for our specific use case. This decision also directly influences the electromechanical design of the device. Choosing to design a continuous scale would force us to utilize components such as variable resistors (potentiometers) or high-resolution digital encoders. In contrast, a discrete scale would suggest the use of rotary switches or multiple mechanical buttons.

Due to the lack of benefit from a continuous scale and our previous decision to use a custom-built mechanism, we decided on a discrete scale, as this is easier to implement using our chosen methods.

3.2.2 Implementation

Based on our previous design decisions and set requirements, we developed a prototype, as shown in Figure 1, resembling a standard bicycle twist gearshift. It can be mounted

¹ <https://www.digikey.com/en/products/detail/EWV-YG9U04B14/3163176>

directly on the handlebars in combination with a shortened bicycle grip, allowing users to control the device by twisting it without removing their hands from the handlebars. Users can rate their subjective experience during cycling on a scale with multiple discrete positions that the device snaps into. We 3D printed all custom parts on a Voron 2.4 3D printer out of PETG filament, but any decently calibrated consumer FDM 3D printer should be able to produce them.

To perform our initial evaluation, we decided to use a simple five-point Likert scale, with each position corresponding to the level of the cyclists' current travel satisfaction. Based on this we also decided on a name for our prototype: **LikertShift**.

Note: We choose PETG because it is a common, easy-to-print material, but has a higher glass transition temperature (T_g) than PLA. For use in the summer or warmer climates, we recommend using an engineering material with an even higher T_g like ABS or Nylon.



(a) CAD model (b) Fully assembled (c) Mounted to the handlebars

Figure 1: The final LikertShift prototype

To mount the prototype, the device is split in half and utilizes a simple clamping mechanism, similar to ones used for commercial bike accessories, to fix it to the handlebars. We tried to add some texture to the clamping surface (see Figure 2), to reduce the necessary clamping pressure, but had mixed results with this. The best solution to stop unwanted rotations and movements has been to apply a piece of masking tape to the handlebars at the desired mounting location.

We modeled a cavity to allow the application of a label sticker to mark the value of each discrete step and also added small markers to indicate the position that is currently selected. The markers are 3D printed with a different filament color to improve visibility. We also used a combination of CAD modeled indents, as well as the so-called “Fuzzy skin” feature in our slicer² to add some texture to the grip area (see Figure 3). This could be improved by utilizing multi-material 3D printing techniques and printing parts of it in more “grippy” materials like TPU, but we decided against such an approach, because multi-material 3D printing either requires special multi-toolhead printers or generates a lot of wasted material.

² https://help.prusa3d.com/article/fuzzy-skin_246186

To satisfy DRQ2 and DRQ5 we designed the device to be fully integrated, without requiring any external power. Figure 4 shows the electronics compartment which houses the main Microcontroller board, as well as a button-cell battery to supply power to the electronics (see Section 3.3).

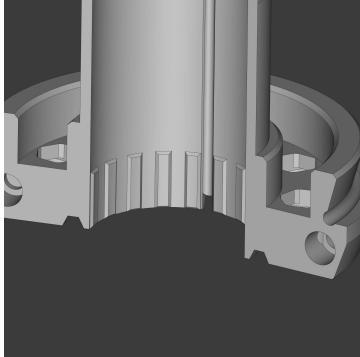


Figure 2: Clamping surface

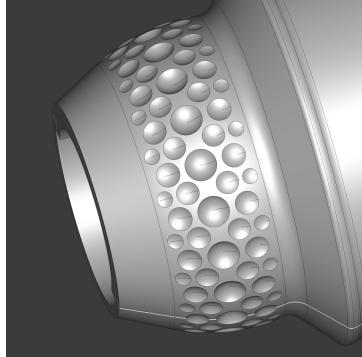


Figure 3: Grip area

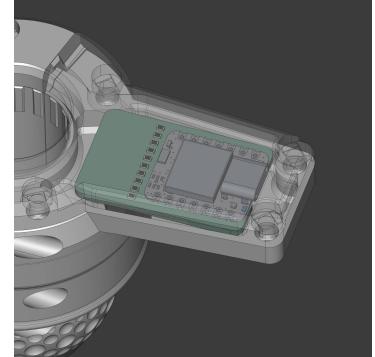


Figure 4: PCB compartment

We used FreeCAD³ (a parametric modelling software) in combination with OpenSCAD⁴ (a code-based solid modelling software) for the mechanical design of our prototype. Both programs qualify as free and open-source software (FOSS), ensuring the design is accessible and modifiable by future contributors without requiring expensive software licenses. The design was created parametrically, allowing for parameters like the number of discrete positions or their spacing to be easily adjusted. All parts have been designed specifically for FDM 3D printing, with a particular print orientation in mind and don't require any additional support structures when printing.

3.2.3 Mechanism

The prototype uses a simple rotary switching mechanism. When the outer Rotator gets rotated, it pulls the Rotator-head along a ramped track (see Figure 5). The Rotator-head, shown in Figure 6, consists of a 4 mm steel bearing ball that is socketed into a 3D printed part and gets pushed into the track by a spring. It is constrained to only move along the rotation axis by ~ 5 mm to be able to overcome the ramps. This way, the mechanism automatically snaps into the depressions between the ramps.

At the bottom of the depressions are nickel-plated neodymium magnets which are being used as the switching contacts while the Rotator-head acts as the contactor. M3 button head screws are used to clamp and connect cables to the contactor and switching contacts. As the cable that goes to the Rotator-head experiences lots of movement and thus relatively high stress, we used special FEP insulated wire⁵ made for the use in cable chains for it. For the contacts we used standard PVC insulated wire. This setup results in a total contact resistance of $\sim 10\Omega$, which is well within useable range for our purposes.

³ <https://www.freecad.org/>

⁴ <https://openscad.org/>

⁵ <https://shop.helukabel.com/de-en/heluflon-fep-6y-m25511/20271>

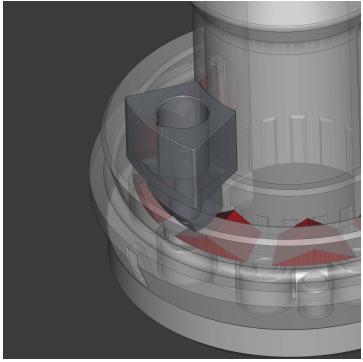


Figure 5: Ramped track

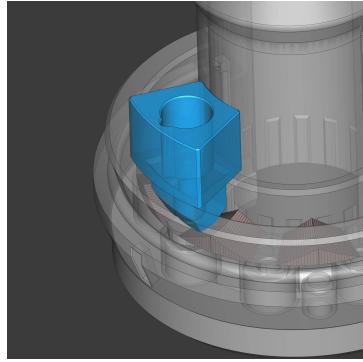


Figure 6: Rotator-head



Figure 7: Assembled PCB

3.3 Electronics

This section describes our electronics design, as well as microcontroller choice and power requirements. For retrieving data from the prototype to our frontend (see Section 3.4), we decided to use the Bluetooth Low Energy (BLE) network, as it offers low-power operation, is supported by most modern smartphones and does not require stationary access points to function. We also only need very short-range communication capabilities, as the user is always situated right next to our device.

3.3.1 Microcontroller Choice

For data acquisition and communication we wanted to use the nRF52 series Microcontroller Units (MCUs) because of their exceptionally low-power usage, Bluetooth Low Energy (BLE) capabilities and wide availability⁶. Specifically, for our prototype we decided to use a relatively affordable (DRQ4) nRF52840 based development board by Seeed Studio⁷. Using a development board allowed us to skip many electrical design challenges and to test the working principle of our prototype on a solderless breadboard.

3.3.2 Power

To satisfy DRQ2 and DRQ5 we designed our device to be fully integrated and enclosed. Due to the given space-constraints by our mechanical design it is only powered by a standard CR2032 button cell battery. Under optimal conditions, it should be able to power our device for at least 1–2 years. We did not optimize our firmware to achieve this kind of power usage though (see 3.3.4).

3.3.3 Printed Circuit Board (PCB)

To minimize wiring and to keep the electronics tightly integrated, we designed a simple breakout-PCB (see Figure 7) that holds the MCU development board and a battery

⁶ <https://docs.nordicsemi.com/category/nrf-52-series>

⁷ https://wiki.seeedstudio.com/XIAO_BLE/

holder and exposes solder pads that allow easy connection of the contactors described in 3.2.3. The solder pads allow connecting of up to ten such contactors, so it can be used with variants of the device with up to nine discrete positions. The solder pads connect directly to the MCUs' General-Purpose Input/Output (GPIO) pins, as well as a high-impedance pull-down resistors to hold the logic state to 0V when not in use.

The two-layer PCB design was implemented using the KiCad Electronics Design Automation Suite (KiCad EDA) and is optimized for affordable manufacturing with design rules allowing for comparably high production tolerances. The KiCad EDA also qualifies as FOSS.

3.3.4 Firmware

When powered up, the prototype starts to advertise itself via BLE, broadcasting its name and that it is ready to accept connections. Once a central device, the smartphone in our case, initiates a connection, the prototype hosts a Generic Attribute (GATT) server providing two services, a custom **LikertShift** service to retrieve the currently selected switch position and a battery service to retrieve the current battery level. Both services each offer a one-byte characteristic that can be read by the central device.

When the **LikertShift** characteristic gets read, first the GPIO pin connected to the contactor at the **Rotator-head** gets set to a HIGH state (~ 3.3 V) then each of the pins connected to the switching contacts gets read, checking for a matching HIGH state. Finally, the detected position gets sent back to the central device, as a value from $1 - n_{positions}$. There are two special cases here:

- (i) No switching contact is set to a HIGH state.

This can happen if the service gets read while the prototype is currently being moved from one position to another. In this case the service will return a value of 0. How exactly this state should be handled is left to the software running of the central device.

- (ii) Multiple switching contacts are set to a HIGH state.

This should not happen during normal operation. In this case the service returns the position corresponding to the last GPIO pin read.

When the battery characteristic gets read, the battery voltage is measured by the MCUs onboard Analog to Digital Converter (ADC). Then, the measured voltage gets linearized and scaled to a value from 0 – 255 and sent back to the central device.

The firmware is written in Rust⁸ using Embassy⁹ (a framework for embedded applications).

⁸ <https://www.rust-lang.org/>

⁹ <https://embassy.dev/>

3.4 Frontend Design

To view and record ratings using the **LikertShift** we need a frontend that can connect to it via BLE and talk to our firmware to retrieve the currently selected position. Because we also want to record location data simultaneously and the **LikertShift** hardware does not possess GPS recording capabilities, we decided to implement this frontend as a smartphone app. This way, we can also extend it to query other information, relevant for our evaluation, from study participants.

3.4.1 Features

The base implementation of the app allows users to:

- connect to nearby **LikertShift** devices and display their currently selected rating.
- record GPS locations and **LikertShift** ratings along arbitrary travel routes.
- view their current location and travelled route on a 2D map.

We later extended the app for performing our own field-study. We took great care to make it as customizable and reusable as possible, so other researchers could reuse it to replicate our results or perform their own studies. Thus, we added the ability for researchers to:

- inquire basic demographic information from study participants.
- predefine routes that participants should follow along and display them in the map view in addition to the actual travelled route.
- add custom surveys that appear when participants finish recording a route.
- record audio interviews.
- bundle up and export all recorded data to a single ZIP archive and reset the app back to a default state.

Surveys and routes can be defined in a human-readable JSON format, so minimal programming knowledge is required to customize the app. For evaluating the **LikertShift** in our own field-study we also added the ability to use the other recording methods described in Section 4.1.

3.4.2 Implementation Details

The app is divided into four main screens which can be switched between by using the bottom bar. When it gets started for the first time, the user is presented with a demographics form to fill-in (see Figure 9a). All form and survey data gets saved locally in a JSON format.

Map Screen

The Map Screen displays the users current location and surroundings on a 2D map (see Figure 8a). It can be moved and rotated using standard touch gestures. The user can

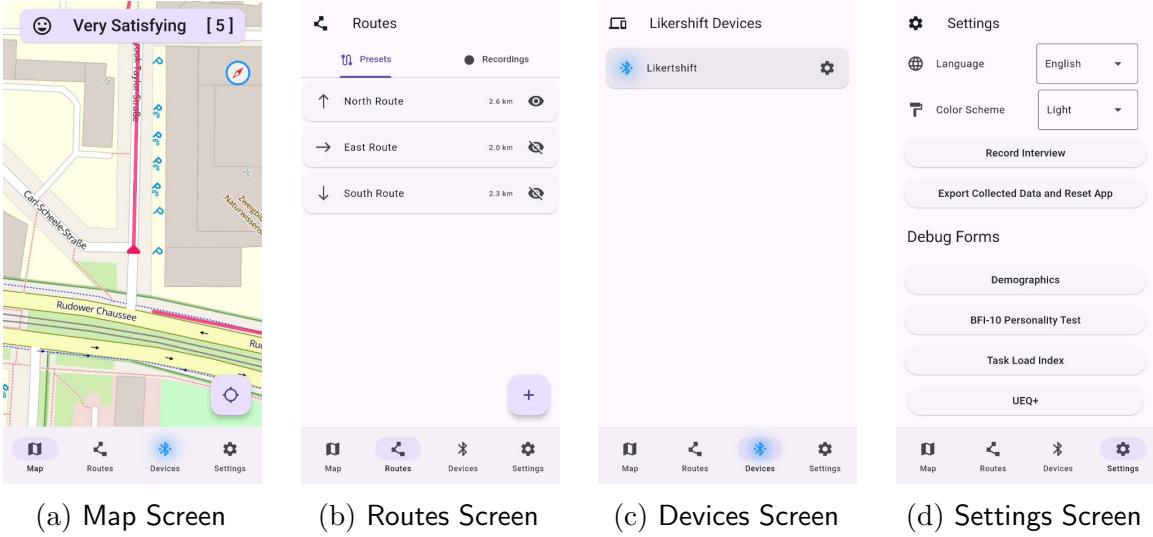


Figure 8: An overview of the four main app pages

press the bottom-right toggle to make the map automatically follow the current location and long-press the compass to lock the rotation to the current movement direction. If the user starts a recording, the **Map Screen** will also display the travelled route, as well as the preset route, if one was selected. Additionally, if a **LikertShift** device is connected, the **Map Screen** will also display the devices' currently selected numeric rating and a description of its meaning. Routes from the **Routes Screen** can also be previewed on the **Map Screen**.

Routes Screen and Route Recording

The **Routes Screen** is divided into two tabs, showing the available route presets and past recordings respectively (see Figure 8b). Route presets can be previewed using the “Eye” buttons and new recordings can be started using the + button.

When starting a new recording the user has to select a recording method and optionally a route preset. Then, whenever the app receives a location update (every $\sim 1\text{--}5\text{ s}$), a new data point will be added to the recording, containing the new location, the time elapsed since starting the recording and the current rating value by the active **LikertShift** device, if available. This data is then immediately written to a new line in a CSV file, to prevent data loss in case of app crashes or battery run out. A recording can be stopped at any time by pressing the “Stop Recording” button that will appear above the main navigation bar. Once a recording is finished, the user can be presented with various forms and surveys (see Figure 9b for an example we used in our field-study).

Devices Screen

The **Devices Screen** allows discovering of and connecting to nearby **LikertShift** devices (see Figure 8c). It displays a list of currently advertising and already known devices. A

new connection can be established by selecting one of the devices from the list. Only one device can be connected to at any time. If a connection gets established successfully, its Bluetooth icon lights up, signaling that it's currently connected. During connection, the LikertShift device is queried for its currently selected rating value every 500 ms (this update rate can be easily adjusted). Invalid ratings (see Section 3.3.4) are ignored, so the previous rating is kept.

Settings Screen

The apps' language and color theme can be adjusted in the **Settings Screen** (see Figure 8d). It has been fully translated to English and German for use in our field-study. Additionally, it allows navigating to the **Interview Recording Screen**, which offers a simple interface to start, pause and stop recording an audio interview (see Figure 9c). Recorded audio interviews will also be saved locally as losslessly encoded FLAC files. Finally, the **Settings Screen** offers a button to “Export collected Data and Reset the App” which will bundle all recorded data into a single ZIP file and store it locally, under a randomly generated participant ID. Said ID will then be reported back to the user, and the app will be reset (see Figure 9d).

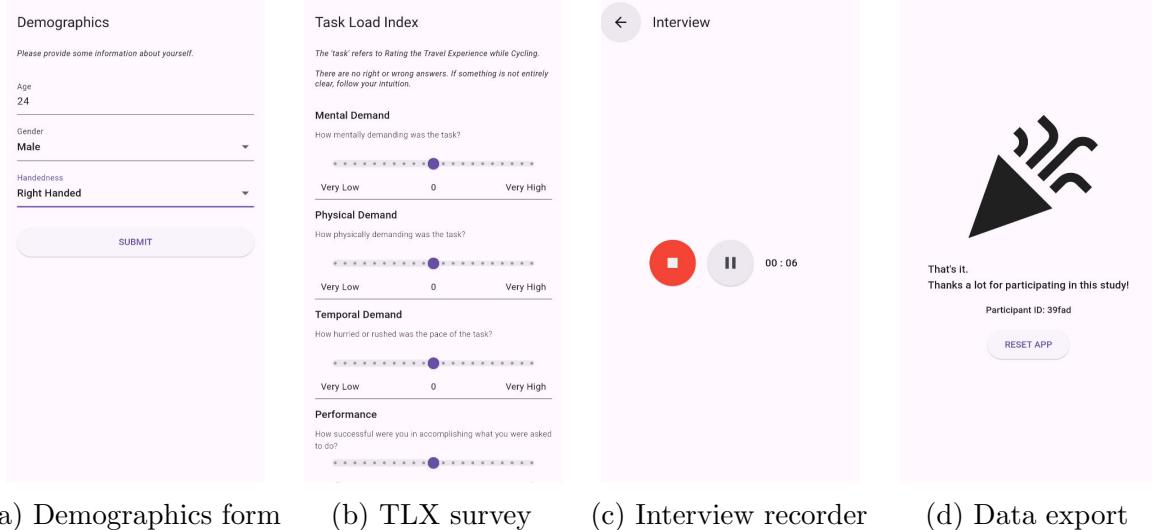


Figure 9: Contextual screens

The app is written using the Flutter¹⁰ framework and Dart¹¹ programming language. While we only tested and used our app on Android devices, the multi-platform nature of Flutter should also allow it to run on iOS powered devices with minimal modifications.

¹⁰ <https://flutter.dev/>

¹¹ <https://dart.dev/>

4 Study Design

We decided to evaluate our prototype in a semi-naturalistic field-study, comparing it against similar methods, capable of recording CSE for individual segments of a larger route. This section describes how we went about designing this study.

4.1 Recording Methods

Our method using the **LikertShift** device works by recording its currently selected value, whenever we receive a new GPS location. Thus, a selected value remains valid until the user selects a new one, and we receive their next GPS location, so a new rating gets recorded whenever they notice a change in their subjective experience and remains valid, onwards from that point in time, until they adjust the rating again. This allows users to choose a rating frequency they consider appropriate by themselves.

We selected two methods from our survey of the related work we regard as suitable to compare against our method and adjusted them, to allow for a fair comparison. The first one, referred to as **Audio Recording** from now on, is taken from the work of *Yamanaka et al.* [36], who captured audio data while cycling and required participants to rate pre-defined route-segments using multiple measures on scales from 1 to 5. We adjusted it to only use one measure and also let participants perform ratings on arbitrary route segments, in contrast to pre-defined ones. The second method, referred to as **Mapping** is the one proposed by *Manton et al.* [23], who let participants color code segments of a previously driven route, based on their risk assessment. In our study, we provided participants with a printed map, picturing the route they drove and let them divide it into multiple segments themselves, assigning a numerical rating to each segment, instead of color coding it.

Choosing a Measure

We had to take great care to select a suitable measure, so we could compare the data collected by different participants and the different methods against each other. This is inherently difficult, as cyclists' subjective experiences are arguably very subjective and will therefore vary drastically from participant to participant.

As mentioned in the introduction to Section 2, most works rely on risk/safety and stress/comfort measures to quantify CSE. The advantage of using risk/safety would be its independence from many external factors and high dependence on the chosen route. Constructing a route with highly varying risk factors would be comparatively easy, but deliberately placing participants in high-risk environments would be extremely unethical, which led us to quickly dismiss that approach. Letting participants rate their comfort level or travel satisfaction comes with its own set of challenges though. Overall comfort level depends on a lot of factors, such as participants initial mood, their experience in riding a bicycle, but also external factors such as weather, traffic, or surrounding scenery of the route. To be able to properly compare our recorded datasets against each other, we needed to find a way to eliminate as many of these factors as possible.

As findings by *McIlroy et al.* [25] revealed the high impact of road surface quality on cyclists overall comfort and satisfaction with their travel, we decided to limit our measure to this, using the metric of “Travel satisfaction, based on the road” as the CSE measure that participants have to use to rate segments of the route. We carefully explained this metric to each participant before conducting the study, instructing them on what should (road condition, road type, available space, slope of the road, etc.) and should not (general mood, current traffic situation, route scenery, etc.) affect their ratings. Ratings were to be performed on a Likert scale from 1 to 5, denoting high dissatisfaction and high satisfaction, respectively. This measure should exert a similar mental demand on participants, as they still need to feel out their comfort level, while increasing data quality, allowing us to perform quantitative comparisons of the different methods used.

4.2 Route Selection

If each participant was to evaluate all three methods after each other on the same route, we would expect to see an increase in accuracy for subsequent route traversals, as by that point the route is already known. Thus, we constructed three different routes and balanced route orders, method orders as well as the number of routes per method in our study. We chose each route to take ~ 10 min to complete for an average cyclist and ensured it contained a multitude of different road types and crossings. As we wanted to conduct the study in a single session per participant, at our university campus, we were limited by the available traffic infrastructure in the immediate vicinity and also had to put all start and endpoints in the approximate same location.

Name	Total Length	Road Type					Number of Crossings
		Road	Bike Path	Mixed Path	Pedestrian Way	Other	
North	2589 m	191 m	480 m	1135 m	359 m	424 m	13
East	2007 m	853 m	465 m	173 m	181 m	335 m	9
South	2289 m	1152 m	677 m	116 m	47 m	297 m	9

Table 1: Overview of the study routes

Table 1 shows the total lengths of the routes, as well as an overview of the different road types and number of crossings per route. Each route is named according to the direction it starts towards. The study was conducted at the Campus Adlershof¹² of the Humboldt University of Berlin, located in a mixed-use area outside the city center. Notably, the roads participants had to drive on were mostly side-roads with very little traffic and therefore received a more positive rating than one might expect. We also opted to use existing cycling infrastructure wherever possible, only evading to drive on the road when no such infrastructure was available. Furthermore, mixed-paths refer to areas mainly intended to be used by pedestrians that are spacious enough to allow for cycling at low speeds.

¹² <https://www.hu-berlin.de/en/about/campus/adlershof>



Figure 10: Elevated path (wood planks)



Figure 11: Tram crossing

We also chose each route to contain at least one section that particularly stands out from the rest of the route. For the “North” route this was an elevated wooden path (Figure 10), as well as an awkward tram crossing (Figure 11). The “East” route contained the objectively worst terrain of our study, a very rugged field participants had to traverse (Figure 12). The “South” route also featured an “off-road” section, traversing a lawn next to a building, this one was noticeably more flat though (Figure 13). The “East” and “South” routes also both featured different sections blocked by the same construction site (Figure 14), where participants had to result to driving on pedestrian ways.



Figure 12: Rugged field



Figure 13: Flat lawn



Figure 14: Construction site

4.3 Data Collection

All data was collected using the **LikertShift** app acting as a frontend for our prototype, as described in Section 3.4 and the **Audio Recording** and **Mapping** methods were added to it. **Audio Recording** is implemented by using the smartphones internal microphone to record an audio track while cycling. The recorded audio will then be added to the recording data as a lossless FLAC file. For the **Mapping** method, we only recorded location and time information in the app. The actual ratings were performed retrospectively on a piece of paper. After resetting the app at the end of each study participation, we simply wrote down the participants ID on said piece of paper, to be able to match it with the recorded data later on. This is described in more detail in Section 5.2.1.

4.3.1 Questionnaires

In addition to this quantitative data, we also used two questionnaires, the *NASA TLX* [11] and the *UEQ+* [33] (a modular extension of the commonly used *UEQ* [19]) to obtain qualitative data, regarding differences between the recording methods used. Participants were asked to fill in these questionnaires right after they finished rating a route with each method. We made sure participants understood that both questionnaires only concerned the respective method used to rate their subjective experience during cycling, but not the act of cycling itself. For the *UEQ+*, we selected the *Attractiveness*, *Efficiency*, *Intuitive Use*, *Hardware Security* and *Social Acceptance* scales and slightly adjusted some questions wordings, as the *UEQ+* is usually used to evaluate the user experience of products, not methods.

4.3.2 Interview

We also conducted a short ~ 10 min interview with every participant, asking them about their frequency of and reasoning for cycling, a ranking of the three methods used to record their subjective experience, as well as problems that could occur and possible improvements for each of the methods. Additionally, we asked what kind of bicycles participants used privately and whether our devices would be able to be mounted to them in its current form, to assess the compatibility of our method with different types of bicycles.

4.4 Procedure

We started by giving an introduction to the study as well as the methods used and explained the general order of operations and approximate duration. We also provided participants with an informed consent sheet, outlining the procedure, associated risks, our data protection guidelines and information about compensation, which they had to sign. They then got introduced to our **LikertShift** app and had to fill in the initial demographics form. Next, participants were shown the bicycle used for performing the study (Figure 15) and then asked to adjust the saddle height to their liking and to drive along a short test route, to get used to it. Then, we randomly selected a row from our balancing sheet (containing three method/route combinations, with each method and route only appearing once). For each of these combinations, we first explained the method used to participants and asked them if they would like to perform another short test drive to simulate using said method. Once done, participants drove the respective route and after submitting their ratings, also filled in our prepared questionnaires in the **LikertShift** app. During the ride, a smartphone was mounted to the bicycle, in between the handlebars (Figure 16), running the **LikertShift** app which displayed the route participants had to follow and recorded the required data. After all three combinations were completed, we finished by carrying out our interview.



Figure 15: The bicycle used in the field-study



Figure 16: Smartphone holder

Safety Precautions

To ensure each participants' safety, the researcher conducting the study followed them on their routes, retaining a safe distance, with another bicycle. Participants were further told to not look at the route on the smartphone for extended periods of time and to focus their attention on the road in front of them and to stop and get in contact with the researcher if a problem occurred, but to otherwise ignore them. They were also told to only drive at a conservative, safe speed.

5 Results

5.1 Participants

We recruited a total of 18 participants for our study using posters we placed on the university campus in combination with snowball sampling. Participants age ranged from 18 to 59 years. Only 3 of the 18 participants identified as female, all others identified as male. Each participant received a compensation of €10.00 for their efforts. The study was conducted from November to December, but except for temperature differences and some light rain, general weather conditions were consistent between study sessions. Each session took about 90 minutes in total. We experienced no study cancellations, and all recorded data was usable.

5.2 Route Data

In this section, we describe the evaluation of the quantitative ratings performed by participants on the measure of “Travel satisfaction, based on the road” using the **LikertShift**, **Audio Recording**, and **Mapping** methods respectively, as described in Section 4.1 and Section 4.3. Figure 17 shows a visually appealing overview of all data collected using all three methods.

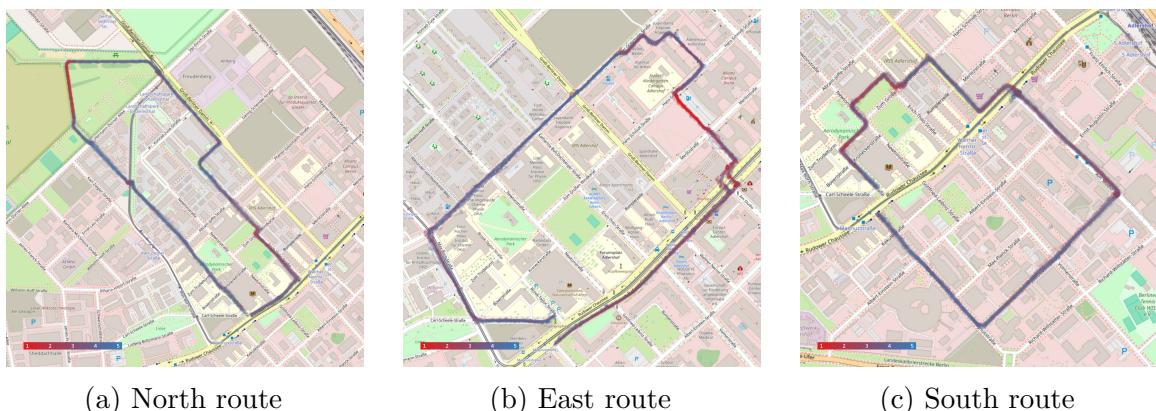


Figure 17: Combined ratings from all methods

5.2.1 Preprocessing

Preprocessing of the recorded data was largely performed by hand, with the help of some simple Python scripts. The start and end positions of all data recordings were adjusted manually, to remove unneeded data. For the **Audio Recording** method, we had to manually listen to each audio file and extract the ratings performed by participants to match them to the recorded timestamps and GPS locations. A similar process was used for the **Mapping** method to match the sections drawn on the paper maps to the recorded data. Our **LikertShift** method required no further preprocessing.

5.2.2 Evaluation

As mentioned in Section 4.1, comparing subjective data is inherently hard. To perform our analysis, we categorized our original reference routes by road type and resampled them to ~ 1 m segments. Then, we projected all recorded data from each method to them. This left us with a combination of rating and road-type for each 1 m segment of each recording. We then gathered all ratings by road-type for each method and computed their mean value, variance, and standard deviation, to see if we could see any drastic differences (see Table 2).

LikertShift	Road Type						
	Road	Bike Path	Mixed Path	Pedestrian Way	Wood Path	Field	Lawn
MEAN	3.7661	3.4624	3.7708	3.0176	1.8114	1.9804	2.8664
VARIANCE	0.9327	0.9140	0.8548	0.8873	0.9661	1.2512	1.3830
STDDEV	0.9658	0.9560	0.9246	0.9420	0.9829	1.1186	1.1760
Audio Recording	Road Type						
	Road	Bike Path	Mixed Path	Pedestrian Way	Wood Path	Field	Lawn
MEAN	3.6622	3.4978	3.7915	2.9559	2.5513	1.6895	2.4405
VARIANCE	0.8519	0.7970	0.8172	1.4158	0.5622	1.2402	0.9290
STDDEV	0.9230	0.8927	0.9040	1.1899	0.7498	1.1137	0.9638
Mapping	Road Type						
	Road	Bike Path	Mixed Path	Pedestrian Way	Wood Path	Field	Lawn
MEAN	3.7476	3.4952	3.7230	3.3553	2.3586	1.0000	3.0291
VARIANCE	0.9192	1.0542	1.0379	1.2773	1.5566	0.0000	1.0891
STDDEV	0.9587	1.0267	1.0188	1.1302	1.2476	0.0000	1.0436

Table 2: Evaluation of the recorded quantitative data

5.3 Questionnaires and Interviews

We evaluated each component of the *TLX* and *UEQ+* surveys separately, without computing an overall index, because we are most interested in what components cause increased task load or affect the user experience.

The *Physical Demand*, *Temporal Demand*, and *Performance* metrics of the *TLX* showed little statistical evidence for differences between methods. The most notable difference is the increased *Mental Demand* observed with the **Mapping** method, as well as higher *Frustration* and perceived *Effort* in comparison to the live-recording methods. There is insufficient statistical evidence for differences in task load between our **LikertShift** method and the **Audio Recording** method.

The evaluation of the *UEQ+* metrics turned out to be more interesting. We observed strong evidence for the *Attractiveness* and *Intuitiveness* of our **LikertShift** method, compared to the other ones. While participants perceived the **LikertShift** and **Audio Recording**

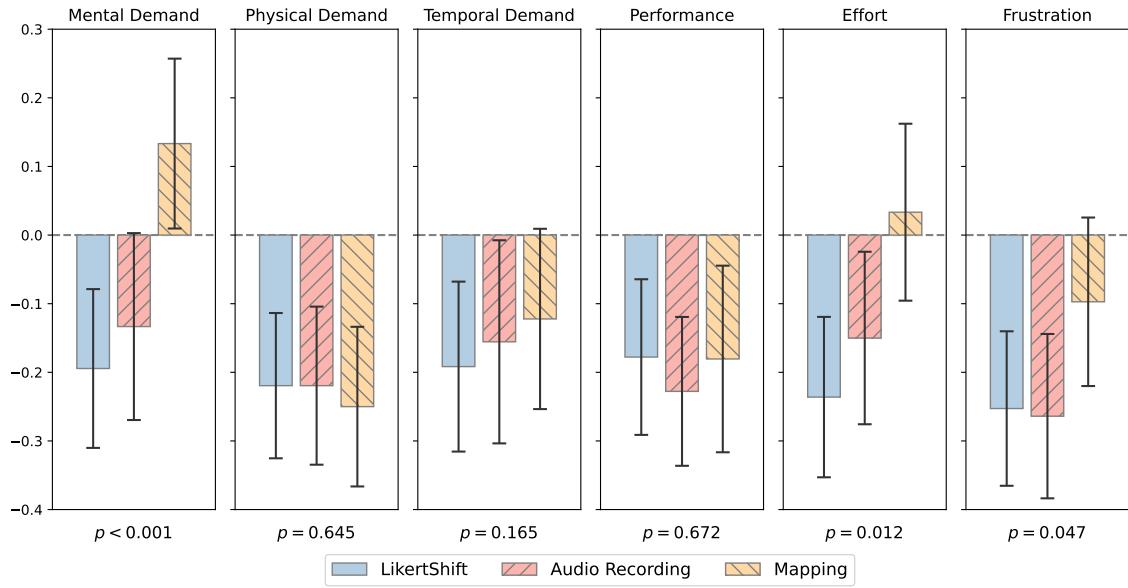


Figure 18: TLX results (error-bars denote the confidence interval based on a student’s distribution; p values are computed using a Friedman test; values are normalized from -1 to 1)

methods to be similarly *Efficient*, in comparison, the **Mapping** method was perceived as highly inefficient. We can also observe low *Social Acceptance* ratings of the **Audio Recording** method.

These findings were confirmed by our performed interviews. Multiple participants noted that the additional time requirements of the **Mapping** method make using it very unattractive to them. Similarly, the majority of participants stated they did not like to talk during cycling and found it “awkward” to perform the **Audio Recording** method. When asked what method participants would choose to conduct a longer (~ 2 weeks) field-study which would require them to record their travel satisfaction on all of their commutes and other trips, 17 out of 18 participants chose our **LikertShift** method.

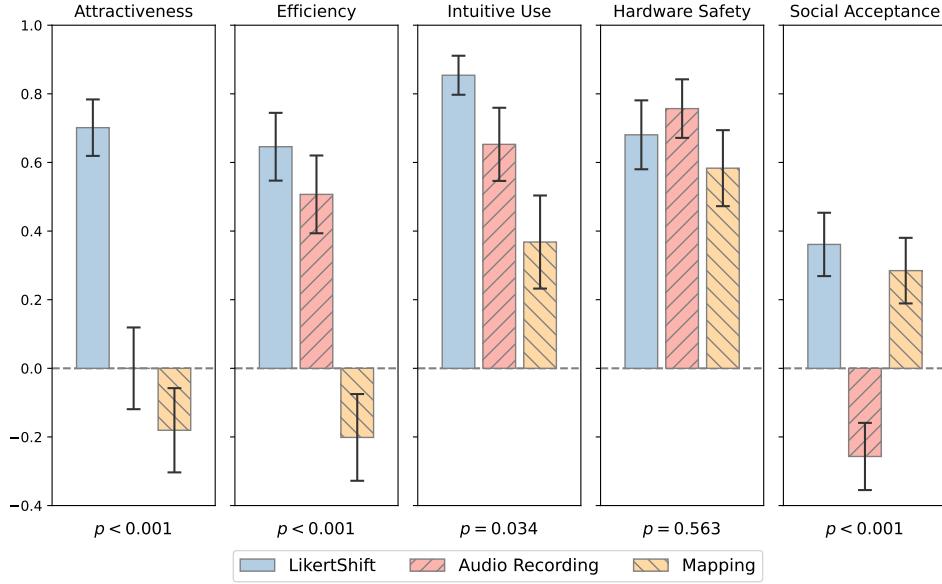


Figure 19: UEQ+ results (error-bars denote the confidence interval based on a student’s distribution; p values are computed using a Friedman test; values are normalized from -1 to 1)

6 Discussion

6.1 Research Questions

6.1.1 Q1: How do different methods to record cyclists’ subjective experiences affect their rating behavior?

We failed to answer this question, as the evaluation of the quantitative data we recorded did not allow us to perceive any large differences. Our best guess is that quantitative differences between methods are very subtle and require larger studies to be found.

6.1.2 Q2: What are cyclists’ preferences regarding methods to record their subjective experiences?

Users highly prefer methods that require minimal additional work and appear invisible to other traffic participants. While none of the methods we compared appear as particularly disturbing, especially to the users’ ability to stay in control and cycle safely, Audio Recording was perceived as unpleasant, while the Mapping method exerted a higher mental demand on them.

6.2 Future Work

Future work could include the combination of multiple methods for recording CSE, most notably hybrid methods, which rely on real-time data acquisition during cycling, but

allow for modification after data collection is complete. The easiest implementation of this would involve combining our **LikertShift** method with an adjustable map included in the smartphone app, similar to the **Mapping** method, but in a digital form. This would allow for corrections, in situations where users performed rating too late or forgot them altogether. Other improvements include the addition of active feedback to the device, for example, in the form of vibrations or sound. This would allow for an even more intuitive rating process. Finally, more sensors and connectivity options could be added to our **LikertShift** device, to make it usable standalone, without requiring a smartphone to actively record the data.

7 Conclusion

In this work, we described the successful design and construction of the **LikertShift**, a prototype device that can be used to record cyclists' subjective experiences, as well as its evaluation against other state-of-the-art methods in a semi-naturalistic field-study. Although quantitative analysis of recorded data revealed little to no statistical evidence supporting a measurable improvement in data quality using our developed **LikertShift** method, qualitative assessments of participants' opinions regarding the different methods provided valuable insights. They highlighted the perceived attractiveness, intuitiveness, and efficiency of using our physical device. Participants also voiced concerns about the social acceptance of speech recording during cycling and the high effort and increased mental demand associated with retrospective mental mapping approaches that rely on the memorization of route segments, particularly for longer routes. From our collected evidence, we see great potential in using physical devices for recording data on cyclists' subjective experiences, especially in longer, more naturalistic field-studies.

References

- [1] I. Alvarez, K. L.-D. Ipiña, S. B. Daily, and J. E. Gilbert. Emotional adaptive vehicle user interfaces : moderating negative effects of failed technology interactions while driving. In *Workshop 1: The Second Workshop on Automotive Natural User Interfaces*, 2012. URL: <https://www.auto-ui.org/12/adjunct-proceedings/w1-06-alvarez.pdf>.
- [2] B. Beck, M. Perkins, J. Olivier, D. Chong, and M. Johnson. Subjective experiences of bicyclists being passed by motor vehicles: The relationship to motor vehicle passing distance. *Accident Analysis and Prevention*, 155:106102, 2021. doi: 10.1016/j.aap.2021.106102.
- [3] D. Bethge, T. Kosch, T. Grosse-Puppendahl, L. L. Chuang, M. Kari, A. Jagaciak, and A. Schmidt. Vemotion: Using driving context for indirect emotion prediction in real-time. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, pages 638–651, New York, NY, USA, 2021. Association for Computing Machinery. doi:10.1145/3472749.3474775.
- [4] M. Braun, J. Li, F. Weber, B. Pfleging, A. Butz, and F. Alt. What if your car would care? exploring use cases for affective automotive user interfaces. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '20, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3379503.3403530.
- [5] M. Braun, F. Weber, and F. Alt. Affective automotive user interfaces - reviewing the state of driver affect research and emotion regulation in the car. *ACM Comput. Surv.*, 54(7), September 2021. doi:10.1145/3460938.
- [6] M. Caon, R. Süss, B. Grelier, O. A. Khaled, and E. Mugellini. Gesturing on the handlebar: A user-elicitation study for on-bike gestural interaction. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, and Y. Fujita, editors, *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, pages 429–439, Cham, 2019. Springer International Publishing. doi:10.1007/978-3-319-96071-5_46.
- [7] H. Dastageeri, P. Rodrigues, and J. Silberer. Happy or scared - detecting emotions of pedelec drivers in urban areas. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-4/W9:27–33, 2019. doi:10.5194/isprs-annals-IV-4-W9-27-2019.
- [8] W. O. Dick de Waard, Paul Schepers and K. Brookhuis. Mobile phone use while cycling: Incidence and effects on behaviour and safety. *Ergonomics*, 53(1):30–42, 2010. PMID: 20069479. doi:10.1080/00140130903381180.
- [9] M. Dozza, G. F. Bianchi Piccinini, and J. Werneke. Using naturalistic data to assess e-cyclist behavior. *Transportation Research Part F: Traffic Psychology and*

Behaviour, 41:217–226, 2016. Bicycling and bicycle safety. doi:10.1016/j.trf.2015.04.003.

- [10] M. Graystone, R. Mitra, and P. M. Hess. Gendered perceptions of cycling safety and on-street bicycle infrastructure: Bridging the gap. *Transportation Research Part D: Transport and Environment*, 105:103237, 2022. doi:10.1016/j.trd.2022.103237.
- [11] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, editors, *Human Mental Workload*, volume 52 of *Advances in Psychology*, pages 139–183. North-Holland, 1988. doi:10.1016/S0166-4115(08)62386-9.
- [12] J. Healey and R. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166, 2005. doi:10.1109/TITS.2005.848368.
- [13] W. Hochleitner, D. Sellitsch, D. Rammer, A. Aschauer, E. Mattheiss, G. Regal, and M. Tscheligi. No need to stop: exploring smartphone interaction paradigms while cycling. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*, MUM ’17, pages 177–187, New York, NY, USA, 2017. Association for Computing Machinery. doi:10.1145/3152832.3152871.
- [14] T.-Y. Hu, X. Xie, and J. Li. Negative or positive? the effect of emotion and mood on risky driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 16:29–40, 2013. doi:10.1016/j.trf.2012.08.009.
- [15] M. Jeon. Don’t cry while you’re driving: Sad driving is as bad as angry driving. *International Journal of Human-Computer Interaction*, 32(10):777–790, 2016. doi:10.1080/10447318.2016.1198524.
- [16] P. Jones. Sensory indiscipline and affect: A study of commuter cycling. *Social and Cultural Geography - SOC CULT GEOGR*, 13:1–14, 09 2012. doi:10.1080/14649365.2012.713505.
- [17] A. Kalra, T. Lim, L. Pearson, and B. Beck. Methods used to capture subjective user experiences in adults while riding bicycles: a scoping review. *Transport reviews*, 43(3):453–477, 2023. doi:10.1080/01441647.2022.2123064.
- [18] K. Kazemzadeh, R. Camporeale, C. D’Agostino, A. Laureshyn, and L. Winslott-Hiselius. Same questions, different answers? a hierarchical comparison of cyclists’ perceptions of comfort: in-traffic vs. online approach. *Transportation Letters*, 13:1–9, 03 2020. doi:10.1080/19427867.2020.1737373.
- [19] B. Laugwitz, T. Held, and M. Schrepp. Construction and evaluation of a user experience questionnaire. *USAB 2008*, 5298:63–76, November 2008. doi:10.1007/978-3-540-89350-9_6.

- [20] T. Lim, A. Kalra, J. Thompson, J. Caldwell Odgers, and B. Beck. Physiological measures of bicyclists' subjective experiences: A scoping review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 90:365–381, 2022. doi:10.1016/j.trf.2022.09.007.
- [21] G. López, A. M. Pérez-Zuriaga, S. Moll, and A. García. Analysis of overtaking maneuvers to cycling groups on two-lane rural roads using objective and subjective risk. *Transportation Research Record*, 2674(7):148–160, 2020. doi:10.1177/0361198120921169.
- [22] L. Ma, R. Ye, and H. Wang. Exploring the causal effects of bicycling for transportation on mental health. *Transportation Research Part D: Transport and Environment*, 93:102773, 2021. doi:10.1016/j.trd.2021.102773.
- [23] R. Manton, H. Rau, F. Fahy, J. Sheahan, and E. Clifford. Using mental mapping to unpack perceived cycling risk. *Accident Analysis and Prevention*, 88:138–149, 2016. doi:10.1016/j.aap.2015.12.017.
- [24] A. Matvienko, J.-B. Durand-Pierre, J. Cvancar, and M. Mühlhäuser. Text me if you can: Investigating text input methods for cyclists. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA '23, New York, NY, USA, 2023. Association for Computing Machinery. doi:10.1145/3544549.3585734.
- [25] R. C. McIlroy, K. L. Plant, and N. A. Stanton. Thinking aloud on the road: Thematic differences in the experiences of drivers, cyclists, and motorcyclists. *Transportation Research Part F: Traffic Psychology and Behaviour*, 83:192–209, 2021. doi:10.1016/j.trf.2021.09.014.
- [26] M. Meenar, B. Flamm, and K. Keenan. Mapping the emotional experience of travel to understand cycle-transit user behavior. *Sustainability*, 11(17), 2019. doi:10.3390/su11174743.
- [27] J. Mesken, M. P. Hagenzieker, T. Rothengatter, and D. de Waard. Frequency, determinants, and consequences of different drivers' emotions: An on-the-road study using self-reports, (observed) behaviour, and physiology. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(6):458–475, 2007. doi:10.1016/j.trf.2007.05.001.
- [28] T. D. Ngo, A. Kashani, G. Imbalzano, K. T. Nguyen, and D. Hui. Additive manufacturing (3d printing): A review of materials, methods, applications and challenges. *Composites Part B: Engineering*, 143:172–196, 2018. doi:10.1016/j.compositesb.2018.02.012.
- [29] J. Y. M. Nuñez, I. P. Teixeira, A. N. R. d. Silva, P. Zeile, L. Dekoninck, and D. Botteldooren. The influence of noise, vibration, cycle paths, and period of

day on stress experienced by cyclists. *Sustainability*, 10(7), 2018. doi:10.3390/su10072379.

- [30] P. Oja, S. Titze, A. Bauman, B. de Geus, P. Krenn, B. Reger-Nash, and T. Kohlberger. Health benefits of cycling: a systematic review. *Scandinavian Journal of Medicine & Science in Sports*, 21(4):496–509, 2011. doi:10.1111/j.1600-0838.2011.01299.x.
- [31] A. Rasch, S. Moll, G. López, A. García, and M. Dozza. Drivers’ and cyclists’ safety perceptions in overtaking maneuvers. *Transportation Research Part F: Traffic Psychology and Behaviour*, 84:165–176, 2022. doi:10.1016/j.trf.2021.11.014.
- [32] G. Rybarczyk, A. Ozbil, E. Andresen, and Z. Hayes. Physiological responses to urban design during bicycling: A naturalistic investigation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 68:79–93, 2020. doi:10.1016/j.trf.2019.12.001.
- [33] M. Schrepp and J. Thomaszewski. Design and validation of a framework for the creation of user experience questionnaires. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(7):88–95, December 2019. doi:10.9781/ijimai.2019.06.006.
- [34] Y. Tan, S. H. Yoon, and K. Ramani. Bikegesture: User elicitation and performance of micro hand gesture as input for cycling. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’17, pages 2147–2154, New York, NY, USA, 2017. Association for Computing Machinery. doi:10.1145/3027063.3053075.
- [35] P. W. Woźniak, L. Dekker, F. Kiss, E. Velner, A. Kuijt, and S. F. Donker. Brotate and tribike: Designing smartphone control for cycling. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI ’20, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3379503.3405660.
- [36] H. YAMANAKA, P. Xiaodong, and J. SANADA. Evaluation models for cyclists’ perception using probe bicycle system. *Journal of the Eastern Asia Society for Transportation Studies*, 10:1413–1425, 2013. doi:10.11175/easts.10.1413.
- [37] P. Zeile, N. Haug, C. Schmidt-Hamburger, N. Motzer, M. Moser, and B. Resch. Emotion sensing for (e-) bicycle safety and mobility comfort. *Journal of Digital Landscape Architecture*, 8:381–389, 2023. doi:10.14627/537740041.
- [38] S. Zepf, M. Dittrich, J. Hernandez, and A. Schmitt. Towards empathetic car interfaces: Emotional triggers while driving. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA ’19, pages 1–6, New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3290607.3312883.

- [39] S. Zepf, J. Hernandez, A. Schmitt, W. Minker, and R. W. Picard. Driver emotion recognition for intelligent vehicles: A survey. *ACM Comput. Surv.*, 53(3), July 2020. doi:10.1145/3388790.
- [40] R. Zhang, M. Te Brömmelstroet, A. Nikolaeva, and G. Liu. Cycling subjective experience: A conceptual framework and methods review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 101:142–159, 2024. doi:10.1016/j.trf.2023.12.021.