VeatherReflect: Employing Weather as Qualitative Representation of Stress Data in Virtual Reality

Nadine Wagener
University of Bremen
Bremen, Germany
nwagener@uni-bremen.de

Marit Bentvelzen
Utrecht University
Utrecht, Netherlands
m.bentvelzen@uu.nl

Bastian Dänekas
University of Bremen
Bremen, Germany
daenekba@uni-bremen.de

Paweł W. Woźniak
Chalmers University of Technology
Gothenburg, Sweden
pawel.wozniak@chalmers.se

Jasmin Niess
University of Oslo
Oslo, Norway
jasminni@ifi.uio.no

Figure 1: VeatherReflect is a VR application that visualises personal tracker stress scores using weather metaphors. The user stands on a platform in space. A semicircle presents a two-hour time frame. Floating videos represent stress levels and can be selected to immerse oneself in the represented weather scenario (sun, snow, rain, thunderstorm).

ABSTRACT

While personal trackers can collect a vast amount of information about their users, the representation of such data has remained unchanged, with bar charts being the most dominant. However, to build systems that facilitate reflection and support well-being, it is crucial to explore alternative ways of representation. Thus, we designed VeatherReflect, a VR application that uses weather metaphors to illustrate tracker stress scores, aiming to encourage users to reflect on their stress data. In a pre-study, we mapped stress scores to weather states. We then compared VeatherReflect with a standard visualisation of stress data presented in VR. VeatherReflect increased participant engagement with personal data and stress awareness. Participants reported reflective insights for stress-reducing behaviour. We contribute findings on how virtual weather as a metaphor for stress can support reflection. We discuss design recommendations for VR applications aiming to facilitate a deeper understanding of complex personal data through engaging qualitative experiences.

DIS '23, July 10–14, 2023, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9893-0/23/07.
https://doi.org/10.1145/3563657.3596125

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**.

KEYWORDS

Virtual Reality; Personal Informatics; Reflection; Fitness Tracker; Stress; Visualisation; Well-being

ACM Reference Format:

Nadine Wagener, Marit Bentvelzen, Bastian Dänekas, Paweł W. Woźniak, and Jasmin Niess. 2023. VeatherReflect: Employing Weather as Qualitative Representation of Stress Data in Virtual Reality. In *Designing Interactive Systems Conference (DIS '23), July 10–14, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3563657.3596125

1 INTRODUCTION

In 2020, the American Psychological Association proclaimed a national mental health crisis due to a significant increase of stress levels across the USA [2]. They report that stress can negatively affect sleeping and eating patterns as well as social and intimate relationships. Further, the report states that stress can lead to being emotionally unstable and experiencing symptoms of depression and burnout. Thus, it is increasingly important to care about one's mental health and well-being. While there are many approaches and applications for self-care, e.g. promoting stress reduction through breathing adaptation [57, 61, 66, 75], an important first step is to get an overview of and start reflecting on one's stress levels. The primary way in which current technologies can support that process is stress tracking with fitness trackers and smartwatches, which became increasingly popular in recent years [40]. Yet, research shows that motivation to use smartwatches decreases over time due to various reasons [18]. For example, users feel as if smartwatches cannot provide effective in-the-moment stress interventions while, when interpreting stress data in retrospect, they often lack knowledge of the underlying technological mechanisms regarding the measurement of stress [20]. This can lead to increasing disinterest to engage with one's data [20]. Further, research in personal informatics indicates that current standard representations for data, such as bar charts and graphs, are unlikely to effectively foster engagement and reflection [4, 8, 24, 41]. To build systems that effectively and lastingly support users in reflecting on their data, we need to study alternative ways to represent personal tracker measures such as stress levels. Thus, there is a necessity to explore alternative forms of data representation in personal informatics to facilitate reflection.

Concurrently, an increasing amount of research shows that Virtual Reality (VR) may be an effective and immersive approach to present personal data and foster reflection. For example, Yoo et al. [74] visualised physical exercises and daily step data in an interactive VR dashboard. However, the system visualised the data both in VR and on a mobile phone in a similar way (i.e. using bar charts). This example illustrates the unused potential for VR to offer immersive personal data representations that go beyond classic fitness tracker visualisations such as bar charts. The potential of VR for stress reduction and relaxation has been explored extensively [29, 46, 58, 65, 73]. Yet, research in representing personal tracking data such as stress data in new way to support reflection by using the unique affordances of VR needs further attention. We hypothesise VR might foster deeper interest and reflection through offering a new perspective on personal data.

The range of experiences that VR provides is larger than what desktop or mobile systems offers. For instance, VR can offer unique forms of data representation through using qualitative, immersive and more abstract forms of data representations. In this paper, we introduce weather as a novel qualitative data representation for stress tracking data. Weather can influence perceived happiness [17, 32], stress [6], and certain weather types are associated with mental health characteristics [39, 63]. Furthermore, albeit not scientifically proven, people associate similar weather scenarios with emotions and states of being relaxed (sunshine) vs danger (hail, thunderstorm). Thus, weather might offer a valuable qualitative alternative to represent stress data and support reflection.

To that end, we designed VeatherReflect—a VR application that supports reflection on personal tracker stress scores using weather as qualitative means of data representation. In a pre-study ($N = 100$), we asked participants to map weather scenarios to a variety of different stress levels. The design of VeatherReflect was informed by those results. We then conducted a user study ($N = 20$) to inquire user engagement and level of reflection with their personal stress data using VeatherReflect and a standard bar chart visualisation of stress data, also presented in VR. We found that VeatherReflect increased stress awareness and led to more reflective insights for stress-reducing behaviour. It significantly increased immersion, mindfulness and user engagement with personal data.

This paper contributes the following: (1) the design and implementation of VeatherReflect—a virtual environment with the aim to support reflection through representing stress levels using weather as a metaphor; (2) a mixed-method user study to evaluate VeatherReflect; and (3) design recommendations for VR applications that aim to facilitate a deeper understanding of complex personal data using qualitative means to design engaging experiences.

2 RELATED WORK

In this section, we first discuss previous work on reflection in HCI to contextualise our work. We then describe the relationship between emotions, health and weather to illustrate the reasons behind choosing weather as a metaphor for qualitative data representations of stress. Finally, we outline how VR can be used to foster reflection.

2.1 Reflection in HCI

The HCI field invested considerable effort in building an understanding of reflection, as it can support life changes [59], offer more self-insight [5], and benefit health and well-being [13, 36, 55].

Recent advancements in technology have led to widespread use of commercial tracking devices (e.g. fitness trackers), that are equipped with more and more sensors and offer an ever-increasing level of accuracy. Such devices enable users to collect previously hidden information about themselves and by reflecting on personal data, a user can notice patterns and trends, which can in turn lead to more knowledge about oneself [22]. However, such systems have been critiqued in the past as they often do not actively encourage reflection [15, 30]. As noted by Baumer [5], personal tracking devices often carry an implicit assumption that by showing a user visualisations of their past data for the purpose of reflection, that reflection will automatically occur. Yet, this conflicts with reflection theories that emphasise the importance of encouraging reflection, as it often does not occur automatically [55].

There are various strategies that can be used to encourage reflection, as shown by Bentvelzen et al. [8]. Through a structured literature review and analysis of mobile applications for reflection, they propose a taxonomy consisting of eleven design resources and 74 design patterns. However, while there are many strategies available to support reflection, the majority of commercial fitness trackers apply just one strategy i.e., depicting personal data as bar charts in a mobile app. To build systems that help users reflect on their data and thereby effectively support well-being, there is a need to study alternative strategies to represent personal tracker measures.

This need is further emphasised by work by Ding et al. [20]. They conducted a study with fitness tracker users who tracked their stress. They found that participants largely did not engage with stress data and tracking did not result in an increased understanding of daily stress, but often led to confusion. There was a mismatch between the physiological stress measured by the device and their personal conceptualisation of stress. These results indicate that the potential for stress tracking to improve well-being is underexplored and alternative ways of reflecting on and engaging with stress tracking data are required to unlock its possible benefits. Our work addresses this gap by exploring alternative ways for supporting reflection on stress data through exploring stress data metaphors presented in VR.

2.2 Influence of Weather on (Mental) Health

Although the relationship between weather and emotions, mood and (mental) health is not as strong as generally assumed [19, 23, 25, 34], health and well-being can be influenced by weather. For example, less sunshine can increase stress [6], and amplify or trigger seasonal affective disorders [39, 63], while intensity and direction of wind can influence the level of energy and fear [11]. Along similar lines, cognitive performance decreases with rising temperature [71]. Further, Yu et al. [76] used machine-learning to predict future self-reports of users regarding their well-being, including their stress level partly based on weather data. In their study, weather was found to be strongly associated with well-being labels. Regarding physical health, headaches and joint pain can increase with higher temperature [33], although a connection of arthritis with weather cannot be finally confirmed [56]. Additionally, numerous weather metaphors are used to express emotional states [78], e.g. connecting thunder with negative emotions and sun with optimism or joy. Such common associations of weather are also used in design to encode emotional information, e.g. linking anger to thunderstorms and joy to a clear blue sky [62].

On another note, weather associations presented in VR have also been explored to support (mental) health. For instance, building on notions that VR elicits thermal perceptions [14], virtual environments have been used to facilitate pain reduction, e.g., immersing patients with severe burns in snowy environments (e.g. [45]). Weather can be also used as a means of improving the user experience in VR. For example, realistic weather in VR can lead to an increase of presence [3], can enhance or reduce motion sickness depending on the weather type [1], or can be used to convey meanings as part of VR story telling [48].

These works highlight the common associations between weather conditions and (mental) health, including emotional responses to them, and highlight the potential of using weather as qualitative metaphor for emotional and health-related content. Yet, its association to stress is unverified and will be addressed in our pre-study. This paper focuses on weather metaphors in VR and how they might foster reflection. Thus, it is situated outside of a clinical context. Instead, we use weather as qualitative representation of real-life stress data in VR and explore if and how such representations can support reflection.

2.3 Supporting Reflection with Virtual Reality

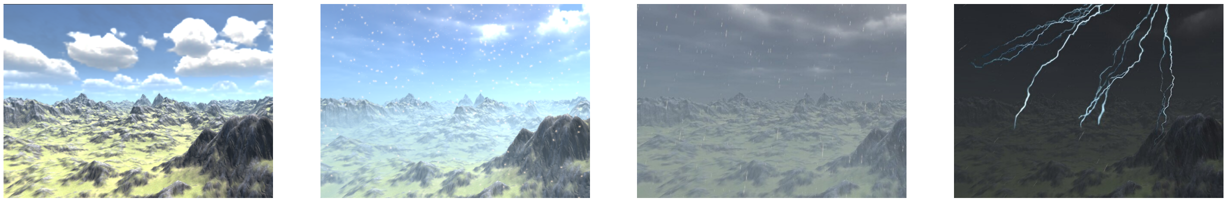
Many VR applications strive to offer stress management interventions [58], mostly through providing gamified approaches for stress reduction. For example, a variety of systems have been designed to adapt to one's breathing thereby supporting relaxation. [57, 61, 66, 75]. While mirroring real-time stress reactions to users can facilitate reflection-in-action [51], such systems do not provide feedback on one's stress level in retrospect. Users have to get to know themselves to realise beforehand, without the system's help, that they may potentially experience stress and need stress treatment. Reflection-on-action can assist in this process. To support reflection-on-action, there is a need to research how stress data can be visualised using the unique affordances of VR which we will address in this paper.

However, most VR applications utilise reflection as means to learn another concept, for example linking it with empathy skills (e.g. [60]), learning (e.g. [31]), or training (e.g. [77]). Some also found that their interactive system facilitated reflection, although it was designed with another purpose in mind, such as autonomous emotional expression [69]. While Jian and Ahmadpour [27] have built a model to understand how reflection should be supported in VR based on reviews of twelve VR applications, they neither focus on personal informatics nor on specific design approaches that facilitate reflection, such as weather.

There have been a few studies studies that demonstrate the connections between personal informatics, virtual reality (VR), and reflection. Two notable examples are studies conducted by Egan et al. [21] and Yoo et al. [74]. Egan et al. [21] introduced health data such as heart rate data as a reliable way to measure the quality of a virtual environment (VE), while Yoo et al. [74] developed a VR dashboard that offers users a reflective overview of their personal sports data, such as step count. Building on these previous studies, our work aims to extend and combine these approaches. Specifically, we focus on designing for reflection and creating engaging visualisations of complex personal data gathered by smartwatches, all within a VR environment.

3 DESIGN

Although previous studies found (perceived) connections between weather and emotions (see. subsection 2.2), little research has linked weather to stress levels. Hence, there is a need to analyse how different weather scenarios relate to different stress levels. To that



(a) ***No stress*** **was mapped to** ***sunny*** **weather.** (b) ***Low stress*** **was mapped to** ***snow*** (c) ***Medium stress*** **was mapped to** ***heavy rain*** (d) ***High stress*** **was mapped to** ***thunderstorm***

Figure 2: In the pre-study, participants mapped pictures of seven weather scenarios to four stress levels. Figure 2 depicts the findings of the most often selected weather scenarios for each stress level, which were then used as qualitative representation of stress in the main study.

end, we first conducted a pre-study to study how participants map weather scenarios to different stress levels. We further examined reasons behind those choices by asking participants to describe a situation they imagined when thinking about a specific stress level.

3.1 Pre-Study: The Interplay Between Weather and Stress

We conducted a within-subjects study utilising a card-sorting task. We asked participants to assign weather conditions to stress levels. Weather conditions were chosen based on common weather types [43] and adapted according to the availability and visualisation possibilities for VR [1]. Thus, we ended up with the following seven weather conditions: Sun with Rainbow, Sun, Fog, Light Snow, Heavy Snow, Heavy Rain, Thunderstorm. Stress levels were informed by common representations of stress levels used by Garmin and Apple smartwatches: No stress = 0–25%, low stress = 26–50%, medium stress = 51–75% and high stress = 76–100%. Thus, participants mapped seven weather states to four stress levels. For each of the stress levels (no stress, low stress, medium stress and high stress), participants further described a situation they had in mind when imagining feeling this level of stress. We chose this approach to understand participants' reasoning behind the mapping, and, effectively, to gather more in-depth insights that might inform the design of VeatherReflect.

3.1.1 Participants. One hundred participants (N = 100) took part in the online study (55 male, 42 female, 2 other, 1 non-binary) with an average-age of 29.75 years (min = 18 years, max = 74 years). Country of residence encompassed 24 different countries. The study took 6,55 minutes on average.

3.1.2 Findings. The results of the pre-study are shown in Figure 3. *No stress* was mainly associated with *sunny* (42 times), *low stress* with *light snow* (31 times), *medium stress* with *heavy rain* (30 times), and *high stress* with a *thunderstorm* (46 times).

Further, we investigated why those weather scenarios were chosen through categorising the reported situations depicting each stress level. A detailed overview of this analysis is included in the supplementary material. *High stress* was primarily associated with feeling time pressure due to upcoming exams or deadlines (53 times). *Medium stress* was experienced mostly at work (27 times) and in public spaces (14 times), e.g. when talking to strangers or waiting in queues. A *low stress* level was mainly associated with being at home (12 times), and among friends, family and pets (10 times). *No stress* was associated with being on vacation (21 times), relaxation (24 times) and good weather (17 times).

3.1.3 Design Implications for VeatherReflect. Based on our pre-study, we identified a clear trend regarding the association of designated weather scenarios with designated stress levels. These insights serve as inspiration for the design of VeatherReflect. Based on the results of the pre-study, we mapped *no stress* to *sunny*, *low stress* to *light snow*, *medium stress* to *heavy rain*, and *high stress* to *thunderstorm*, as depicted inFigure 2. Consequently, these weather conditions were chosen to visualise the stress levels in our main study.

Yet, we found some interesting outliers in our data, which show that a certain level of individual interpretation needs to be taken into account. To give an example, *no or low stress* were sometimes mapped to thunderstorm or heavy rain. Those participants explained, that they - as music lovers - imagined listening to loud music or thought about positive sparks "similar to lightning" within a relationship. In some other cases, *high stress* was mapped to sun, which was a decision informed by this year's drought.

Moreover, participants often mentioned that the sound of the weather has a big impact on their perception of weather. For instance, listening to a light rain or snowfall is soothing, while thunder makes them afraid. Accordingly, we decided to add sound effects to VeatherReflect as well. While this adds another modality, we are interested in researching the effects of a cohesive weather scenario utilising a variety of modalities VR has to offer. We hypothesise that the experience of an immersive weather metaphor in a virtual environment can support reflection on stress data. Hence, based on the insights of our pre-study and the aim of this paper, we opted to add sound to the overall experience. For a more detailed discussion of this aspect, see subsection 6.2. Sun was accompanied by birds' chirping, snow by the sound of light wind, rain by the sound of drops falling on the ground, and thunderstorm by loud rolling thunder.

[1]To represent those weather scenarios in VR, we used the VR package *Enviro - Sky and Weather* in the main study. Some weather types, such as wind/storm and frost, are hard to depict in photos and in VEs, which is why we decided against them.

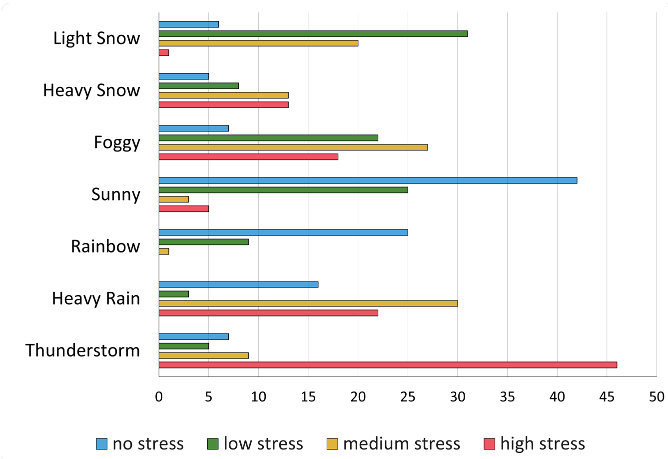


Figure 3: Results of the pre-study. Seven weather scenarios were mapped to four stress levels (no stress, low stress, medium stress, high stress).

3.2 Final Prototype

The design considerations derived from the pre-study (see subsection 3.1) informed the design of VeatherReflect. The aim of the system is to support users in reflecting on their stress data. To that end, VeatherReflect strives to immerse users in weather conditions that match prior measured stress levels. The relative times of the two hours of measured stress data was placed in a half circle around the user. This represents the x-axis of a standard visualisation of graphs in Garmin smartwatches. Each 15 min segment of the graph (for more information, see subsection 4.1) is mapped to one weather scenario. If participants did not experience any stress (*no stress*) during that time interval, this was represented by *sunny* weather, *low stress* by *light snow*, *medium stress* by *heavy rain* and *high stress* by being immersed in a *thunderstorm*. These weather conditions are showcased via a floating pre-view video depicting a short clip of that weather scenario, e.g. where it is snowing. Peaks, i.e. deviations from the general common stress level in the time interval (for more information see subsection 4.1), were also depicted via floating pre-view videos that were placed at the correct time in the semicircle and placed a bit higher to make them stand out more. Apart from that, the virtual environment consisted only of a barren-looking rectangular plane in outer space. We did not offer any terrain or landscape as this could unconsciously affect the user—a thunderstorm might feel different when standing on a mountain than next to a lake. As skybox, we chose a space environment (see Figure 1 as reference). First, this backdrop is often used in prior research (e.g. [69, 70]) as it is non-distracting, mostly dark, and the backdrop highlights changes in the surrounding, such as added colour or weather. Second, stars and the slightly pink coloured milky way provide enough light to not influence users' mood in a negative way [69, 70]. An exemplary setup of VeatherReflect is shown in Figure 4.

By stepping closer to the floating video previews, participants could select the depicted weather condition by pressing the controller's trigger button. Upon selection, participants activate that weather scenario, e.g. the cloud cover changes, and they see and hear the rain. To provide as little distraction to the feeling of being surrounded by each weather scenario, the timeline of the semi-circle with the attached other weather scenarios disappeared (see Figure 4). Everything else, such as the barren-looking rectangular field in outer space remained unchanged. Schematically, this is depicted in Figure 1. In order to return, participants selected a small preview picture of the semi-circle starting scene which is placed on a column in the centre of the stage behind the user.

4 EVALUATION

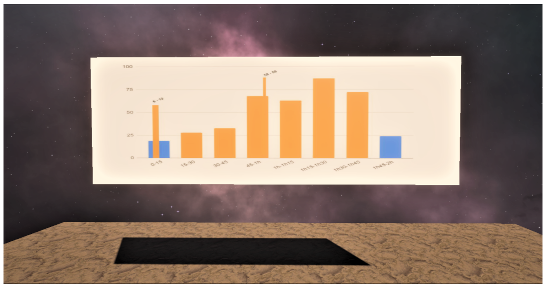
We conducted a user-study to evaluate VeatherReflect and compare it with the baseline 2D-Plot. The baseline presents participants with a standard visualisation of stress levels using bar charts in VR. Consequently, we conducted a mixed-method within-subject study with two conditions. The order of the conditions was counterbalanced and randomised.

We decided to compare VeatherReflect to a 2D graph in VR as we endeavoured to investigate if and how an immersive, qualitative representation of stress data in VR can support reflection on stress indicators compared to the standard representation of stress data used in consumer products. An alternative that we considered but rejected was using graphs provided on a smartwatch or another mobile or desktop device. We decided to not compare a stress data representation on a mobile or desktop device to an experience in VR. Presenting data on a mobile device cannot immerse users as VR can, thus potentially adding confounding variables. Also, due to the novelty effect of VR, users might rate any experience as better just because it is in VR. Thus, we opted to translate the 2D graph to VR, hence keeping the same mode of presentation while still providing the same visualisation (i.e. a bar chart) as on a smartwatch.

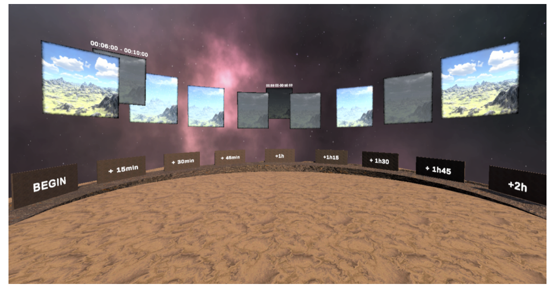
Consequently, both conditions are experienced in VR and share a similar-looking environment of a barren-looking rectangular field in outer space. Regarding the stress data visualisation, currently commercially available consumer products display stress data differently. For instance, smartwatches from Apple present stress levels in circular bar charts, while wearables from Fitbit and Garmin additionally provide bar charts over time. Yet, in both representation forms, it remains unclear in which time intervals stress is measured. To disambiguate, we chose to visualise more granular level using a continuous graph. Thus, we present the users with stress measurements taken during an activity related to stress and relaxation, e.g. Yoga. Stress levels during the exercise are then presented as bar charts. Time is then as a relative factor, counting from $0:00:00$ upwards from the start of the the activity. Time is presented on the x-axis, and stress levels on the y-axis of the graph. For added ecological validity, we use a colouring scheme based on a consumer-grade product, the Garmin Connect App. Garmin colour codes the stress levels in the following manner: *no stress* is colour-coded in blue, all the other stress levels uniformly in orange.

4.1 Conditions

We conducted a within-subjects study with two conditions. The baseline C0 is called 2D-Plot, and condition C1 VeatherReflect. Based on prior considerations as elaborated above, 2D-Plot was also set in VR. Here, participants experienced their tracker stress



(a) Condition 2D-Plot. An aggregated graph of a participant's stress data is shown in VR.



(b) Condition VeatherReflect. Participant's stress level is visualised by weather scenarios.

Figure 4: Representation of both conditions experienced in VR with the same data as basis. Within the two hours, the participant encountered two spikes, which were added to the eight bars or weather scenarios.

scores in form of a graph, based on two hours of recorded stress data. This graph visually remains close to the actual data from smartwatches, but its data is prior processed to keep both conditions as comparable as possible. Details about the data preparation can be found in subsection 4.2. This graph is employed as material of a virtual plane in VR. The plane, sized 2m in width and 1m in height, was placed approximately on eye level of the participants in a distance of 2m. Thus, the bar chart can be seen in whole without moving the head, and participants can also go up to certain segments and have a closer look. On average, participants spend 00:02:20 minutes (*SD* = .034) with this condition.

Condition C1, VeatherReflect, is based on the same pre-processed stress scores measured by a smartwatch as condition 2D-Plot. In contrast, here participants' tracker stress scores are represented in form of four types of weather conditions. The design of VeatherReflect is explained in detail in subsection 3.2. On average, participants spend 00 : 05 : 56 minutes (*SD* = .116) in that condition.

4.2 Data Preparation

In our study, we recorded two hours of stress data with a Garmin smartwatch (also see subsection 4.6). We used a two-hour recording period to balance regarding gathering conclusive stress data (that is granular enough to show spikes in stress) while minimising participant burden. However, to further investigate stress patterns over time, future studies could adopt a longitudinal design with stress data collected at regular intervals throughout the day. The data points of the two-hour recording were then split in eight segments of fifteen minutes each. Each segment is assigned with the average of its stress values (e.g. when the majority of bars show a low stress level - between 26–50% - then the whole segment will be depicted as having a low stress level). As this somewhat evens out the data, up to four peaks per participant were additionally displayed. Peaks are defined as the strongest deviations from the computed average, e.g. a peak of high stress in a segment of low stress. Peaks were extracted together with the specific time interval they apply to and also averaged. As an example, in a 15min segment of low stress (26–50%), we identified 7min of high stress, split up into three bars with a height of 78%, 80% and 91%. These bars will be depicted as one peak with the calculated average height of 83%. The time interval of each peak is shown on top. X- and y-axis as well as the colour-coding were adopted from the original bar chart taken from Garmin smartwatches. Based on our data processing, we obtained graphs similar to Figure 4.

4.3 Data Collection

Quantitative data was collected from four questionnaires (UES-SF, TSRI, SMS, IPQ-s) which the participants completed after experiencing each condition. Further, we gained qualitative insights through post-test interviews.

4.3.1 Measures. The following dependent variables were measured for both conditions:

User Engagement can be used to measure the quality of the user experience, and the participants' ability to engage and sustain engagement in digital environments [44]. As reflection often requires encouragement [5, 55] and, thus, active involvement from the user, we explore if users engage differently in the two conditions. It was measured with the User Engagement Scale-SF [44]. This 12-item scale, answered on a 5-point-Likert scale is divided into four subscales: *Aesthetic Appeal (AE)*, *Focused Attention (FA)*, *Perceived Usability (PU)*, *Reward Factor (RW)*. All items can be summed and divided by twelve to calculate an overall engagement score.

Reflection that an interactive system can evoke was measured utilising the Technology-Supported Reflection Inventory (TSRI), as proposed by Bentvelzen et al. [7]. The TSRI supports understanding the qualities of an interactive technology which supports reflection, and evaluates how effectively such a system supports reflection. Participants answer nine items on a 7-point-Likert scale. They are categorised in three dimensions, *Insight*, *Exploration* and *Comparison*.

Mindfulness practice has been shown to positively correlate with reflection, through opening up headspace which facilitates reflection [28, 42], and by being focused on the present moment [9]. This can enable awareness. Mindfulness was also found to be a positive contributing factor when reflecting on biosignals [35]. Consequently, if a system measurably increases mindfulness, the probability for reflective practice potentially increases as well. Thus, we also measure mindfulness, using the state mindfulness scale (SMS) [64]. SMS reflects mindfulness as the objects of mindful awareness (i.e. what experience a person focuses on, including physical sensations and mental events) and the qualities of mindful awareness (i.e. how a person attends to experience, including perceptual sensitivity to stimuli, deliberate attention, willingness to feel, curiosity) [49, 64]. The 21 items, rated on a 5-point Likert scale, are divided into two subscales that measure bodily sensations (SMS-body, $max = 24$) and attention to and awareness of mental events such as emotions and patterns of thought (SMS-mind, $max = 60$). Higher scores reflect higher levels of state mindfulness. A sum of the subscales results in the total score of mindfulness.

Presence, defined as the sense of being there in the VE [54], can be linked to reflection as well, as present-moment attention can facilitate reflection processes [9]. Presence was measured using the Igroup Presence Questionnaire (IPQ) [52]. The IPQ is a 14-item scale answered on a 7-point-Likert scale. It is divided into four subscales: general presence (sense of being there), involvement (attention devoted to the VE), experienced realism (subjective experience of realism in the VE), and spatial presence (feeling physically present in the VE).

4.4 Interview Protocol

We conducted semi-structured interviews that lasted 14 minutes on average. All audio recordings were transcribed verbatim and imported into dovetail software. Two authors coded three interviews using open coding. Next, a coding tree was established through iterative discussion. The remaining transcripts were coded individually by one author using the coding tree. A final discussion session between two authors was conducted to identify emerging themes applying thematic analysis [12]. Within the interview, we asked about differences between the visualisations, focusing on exploration, understanding, engagement and insights that participants could gain. The interview protocol can be found in the supplementary material.

4.5 Participants

We used our extended social network and snowball sampling to recruit participants. In total, $N = 20$ participants (9 females, 11 males) took part in the study ($M = 27.26$ years, min : 20 years, max : 35 years). They received 20€ as remuneration. Twelve of the participants identified as German, two as Russian, two as Indian, and one each as Italian, Greek, Bangladeshi, and Chinese. Thirteen participants did not own a smartwatch, three owned a Fitbit, two an Apple watch, one a Xiaomi band and one a Mi fit band. If used, they mainly tracked sleep data and step counts, only two reported checking stress, albeit on an irregular basis. Four participants had never tried out VR before, five tried VR once or twice in their lives, seven experienced VR several times a year, one used VR several times a month, and two several times a week.

4.6 Study Procedure

Participants were asked to wear either a Garmin Venu sq or a Garmin Vivoactive 4 for at least two hours during their daily routine. To enable granular stress tracking (see subsection 4.1), we started a Yoga activity for these two hours. Our study represents an initial step towards understanding the potential of qualitative data representations in VR to foster reflection. Within the next 24 hours, the participants would take part in the second phase of the study, experiencing two different approaches to represent personal stress data in VR.

For this second part of the study, the participants gave consent, and completed demographics, SRIS and iPAQ-s questionnaires. Afterwards, the experimenter explained the procedure, emphasising that in each condition the participants will experience different forms of visualisations of their stress data that they can freely explore, thus without having a specific task. Before experiencing the VeatherReflect condition, participants were additionally informed how to use the controller to enter weather scenes. They were informed that weather scenarios will be used to represent their stress data, but were not instructed about which weather scenarios was mapped to which stress level beforehand. Screen and audio were recorded, in case that participants started describing their experience in line with a thinking-aloud method. After each condition, participants filled out the UES-SF, TSRI and IPQ questionnaires. Post-test, the experimenter conducted a semi-structured interview. All participants were also informed that they could stop the sessions at any time without giving a reason and without any negative consequences.

5 RESULTS

Based on the evaluation, we gathered quantitative results from the questionnaires as well as qualitative insights from the interviews. Our findings will be presented in this section.

5.1 Quantitative Results

Our quantitative results comparing 2D-Plot and VeatherReflect in regard to user engagement, mindfulness, reflection and presence are presented in this section. A data set of one participant had to be excluded from the calculations. Specifics about the calculations can also be found in the supplementary material.

5.1.1 User Engagement (UES-SF). We used a paired t-test to investigate the effect of VeatherReflect on the UES-SF and all its subscales. In regard to *Aesthetic Appeal (AE)*, we found a significant effect, $t(18) = 4.07$, $p < .01$. We further found a significant effect of VeatherReflect on the UES-SF subscale *Focused Attention (FA)*, $t(18) = 7.1$, $p < .001$. We used a paired t-test to investigate the effect of VeatherReflect on the UES-SF subscale *Reward Factor (RW)*. We found a significant effect, $t(18) = 3.00$, $p = .04$. However, we did not find a significant effect of VeatherReflect on *Perceived Usability (PU)*, $t(18) = 1.2$, $p = .25$. Finally, we used a paired t-test to investigate the effect of VeatherReflect on the UES-SF *Total (TOT)*. We found a significant effect, $t(18) = 4.9$, $p < .001$. All p-values were corrected using Bonferroni correction. The significant results are illustrated in Figure 5.

5.1.2 Mindfulness (SMS). We used a paired t-test to investigate the effect of VeatherReflect on the SMS subscale *SMS-Mind*. We found a significant effect, $t(18) = 2.97$, $p = .03$. We also found a significant effect of VeatherReflect on the SMS subscale *SMS-Body* when using a paired t-test. We found a significant effect, $t(18) = 3.12$, $p = .02$. We used a paired t-test to investigate the effect of VeatherReflect on the total SMS score *SMS-TOT*. We found a significant effect, $t(18) = 3.23$, $p = .02$. The results are shown in Figure 6. All p-values were corrected using Bonferroni correction.

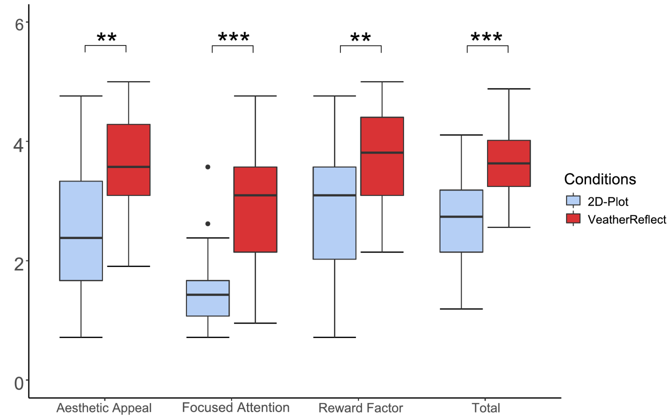


Figure 5: Significant results of the User Engagement Scale. Condition 2D-Plot is shown in blue, condition VeatherReflect is represented in red. Significance levels are indicated with * for $p \leq .05$, ** for $p \leq .01$ and * for $p \leq .001$.**

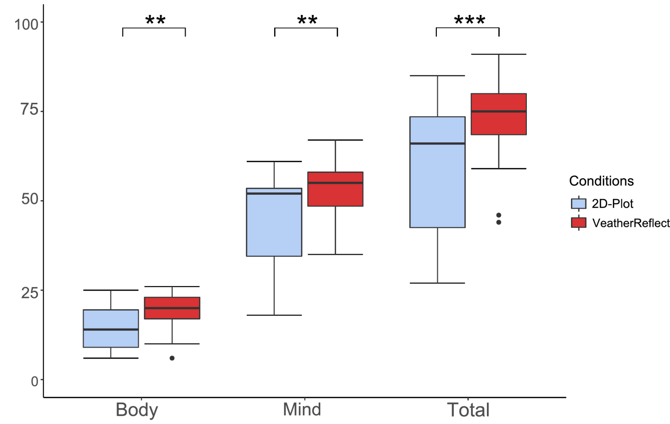


Figure 6: Significant results for mindfulness as measured by the SMS. Condition 2D-Plot is shown in blue, condition VeatherReflect is represented in red. Significance levels are indicated with * for $p \leq .05$, ** for $p \leq .01$ and * for $p \leq .001$.**

5.1.3 Reflection (TSRI). To measure reflection quantitatively, we used a non-parametric measure as recommended by Bentvelzen et al. [7]. Thus, we used the Wilcoxon paired rank sum test to measure the level of reflection. For the subscale *Comparison*, we did not find any significant differences, $V(18) = 76.5$, $p = .41$. In regard to the subscale *Exploration*, we also did not find any significance, $V(18) = 39$, $p = .72$. Similar results were found for the subscale *Insight*, $V(18) = 25.5$, $p = .87$. All p-values were corrected using Bonferroni correction.

5.1.4 Presence (IPQ). We used a paired t-test to investigate the effect of VeatherReflect on IPQ and all subscales. In regard to *IPQ-General*, we did not find a significant effect, $t(18) = 1.32$, $p = .78$. Further, we could not find a significant effect in the subscale *IPQ-Inv*, $t(18) = 2.9$, $p = n.s.$ as well as in *IPQ-Spatial*, $t(18) = 2.22$, $p = .14$ and in *IPQ-Real*, $t(18) = 0.31$, $p = .76$. All p-values were corrected using Bonferroni correction.

5.2 Qualitative Findings

Based on our qualitative inquiry, we identified four themes: *Interaction and Movement*, *Task-Load and Understanding*, *Recollection and Reflection*, and *Caring and Sharing*. Our findings are described below and illustrated with excerpts from the interviews.

5.2.1 Interaction and Movement. In the condition 2D-Plot, we found that participants would have liked to have more interaction with their data. To illustrate, some participants tried to move items such as single bars to get more information about their data. In contrast, participant enjoyed the interaction possibilities provided in VeatherReflect. They felt empowered to choose if they want to be immersed in a certain weather scenario. Participants reported that having the freedom of choice as well as physically moving towards the weather previews and actively selecting them led to higher engagement, interest and joy when interacting with their data in VeatherReflect, and to increased feelings of immersion. Participants also explained that they felt the urge to move around and stay longer within the weather scenarios to absorb the atmosphere. As an effect, some participants physically reacted to the weather phenomena in that condition. For example, they mentioned shivering and pulling up their shoulders in the snow scenario.

> [VeatherReflect] triggered that I moved around in the environment. The first [condition] didn't do this. And I think that helped me also to get a little bit more invested into the whole experience in the second prototype because I felt a bit more immersed in the environment. (P5)

5.2.2 Task-Load and Sensemaking. Based on our findings, the 2D-Plot offered an objective and detailed visualisation that especially supported getting a quick overview of stress data. Participants appreciated being presented with the maximum value that their stress level could reach. Still, it remained unclear to them how their stress level has to be interpreted, especially in relation to others. For example, participants reported being unsure about what they should aim for: having a stress level similar to the average of 50% or should it be as low as possible. In contrast, participants reported that the aim becomes clear in VeatherReflect, namely being in good weather. Nevertheless, this visualisation approach provides less granular insights, which made it more challenging for some participants to interpret. For example, participants struggled to understand if they were at the lower or upper end within a specific stress level. Along similar lines, they would have appreciated knowing the number of different weather scenarios as a reference point, as they often only experienced two or three weather scenarios altogether. Some participants also questioned why high stress is always mapped to something with negative connotations. As an example, they mentioned that doing sports is a high-stress but positive activity, which would still be represented with a thunderstorm in VeatherReflect. Moreover, VeatherReflect was considered to offer a more subjective visualisation, thus already interpreting the stress scores to some extent. While some enjoyed having a lower task-load through that, others were of the opinion that VeatherReflect was over-interpreting their stress level. Those participants questioned or rejected the visualisation, as the virtual weather was mapped to a more intense stress level than their subjectively perceived stress level. These participants felt that the VR app not only visualised the different levels but also made an assessment of the stress levels in the process:

> [The 2D-Plot] is a more neutral way of visualising data in the sense that the visualisation itself does not attribute the meaning to the stress level. It does not make a judgement on whether it was good or bad stress. It just shows you the level. The second visualisation [VeatherReflect] certainly does make this judgement, because high stress level equals bad [weather]. Thunderstorms, rain. (P19)

5.2.3 Recollection and Reflection. Both visualisations stimulated participants to think about past potentially stressful situations. For instance, they wondered which situation might be depicted by which bar or by which weather. However, some participants mentioned having different and deeper insights about the origin of their stress through VeatherReflect. It helped them understand that external factors such as places or people influenced their experienced stress level. One participant elaborated:

> I was literally in a stressful environment [in VeatherReflect] that triggered stress in me. So, I think that helped me with reflecting: It was not just me who was stressed, but also part of the environment and the whole surrounding was stressful to me. So I think it [VeatherReflect] provided another [conceptual] level of what I could look at and why I experienced stress. (P5)

Furthermore, participants also shared that they spent more time looking at the 2D-Plot in the study than they would normally do due to being in a study. However, they still mentioned spending more time (and wanting to spend more time) in VeatherReflect. It sparked their interest in their personal data, and triggered reflection about behaviour changes regarding dealing with stress. These two quotes highlight these aspects:

> [VeatherReflect] invites you to more exploration. It triggers more curiosity. [...] And it forces you to spend a little bit more time with your data, but in a nice way. (P10)

> I realised [in VeatherReflect] how far I can go. So, now I think about my boundaries outside of this experiment, how much stress is okay, like [how much] rain [is okay], and when to stop. (P15)

5.2.4 Caring and Sharing. Some participants reported that VeatherReflect made them care more about their personal data; it became a personal and emotional experience. While they remained neutral when analysing the 2D-Plot, through VeatherReflect they could recall moments of stress and reflect upon them, which made them relate more to their personal data:

> VeatherReflect makes me more interested in my data. At least in my case, when I am thinking of a stressful situation, I immediately relate it to my feelings. And my feelings for me are not numbers! I feel more connected [using VeatherReflect]. (P17)

Further, VeatherReflect led to emotional experiences. Some participants described that they became more agitated or more relaxed due to the weather they were exposed to. In addition, participants emphasised that the additional sound made the experience more intense than only visually experiencing the weather around them. In general, participants appreciated the experience very much, with some calling VeatherReflect "a poetic and beautiful experience" (P17) that they wished to experience over and over again. Additionally, some participants expressed the desire to invite friends to VeatherReflect, to show and talk to them about their stress levels and discuss how to find long term stress relief:

> Also showing it to others: 'Look, that is how I felt at that moment. So, I think this is easier to communicate and to relive than rather showing 'I had a stress level of 70.' What does this mean anyway? (P8)

On a general note, eleven participants preferred VeatherReflect over the 2D-Plot while six wanted both. Participants emphasised that both complement each other and can be useful in their own right depending on the context and purpose that the user has in mind. One participant suggested:

> You definitely need both, maybe in a sort of layer thing where the weather gives you a nice overview of the whole day so you can see, ah, this were the moments where I might need to dig deeper. And then, when you go inside this particular time frame, then you could add more details there, and maybe show different visualisations or map it to the weather with more granularity. (P8)

6 DISCUSSION

In this work we endeavoured to understand the effect of presenting stress data with qualitative visualisations such as weather in VR. In this section we first discuss four design recommendations that were derived based on our data. We then reflect on limitations and opportunities for future work.

6.1 Design Recommendations

We investigated how a VR application could be designed when using qualitative instead of quantitative representations of complex personal data. On the one hand, the quantitative plot was considered to be more objective, detailed and helpful to get an overview of one's data. Yet, it falls short in creating interest in understanding personal data, which negatively affected engagement and depth of reflection.

On the other hand, VeatherReflect has shown that weather as metaphor for stress data is intuitively understood and can mitigate most of the aforementioned challenges. Although users intuitively understood their data representation, they missed granularity in the presentations. However, higher granularity does not necessarily equal quantitative values and presentations. Instead, this could be realised via levels regarding the intensity of weather or by including a higher number of weather scenarios. Participants also voiced an interest in rankings and a classification system to fully grasp what they experience and to contextualise and compare their data with other people.

Recommendation 1: Qualitative representations of stress data should be accompanied by a reference point that serves as guide to understanding one's data.

Our findings also revealed some insights into specific aspects that created a significantly more engaging experience (UES-Total: p = .0006). First of all, participants emphasised the importance of feeling empowered when interacting with the environment. In our case, autonomously being able to decide if, when and how to enter one's data generated engagement. Participants desired even more flexibility, especially in regard of the mapping of weather to their stress levels. Future work could investigate the effects on engagement and reflection when users autonomously select or create the weather themselves. This could mitigate the perception of some participants that the system is over-interpreting their stress levels. Empowering users in that way could reduce dissatisfaction and motivate users long-term, as it addresses one of the universal psychological needs, namely autonomy [50, 68]. Besides significantly better scores as measured with the UES, users also spent more time with condition VeatherReflect (2D-Plot: 2:20min vs VeatherReflect: 5:56min), and explored their data more extensively. This led, in turn, to a perceived increase in depth of reflection, e.g. some participants started considering their environment as one possible stress factor due to VeatherReflect.

Recommendation 2: Qualitative personal data representations in VR should provide interaction possibilities and support autonomy.

Interestingly, our findings showed a significant increase in spatial presence (IPQ SP: p = .016) and in mindfulness (SMS-Total: p = .001). Based on literature, presence - the sense of being in the VE [54] - can be increased by addressing sensory, vestibular, proprioceptive and interoceptive channels [10, 47], including haptics [38, 67], olfactory elements [37], and thermoception [53]. However, our environment was quite plain, neither providing high realism [26] nor natural settings [16]. Still, the weather scenarios were designed for high realism and included sound. Although participants only stayed a few minutes in each weather scenario, they reported feeling immersed and allowed the weather to affect them, which is also shown in significant increases in presence and mindfulness. Some participants even experienced physical reactions, showcasing an intuitive understanding of weather as a metaphor for stress. While being immersed in weather was the decisive factor, prior literature has also shown that mindfulness and presence can mutually benefit from each other [72], increasing each others effects. Future VR researchers have to master the dichotomy of providing enough feedback to generate presence while designing for simplicity to set a clear focus on key elements.

Recommendation 3: Qualitative representations of data in VR should focus on increasing presence and mindfulness.

Additionally to the aspects discussed above, participants enjoyed the emotional component they experienced in the condition VeatherReflect. They were emotionally involved by the weather scenarios as qualitative representations of their personal data, becoming relaxed or more stressed by being immersed in the weather, and linked these feelings to their personal data. As a result, they increasingly cared about their stress levels, and also reported getting emotionally attached to the application VeatherReflect itself as it provided an engaging experience. This indicates motivation that could be valuable support for long-term usage of our system.

Recommendation 4: To represent personal data in VR, qualitative metaphors that address users on an emotional level could be particularly valuable.

6.2 Limitations & Future Work

We acknowledge some limitations regarding our study design. First of all, we decided to reset the smartwatch for every participant, to not have relics that might affect the stress level. However, stress level is partly calculated based on the baseline of resting heart rate, which the smartwatch cannot learn to adequately adjust in such a short time frame. Thus, participants' dissatisfaction with inadequate representation of their stress could be partly due to that aspect. Further, our participant sample is still rather small (N=20) and it included participants with basic and also with very little prior experience of engaging with smartwatch data. Having a homogeneous sample was not a pre-requisite for our study, as we did not strive for interpersonal comparability. Instead our aim was to focus on exploring how new ways of representing data in VR can potentially facilitate reflection and engagement. Here, we focused on the individual framing and understanding of participants. Additionally, the participant sample had varying prior knowledge with VR systems. While all have been familiar with bar charts, this might have led to being more excited by VeatherReflect. To counter this potential novelty effect and to increase comparability of conditions, both conditions were conducted in VR. Yet, we cannot rule out some influence of a novelty effect in the VeatherReflect condition for the more inexperienced users. Furthermore, participants only wore the smartwatch for two to three hours to gather data during their normal day. Some participants reported having sat in front of the laptop while others were physically active during that time. Although we did not strive for interpersonal comparability, some participants received more diverse stress data visualisations than others, which could have led to differences in the amount of insights that were possible. Tracking data for a whole day could have mitigated that problem. Based on our results, we recommend exploring individual stress mappings in future work, i.e. that participants create their own weather representation matching the subjective perception of their stress levels. Additionally, in this study, we explored the potential of weather as a cohesive immersive experience. Thus, the weather scenarios that we explored consisted of different modalities such as soundscapes, lighting, and colour saturation. The decision which weather scenario was accompanied by which soundscape, e.g. that sunny weather was mapped to birds' chirping and the thunderstorm to a loud rolling thunder, was based on realistic scenarios. Future work could investigate the influence that each specific factor presents on the dependent variables, with a specific focus on the role of sound.

Future work, while addressing and validating the design recommendations established in this paper, could further research the following ideas. First, as participants reported becoming emotional during the study, feeling stressed in thunderstorms, one could track and compare their stress levels gathered during the study. Second, to provide more granularity when using qualitative representations one could immerse users in weather as a time lapse, gradually showing each spike. While this would mitigate that problem, it removes autonomy and interaction possibilities from users. Hence we recommend careful consideration, although this research approach appears interesting and valuable to be compared to a more active approach.

7 CONCLUSION

This paper explored how personal complex data (i.e. stress level) could be visualised in a qualitative and engaging way in VR, with the aim to foster reflection. To that end, we presented VeatherReflect – a VR application that visualises tracker stress scores using weather as metaphorical means of representation. Quantitative and qualitative results of N=20 participants showed that VeatherReflect significantly increased engagement with personal stress data and enhanced insights into stress-reducing behaviour. We found that participants felt a significant increase of being aware of one-self, higher interest in engaging with data and more motivation to reflect on their data.

Based on our findings, we derived four design recommendations for VR apps that aim to visualise personal data in a qualitative way: Within a framework, they should provide engaging autonomous interaction that support presence, mindfulness and emotionally engaging experiences. Our work can provide a first step in showing that weather can be a meaningful metaphor to represent tracker stress scores, and that VR can be an effective tool to support engaging with complex personal data. We hope that this paper will inspire further inquiry into qualitative visualisation approaches that effectively support well-being.

ACKNOWLEDGMENTS

We thank all our participants for taking part in our study. This research is funded by the German Research Foundation (DFG) under Germany's Excellence Strategy (EXC 2077, University of Bremen), and by the Minds-Media-Machines project (54030724).

REFERENCES

[1] Wadhah Al-Ashwal, Houshyar Asadi, Shady Mohamed, Shehab Alsanwy, Lars Kooijman, Darius Nahavandi, Ahmad Abu Alqumsan, and Saeid Nahavandi. 2021. Cybersickness Measurement and Evaluation During Flying a Helicopter in Different Weather Conditions in Virtual Reality. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2152–2157. https://doi.org/10.1109/SMC52423.2021.9659215

[2] American Psychological Association. 2020. Stress in America™ 2020: A National Mental Health Crisis. Retrieved 11.10.2022 from https://www.apa.org/news/press/releases/stress/2020/report-october

[3] M. Barton. 2008. How's the weather: Simulating weather in virtual environments. 8 (09 2008).

[4] Lyn Bartram. 2015. Design challenges and opportunities for eco-feedback in the home. *IEEE Computer Graphics and Applications* 35, 4 (2015), 52–62.

[5] Eric P.S. Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) *(DIS '14)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/2598510.2598598

[6] Mark E. Beecher, Dennis Eggett, Davey Erekson, Lawrence B. Rees, Jennie Bingham, Jared Klundt, Russell J. Bailey, Clark Ripplinger, Jessica Kirchhoefer, Robert Gibson, Derek Griner, Jonathan C. Cox, and R.D. Boardman. 2016. Sunshine on my shoulders: Weather, pollution, and emotional distress. *Journal of Affective Disorders* 205 (2016), 234–238. https://doi.org/10.1016/j.jad.2016.07.021

[7] Marit Bentvelzen, Jasmin Niess, Mikołaj P. Woźniak, and Paweł W. Woźniak. 2021. The Development and Validation of the Technology-Supported Reflection Inventory. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 366, 8 pages. https://doi.org/10.1145/3411764.3445673

[8] Marit Bentvelzen, Paweł W. Woźniak, Pia S.F. Herbes, Evropi Stefanidi, and Jasmin Niess. 2022. Revisiting Reflection in HCI: Four Design Resources for Technologies That Support Reflection. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 1, Article 2 (mar 2022), 27 pages. https://doi.org/10.1145/3517233

[9] Elisabeth Sophie Blanke, Mirjam Schmidt, Michaela Riediger, and Annette Brose. 2020. Thinking mindfully: How mindfulness relates to rumination and reflection in daily life. *Emotion* (2020).

[10] Corey Bohil, Bradly Alicea, and Frank Biocca. 2011. Virtual reality in neuroscience research and therapy. *Nature reviews. Neuroscience* 12 (11 2011), 752–62. https://doi.org/10.1038/nrn3122

[11] Elisabeth Bos, Rogier Hoenders, and Peter Jonge. 2012. Wind direction and mental health: A time-series analysis of weather influences in a patient with anxiety disorder. *BMJ case reports* 2012 (06 2012). https://doi.org/10.1136/bcr-2012-006300

[12] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[13] Fred B. Bryant, Colette M. Smart, and Scott P. King. 2005. Using the past to enhance the present: Boosting happiness through positive reminiscence. *Journal of Happiness Studies* 6, 3 (2005), 227–260. https://doi.org/10.1007/s10902-005-3889-4

[14] Zikun Chen, Roshan Lalintha Peiris, and Kouta Minamizawa. 2017. A Thermally Enhanced Weather Checking System in VR. In *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 123–125. https://doi.org/10.1145/3131785.3131825

[15] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, and Nathalie Henry Riche. 2017. Understanding self-reflection: How people reflect on personal data through visual data exploration. In *ACM International Conference Proceeding Series (PervasiveHealth '17)*. Association for Computing Machinery, New York, NY, USA, 173–182. https://doi.org/10.1145/3154862.3154881

[16] Karen Cochrane, Lian Loke, Matthew Leete, Andrew Campbell, and Naseem Ahmadpour. 2021. Understanding the First Person Experience of Walking Mindfulness Meditation Facilitated by EEG Modulated Interactive Soundscape. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Salzburg, Austria) *(TEI '21)*. Association for Computing Machinery, New York, NY, USA, Article 18, 17 pages. https://doi.org/10.1145/3430524.3440637

[17] Marie Connolly. 2011. Some Like It Mild and Not Too Wet: The Influence of Weather on Subjective Well-Being. *Journal of Happiness Studies* 14 (07 2011). https://doi.org/10.2139/ssrn.1884996

[18] Lynn Coorevits and Tanguy Coenen. 2016. The rise and fall of wearable fitness trackers. In *Academy of Management*. https://doi.org/10.5465/AMBPP.2016.17305abstract

[19] Jaap Denissen, Ligaya Butalid, Lars Penke, and Marcel Aken. 2008. The Effects of Weather on Daily Mood: A Multilevel Approach. *Emotion (Washington, D.C.)* 8 (11 2008), 662–7. https://doi.org/10.1037/a0013497

[20] Xianghua (Sharon) Ding, Shuhan Wei, Xinning Gui, Ning Gu, and Peng Zhang. 2021. Data Engagement Reconsidered: A Study of Automatic Stress Tracking Technology in Use. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 535, 13 pages. https://doi.org/10.1145/3411764.3445763

[21] Darragh Egan, Sean Brennan, John Barrett, Yuansong Qiao, Christian Timmerer, and Niall Murray. 2016. An evaluation of Heart Rate and ElectroDermal Activity as an objective QoE evaluation method for immersive virtual reality environments. In *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*. 1–6. https://doi.org/10.1109/QoMEX.2016.7498964

[22] Daniel A Epstein, Clara Caldeira, Mayara Costa Figueiredo, Xi Lu, Lucas M Silva, Lucretia Williams, Jong Ho Lee, Qingyang Li, Simran Ahuja, Qiuer Chen, Payam Dowlatyari, Craig Hilby, Sazeda Sultana, Elizabeth V Eikey, and Yunan Chen. 2020. Mapping and Taking Stock of the Personal Informatics Literature. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4 (12 2020). https://doi.org/10.1145/3432231

[23] John M. Grohol. 2014. Can Weather Affect Your Mood? https://psychcentral.com/blog/can-weather-affect-your-mood#1

[24] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. 2014. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics* 21, 3 (2014), 420–433.

[25] Marcus J.H. Huibers, L. Esther de Graaf, Frenk P.M.L. Peeters, and Arnoud Arntz. 2010. Does the weather make us sad? Meteorological determinants of mood and

depression in the general population. *Psychiatry Research* 180, 2 (2010), 143–146. https://doi.org/10.1016/j.psychres.2009.09.016

[26] Jonatan Hvass, Oliver Larsen, Kasper Vendelbo, Niels Nilsson, Rolf Nordahl, and Stefania Serafin. 2017. Visual realism and presence in a virtual reality game. In *2017 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON).* 1–4. https://doi.org/10.1109/3DTV.2017.8280421

[27] Jade Jiang and Naseem Ahmadpour. 2021. Beyond Immersion: Designing for Reflection in Virtual Reality. In *33rd Australian Conference on Human-Computer Interaction* (Melbourne, VIC, Australia) *(OzCHI '21)*. Association for Computing Machinery, New York, NY, USA, 208–220. https://doi.org/10.1145/3520495.3520501

[28] Silvia Jordan, Martin Messner, and Albrecht Becker. 2009. Reflection and Mindfulness in Organizations: Rationales and Possibilities for Integration. *Management Learning* 40, 4 (2009), 465–473. https://doi.org/10.1177/1350507609339687 arXiv:https://doi.org/10.1177/1350507609339687

[29] Alexandra Kitson, Mirjana Prpa, and Bernhard E. Riecke. 2018. Immersive Interactive Technologies for Positive Change: A Scoping Review and Design Considerations. *Frontiers in Psychology* 9 (2018), 1354. https://doi.org/10.3389/fpsyg.2018.01354

[30] Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection Companion: A Conversational System for Engaging Users in Reflection on Physical Activity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 2 (7 2018). https://doi.org/10.1145/3214273

[31] Chongsan Kwon. 2019. Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR technologies. *Virtual Reality* 23, 1 (2019), 101–118.

[32] Sylvia Kämpfer and Michael Mutz. 2013. On the Sunny Side of Life: Sunshine Effects on Life Satisfaction. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement* 110, 2 (2013), 579–595. https://doi.org/10.1007/s11205-011-9945-z

[33] Mihye Lee, Sachiko Ohde, Kevin Y. Urayama, Osamu Takahashi, and Tsuguya Fukui. 2018. Weather and Health Symptoms. *International Journal of Environmental Research and Public Health* 15, 8 (2018). https://doi.org/10.3390/ijerph15081670

[34] Jiwei Li, Xun Wang, and Eduard Hovy. 2014. What a Nasty Day: Exploring Mood-Weather Relationship from Twitter. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management* (Shanghai, China) *(CIKM '14)*. Association for Computing Machinery, New York, NY, USA, 1309–1318. https://doi.org/10.1145/2661829.2662090

[35] Lika Haizhou Liu, Xi Lu, Richard Martinez, Dennis Wang, Fannie Liu, Andrés Monroy-Hernández, and Daniel A. Epstein. 2022. Mindful Garden: Supporting Reflection on Biosignals in a Co-Located Augmented Reality Mindfulness Experience. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing* (Virtual Event, Taiwan) *(CSCW'22 Companion)*. Association for Computing Machinery, New York, NY, USA, 201–204. https://doi.org/10.1145/3500868.3559708

[36] Sonja Lyubomirsky, Kennon M Sheldon, and David Schkade. 2005. Pursuing happiness: The architecture of sustainable change. *Review of general psychology* 9, 2 (2005), 111–131.

[37] Emanuela Maggioni, Robert Cobden, Dmitrijs Dmitrenko, Kasper Hornbæk, and Marianna Obrist. 2020. SMELL SPACE: Mapping out the Olfactory Design Space for Novel Interactions. *ACM Trans. Comput.-Hum. Interact.* 27, 5, Article 36 (Aug. 2020), 26 pages. https://doi.org/10.1145/3402449

[38] Maud Marchal, Gabriel Cirio, Yon Visell, Federico Fontana, Stefania Serafin, Jeremy Cooperstock, and Anatole Lécuyer. 2013. Multimodal rendering of walking over virtual grounds. In *Human Walking in Virtual Environments*. Springer, 263–295. https://doi.org/10.1007/978-1-4419-8432-6_12

[39] Erin Michalak, Edwin Tam, Chinnapalli Manjunath, Anthony Levitt, Robert Levitan, and Raymond Lam. 2005. Quality of Life in Patients with Seasonal Affective Disorder: Summer vs Winter Scores. *Canadian journal of psychiatry. Revue canadienne de psychiatrie* 50 (05 2005), 292–5. https://doi.org/10.1177/070674370505000510

[40] Vivian Genaro Motti. 2019. Wearable health: opportunities and challenges. In *Proceedings of the 13th EAI international conference on pervasive computing technologies for healthcare.* 356–359. https://doi.org/10.1145/3329189.3329226

[41] Jasmin Niess, Kristina Knaving, Alina Kolb, and Paweł W. Woźniak. 2020. Exploring Fitness Tracker Visualisations to Avoid Rumination. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services* (Oldenburg, Germany) *(MobileHCI '20)*. Association for Computing Machinery, New York, NY, USA, Article 6, 11 pages. https://doi.org/10.1145/3379503.3405662

[42] Polly Nugent, Duncan Moss, Rebecca Barnes, and Jenny Wilks. 2011. Clear(ing) space: Mindfulness-based reflective practice. *Reflective Practice* 12 (02 2011), 1–13. https://doi.org/10.1080/14623943.2011.541088

[43] Met Office. 2022. Types of Weather. Retrieved 12.10.2022 from https://www.metoffice.gov.uk/weather/learn-about/weather/types-of-weather

[44] Heather L. O'Brien, Paul Cairns, and Mark Hall. 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies* 112 (2018), 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

[45] Ivan Phelan, Penny J. Furness, Maria Matsangidou, Nathan T. Babiker, Orla Fehily, Andrew Thompson, and Shirley A. Lindley. 2021. Designing effective virtual reality environments for pain management in burn-injured patients. *Virtual Reality* (2021). https://doi.org/10.1007/s10055-021-00552-z

[46] Denise Quesnel, Ekaterina Stepanova, Ivan Aguilar, Patrick Pennefather, and Bernhard Riecke. 2018. Creating AWE: Artistic and Scientific Practices in Research-Based Design for Exploring a Profound Immersive Installation. 1–207. https://doi.org/10.1109/GEM.2018.8516463

[47] Giuseppe Riva, Brenda Wiederhold, and Fabrizia Mantovani. 2019. Neuroscience of Virtual Reality: From Virtual Exposure to Embodied Medicine. *Cyberpsychology, Behavior, and Social Networking* 22 (01 2019), 82–96. https://doi.org/10.1089/cyber.2017.29099.gri

[48] Scott Roberts. 2018. *The Narrative of Weather: Understanding the Impact of Photorealistic Procedural Weather on Narrative in Virtual Reality.* Ph. D. Dissertation. Queensland College of Art. https://doi.org/10.25904/1912/2663

[49] Liad Ruimi, Yuval Hadash, Galia Tanay, and Amit Bernstein. 2019. State Mindfulness Scale (SMS). https://doi.org/10.31231/osf.io/3x7ck

[50] Richard M Ryan and Edward L Deci. 2017. *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* Guilford Publications, New York.

[51] Pedro Sanches, Kristina Höök, Elsa Vaara, Claus Weymann, Markus Bylund, Pedro Ferreira, Nathalie Peira, and Marie Sjölinder. 2010. Mind the Body! Designing a Mobile Stress Management Application Encouraging Personal Reflection. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems* (Aarhus, Denmark) *(DIS '10)*. Association for Computing Machinery, New York, NY, USA, 47–56. https://doi.org/10.1145/1858171.1858182

[52] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 266–281. https://doi.org/10.1162/105474601300343603

[53] Liam Jordan Sheehan, Andre Brown, Marc Aurel Schnabel, and Tane Moleta. 2021. The Fourth Virtual Dimension-Stimulating the Human Senses to Create Virtual Atmospheric Qualities, Vol. 2. Proceedings of the 26th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), 213–222. https://doi.org/10.25455/wgtn.14821329.v1

[54] Mel Slater, R. Lotto, Maria Arnold, and Maria Sanchez-Vives. 2009. How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicologia* 40 (09 2009).

[55] Petr Slovák, Christopher Frauenberger, and Geraldine Fitzpatrick. 2017. Reflective Practicum: A Framework of Sensitising Concepts to Design for Transformative Reflection. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2696–2707. https://doi.org/10.1145/3025453.3025516

[56] Geir Smedslund and Kåre Hagen. 2011. Does rain really cause pain? A systematic review of the associations between weather factors and severity of pain in people with rheumatoid arthritis. *European journal of pain (London, England)* 15 (01 2011), 5–10. https://doi.org/10.1016/j.ejpain.2010.05.003

[57] Tobias Sonne, Timothy Merritt, Paul Marshall, Johanne J. Lomholt, Jörg Müller, and Kaj Grønbæk. 2017. Calming Children When Drawing Blood Using Breath-Based Biofeedback. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 725–737. https://doi.org/10.1145/3064663.3064742

[58] Florian Soyka, Markus Leyrer, Joe Smallwood, Chris Ferguson, Bernhard E. Riecke, and Betty J. Mohler. 2016. Enhancing Stress Management Techniques Using Virtual Reality. In *Proceedings of the ACM Symposium on Applied Perception* (Anaheim, California) *(SAP '16)*. Association for Computing Machinery, New York, NY, USA, 85–88. https://doi.org/10.1145/2931002.2931017

[59] Ursula M. Staudinger. 2001. Life Reflection: A Social-Cognitive Analysis. *Review of General Psychology* 5, 2 (2001), 148–160. https://doi.org/10.1037/1089-2680.5.2.148

[60] Kalliopi-Evangelia Stavroulia and Andreas Lanitis. 2019. Enhancing Reflection and Empathy Skills via Using a Virtual Reality Based Learning Framework. *International Journal of Emerging Technologies in Learning (iJET)* 14 (04 2019), 18. https://doi.org/10.3991/ijet.v14i07.9946

[61] Ekaterina R. Stepanova, John Desnoyers-Stewart, Philippe Pasquier, and Bernhard E. Riecke. 2020. JeL: Breathing Together to Connect with Others and Nature. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 641–654. https://doi.org/10.1145/3357236.3395532

[62] Zhida Sun, Sitong Wang, Wenjie Yang, Onur Yürüten, Chuhan Shi, and Xiaojuan Ma. 2020. "A Postcard from Your Food Journey in the Past": Promoting Self-Reflection on Social Food Posting. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1819–1832. https://doi.org/10.1145/3357236.3395475

[63] Daryl Talks. [n. d.]. How In-Game Weather can Regulate Mental Health. https://www.youtube.com/watch?v=W7G1j4Su6bs

[64] Galia Tanay and Amit Bernstein. 2013. State Mindfulness Scale (SMS): Development and Initial Validation. *Psychological assessment* 25 (09 2013). https://doi.org/10.1037/a0034044

[65] Jennifer G. Tichon and Timothy Mavin. 2019. Using the Experience of Evoked Emotion in Virtual Reality to Manage Workplace Stress: Affective Control Theory (ACT). https://doi.org/10.4018/978-1-5225-8356-1.ch011

[66] Jay Vidyarthi, Bernhard E. Riecke, and Diane Gromala. 2012. Sonic Cradle: Designing for an Immersive Experience of Meditation by Connecting Respiration to Music. In *Proceedings of the Designing Interactive Systems Conference* (Newcastle Upon Tyne, United Kingdom) *(DIS '12)*. Association for Computing Machinery, New York, NY, USA, 408–417. https://doi.org/10.1145/2317956.2318017

[67] Nadine Wagener, Alex Ackermann, Gian-Luca Savino, Bastian Dänekas, Jasmin Niess, and Johannes Schöning. 2022. Influence of Passive Haptic and Auditory Feedback on Presence and Mindfulness in Virtual Reality Environments. *International Conference on Multimodal Interaction (ICMI '22).* https://doi.org/10.1145/3536221.3556622

[68] NADINE WAGENER, JASMIN NIESS, and YVONNE ROGERS. 2022. Considering Fundamental Psychological Needs in Virtual Reality Experiences. (2022).

[69] Nadine Wagener, Jasmin Niess, Yvonne Rogers, and Johannes Schöning. 2022. Mood Worlds: A Virtual Environment for Autonomous Emotional Expression. In *CHI Conference on Human Factors in Computing Systems*. 1–16.

[70] Nadine Wagener, Leon Reicherts, Nima Zargham, Natalia Bartłomiejczyk, Ava Elizabeth Scott, Katherine Wang, Marit Bentvelzen, Evropi Stefanidi, Thomas Mildner, Yvonne Rogers, and Jasmin Niess. 2023. SelVReflect: A Guided VR Experience Fostering Reflection on Personal Challenges. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 323, 17 pages. https://doi.org/10.1145/3544548.3580763

[71] Adrian F. Ward. 2013. Winter Wakes Up Your Mind – and Warm Weather Makes it Harder to Think Straight: How temperature shapes difficult decisions. https://www.scientificamerican.com/article/warm-weather-makes-it-hard-think-straight/

[72] Erik Wolf, Sara Klüber, Chris Zimmerer, Jean-Luc Lugrin, and Marc Erich Latoschik. 2019. "Paint That Object Yellow": Multimodal Interaction to Enhance Creativity During Design Tasks in VR. In *2019 International Conference on Multimodal Interaction* (Suzhou, China) *(ICMI '19)*. Association for Computing Machinery, New York, NY, USA, 195–204. https://doi.org/10.1145/3340555.3353724

[73] Kieran Woodward, Eiman Kanjo, David Brown, Thomas McGinnity, Becky Inkster, Athanasios Tsanas, and Donald Macintyre. 2020. Beyond Mobile Apps: a Survey of Technologies for Mental Well-being. *IEEE Transactions on Affective Computing* (2020), 21. https://doi.org/10.1109/TAFFC.2020.3015018

[74] Soojeong Yoo, Phillip Gough, and Judy Kay. 2018. VRFit: An Interactive Dashboard for Visualising of Virtual Reality Exercise and Daily Step Data. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction (OzCHI '18)*. Association for Computing Machinery, New York, NY, USA, 229–233. https://doi.org/10.1145/3292147.3292193

[75] Bin Yu, Mathias Funk, Jun Hu, and Loe Feijs. 2017. StressTree: A Metaphorical Visualization for Biofeedback-Assisted Stress Management. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 333–337. https://doi.org/10.1145/3064663.3064729

[76] Han Yu, Elizabeth B. Klerman, Rosalind W. Picard, and Akane Sano. 2019. Personalized Wellbeing Prediction using Behavioral, Physiological and Weather Data. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. 1–4. https://doi.org/10.1109/BHI.2019.8834456

[77] Hangyu Zhou, Yuichiro Fujimoto, Masayuki Kanbara, and Hirokazu Kato. 2021. Virtual Reality as a Reflection Technique for Public Speaking Training. *Applied Sciences* 11, 9 (2021). https://doi.org/10.3390/app11093988

[78] Izabela Żołnowska. 2011. Weather as the source domain for metaphorical expressions. *Avant: Journal of the Philosophical-Interdisciplinary Vanguard* 2 (2011), 165–179.