Constructing Metric Ecologies: How Fitness Tracker Users Distinguish Personal Health and Well-Being Metrics

Nina Kersting
Utrecht University
Utrecht
The Netherlands
n.kersting@students.uu.nl

Bradley Rey
University of Winnipeg
Winnipeg
Canada
b.rey@uwinnipeg.ca

Marloes Vredenborg
Utrecht University
Utrecht
The Netherlands
m.t.r.vredenborg@uu.nl

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

Fitness trackers increasingly promise insight into both health *and* well-being. However, these constructs are hard to operationalise given they are not directly measurable via a single metric. Instead, people must interpret combinations of metrics—or *metric ecologies*—to make sense of their health and well-being. Yet, little is known about what metrics form these ecologies in practice, or how people distinguish health from well-being. To address this gap, we conducted an exploratory study with Garmin fitness tracker users ($n = 21$) using gaze-cued retrospective think-aloud sessions. Our findings show that participants treated health and well-being as related but distinct reflective constructs. Health was commonly understood as bodily and measurable, making it easier to interpret through objective metrics. Well-being, by contrast, was understood as emotional and contextual. We contribute empirical insights into how users define health and well-being and the relationship between these constructs, the metrics included and excluded within ecologies, and discuss opportunities for personal informatics systems to better support reflection of health *and* well-being.

Health, Well-being, Personal Informatics, Fitness Trackers, Metric Ecologies, Gaze-cued RTA

1. INTRODUCTION

Wearable devices such as smartwatches and fitness trackers have become embedded in many peoples' everyday lives. In 2022, 41% of US adults reported owning such a device, and 90% of them used it to track fitness or monitor health (Consult 2024). These devices, often described as personal informatics (PI) tools, enable people to collect data about themselves and their activities and to reflect on it as a way of gaining insight and self-knowledge (Li, Dey, and Forlizzi 2010; Choe et al. 2014). For many people, this reflection is closely tied to broader goals of understanding, maintaining, or improving their health and well-being (Choe et al. 2014; Rooksby et al. 2014; Lupton 2017; Niess and Woźniak 2018).

Importantly, fitness trackers do not measure health or well-being directly. Instead, they provide increasingly detailed physiological and behavioural measurements from which people can infer their health, well-being, and other broader understandings of themselves (Niess and Woźniak 2018). To make these inferences, people will in part draw on combinations of available metrics (e.g., steps, heart rate, stress)—creating *metric ecologies* (Bentvelzen, Niess, and Woźniak 2022). Yet, creating meaningful metric ecologies is not straightforward. Prior work focused on general personal health reflection shows that people often encounter conceptual mismatches between the metrics provided by trackers and what they personally consider meaningful or representative of themselves (Bentvelzen, Niess, and Woźniak 2021; Ding et al. 2021).

While this prior work identifies broader difficulties in making personal health data meaningful, less is known about how people construct metric ecologies specifically for health and well-being. This distinction is important because health and well-being, although frequently discussed in conjunction, are conceptually distinct. More specifically, it remains unclear how people generally conceptualise health and well-being, how they map available metrics onto these constructs, and why certain metrics become meaningful indicators of one, the other, or both. Without this understanding, designers risk creating PI systems that offer narrow or predefined interpretations of health and well-being, limiting

© Kersting et al. Published by BCS Learning and Development Ltd.

people's ability to engage in increasingly meaningful reflection and sense-making.

To address this gap, we investigate how fitness tracker users construct and reason through metric ecologies related to health and well-being. Specifically, we explore the following research questions: **RQ1:** How do fitness tracker users conceptually differentiate between health and well-being, and how do they view the relationship between the two? **RQ2:** How do users integrate and interpret tracked metrics within their personal metric ecologies for health and well-being? **RQ3:** What factors lead fitness tracker users to exclude or reject certain tracked metrics from their personal metric ecologies? To address these questions, we conducted a study with Garmin fitness tracker users ($n = 21$) using gaze-cued retrospective think-aloud sessions. Participants first explored their own data naturally within Garmin Connect while their gaze was tracked. These recordings were then replayed to cue participants to articulate their interpretations, reasoning processes, and reflections surrounding their desired metric ecologies.

Our findings show that participants drew a clear distinction between health and well-being within their metric ecologies. Health was typically framed as factual, measurable, and tied to bodily function, making it easier for participants to interpret through existing tracked metrics. In contrast, well-being was described as nuanced, situational, and emotionally grounded, extending beyond numerical representations to include social, environmental, and contextual factors. Interestingly, participants often relied on the same metrics (e.g., sleep or stress scores) across both concepts, but interpreted them differently depending on the broader ecology the metrics were included within. For example, sleep scores were viewed as indicators of physical recovery in relation to health, but as determinants of mood, energy, and daily functioning in relation to well-being. Participants also identified important perceived gaps in current tracking systems, noting that factors such as mood, social interaction, and time outdoors were central to well-being despite being poorly represented in existing PI tools.

Based on our retrospective think-aloud study with $n = 21$ participants, we contribute: (1) empirical insights into how fitness tracker users define health and well-being, (2) metric ecologies showing how participants relate their collected metrics to these concepts, further illustrated through an interactive visualisation; and (3) a discussion of how our findings can inform design of future PI systems.

2. RELATED WORK

2.1. Personal Informatics Models and Reflection

Personal informatics (PI) encompasses practices of collecting, reflecting on, and acting upon personal data to gain self-knowledge and achieve personal goals (Li, Dey, and Forlizzi 2010). Foundational PI models characterise tracking as an iterative process shaped by goals, context, and evolving interpretations of data. For example, Li et al.'s *Stage-Based Model of Personal Informatics Systems* (Li, Dey, and Forlizzi 2010) frames tracking as a progression from collection to reflection and action, while later work such as the *Lived Informatics Model* (Epstein, Ping, et al. 2015) and *Tracker Goal Evolution Model* (Niess and Woźniak 2018) emphasise the dynamic and evolving nature of tracking practices and sense-making.

Across these models, reflection depends on users translating metrics into interpretations that are personally meaningful and actionable. However, this process is often difficult when metrics are abstract, lack contextual grounding, or fail to align with users' goals and lived experiences (Niess and Woźniak 2018; Bentvelzen, Niess, and Woźniak 2021, 2022). Such difficulties can reduce potentially meaningful insights to narrow, metric-specific interpretations that provide limited support for broader understanding or behaviour change (Dijk et al. 2017).

To better characterise these barriers, the *Technology Mediated Reflection Model* (TMRM) was proposed to describe reflection in PI systems as shaped by temporal and conceptual cycles (Bentvelzen, Niess, and Woźniak 2021). Particularly relevant to our work is the notion of *conceptual mismatch*, where users struggle to connect tracked metrics to personally meaningful concepts or goals. Prior work has identified such mismatches in domains including stress tracking, where physiological measures often fail to align with users' lived experiences of stress (Ding et al. 2021). These mismatches can fragment users' understanding of their broader health and well-being, limiting meaningful reflection and actionability (Dijk et al. 2017).

In response, users may adapt, reinterpret, or abandon tracking practices altogether (Rooksby et al. 2014; Bentvelzen, Niess, and Woźniak 2021). Understanding how people combine and relate metrics within broader *metric ecologies* offers one way to better align PI systems with users' own conceptualisations of health and well-being. By examining how users construct meaning across multiple interconnected metrics, designers may better support conceptual alignment. In turn, this alignment can foster facilitated reflection

(Bentvelzen, Niess, and Woźniak 2021), ideally supporting increased actionability of insights with respect to overall health and well-being.

2.2. Metrics in Personal Informatics

PI systems enable people to monitor different aspects of their lives through tracked metrics. Early systems primarily focused on directly measurable indicators such as steps (Lin et al. 2006; Consolvo et al. 2008; Epstein, Caldeira, et al. 2020), heart rate (Han et al. 2015; Liu, Dabbish, and Kaufman 2017), and weight (Eikey and Reddy 2017; Ryan, Linehan, and Dockray 2021). These metrics are comparatively easy to interpret because they can be directly measured, compared against baselines, and tied to concrete goals (Sefidgar et al. 2024). More recently, commercial fitness trackers have introduced increasingly abstract and derived metrics (Bentvelzen, Niess, and Woźniak 2022), including stress levels, recovery scores, body battery, cardio load, and sleep stages. However, while trackers provide access to many individual metrics, users often frame their goals in broader qualitative terms such as "being healthy" or "feeling well" (Niess and Woźniak 2018).

Interpreting broader constructs requires people to combine tracked metrics with lived experience and contextual knowledge (Bentvelzen, Niess, and Woźniak 2022). For example, fertility tracking may involve relating basal body temperature, cycle day, and cervical mucus (Costa Figueiredo et al. 2018), while understandings of sleep quality may emerge through combinations of sleep duration, sleep scores (Liang, Ploderer, and Chapa-Martell 2017), and subjective feelings of restfulness. Similarly, a recent study by Searle et al. (Searle, Potapov, and Ayobi 2025) explored how people understand and conceptualise energy, showing that individuals adopt a personal conception of energy that is expressed through different experiential states. These include energised states such as boredom, interest, creativity, excitement and gratitude. These studies illustrate that people construct *metric ecologies* (Bentvelzen, Niess, and Woźniak 2022; Karahanoğlu, Coskun, et al. 2024; Loerakker et al. 2025): user-constructed combinations of metrics that collectively support understanding of abstract, multifaceted concepts that cannot be represented through any single metric alone.

Recent work has begun to explore how people construct and adapt these ecologies over time. For instance, novice runners may focus on pace or distance, whereas experienced runners combine objective signals with subjective feelings of exertion and historical performance (Karahanoğlu, Coskun, et al. 2024). Similarly, Loerakker et al. examined which metrics fitness tracker users considered meaningful and how those metrics shaped their goals and interpretations (Loerakker et al. 2025). While this work provides insight into metric interpretation and goal-based reflection, less is known about metric ecologies used to reason about the broader concepts of health and well-being.

2.3. Defining Health & Well-being

Engagement with PI systems is often motivated by desires to monitor or improve health and well-being (Choe et al. 2014; Epstein, Ping, et al. 2015; Paré, Leaver, and Bourget 2018; Ayobi, Marshall, and Cox 2020). Commercial fitness trackers similarly position themselves as tools for supporting both concepts. Yet, despite their frequent use, health and well-being remain complex, multifaceted, and difficult to define or measure directly through any single metric. Definitions often overlap: even dictionary definitions of well-being commonly reference being healthy (Merriam-Webster 2024b; Oxford English Dictionary 2025b), while definitions of health similarly invoke well-being (Merriam-Webster 2024a; Oxford English Dictionary 2025a). A widely cited example is the World Health Organisation's definition of health as "a state of complete physical, mental, and social well-being, not merely the absence of disease or infirmity" (WHO 2025), though this definition has also been criticised as difficult to operationalise (Schramme 2023).

This ambiguity extends into HCI and PI research, where health and well-being are frequently invoked but inconsistently defined. Some work distinguishes health as primarily physical and well-being as more psychological or holistic (Stiglbauer, Weber, and Batinic 2019), while other studies use the terms interchangeably (Epstein, Caldeira, et al. 2020). Researchers have also drawn on broader frameworks such as Seligman's PERMA model (Seligman 2011) or notions of "computing for well-being" spanning physical, psychological, and societal factors (Dang et al. 2024). Across this work, however, definitions are often researcher- or system-driven, with limited attention paid to how people themselves conceptualise these constructs (Smits et al. 2022).

This definitional ambiguity creates challenges for both PI research and system design. Without understanding how users themselves conceptualise health and well-being, PI systems risk constraining reflection to narrow, device-generated interpretations. Understanding how people construct meaning around these concepts is therefore essential for designing systems that better support reflection, sense-making, and meaningful action around personal data and the concepts of health and well-being.

Table 1: *An overview of the participants in the study. The duration indicates how long participants had been using their Garmin device prior to the study, while the goal reflects their main motivation for using it.*

PID	Gender	Age	Garmin Model	Tracker Use Duration	Goal	Highest Education, Occupation
P1	Male	49	Vivoactive 4s, Edge 530	4 years	Maintain & improve fitness	Bachelors, Computer Art and Animation
P2	Male	28	Venu 2s	3 years	Track sleep & steps, Maintain & improve fitness	Bachelors, PhD Student
P3	Male	23	Venu 2	3 years	Track steps, Maintain & improve fitness	Bachelors, Public International Law
P4	Female	29	Forerunner 45	2 years	Maintain & improve fitness	Professional Degree, PhD Student
P5	Male	32	Fenix 7	1.5 years	Track steps & fitness age, Maintain & improve fitness	Professional Degree, Operational Specialist
P6	Female	42	Not specified	2 months	Track stress, Maintain & improve fitness	Bachelors, Store Manager
P7	Male	21	Vivosmart HR+	7 years	Track steps	Secondary School, Not provided
P8	Male	23	Forerunner 745, Edge 830	5 years	Track sleep, Monitor sports performance	Bachelors, Not provided
P9	Female	24	Forerunner 55	5 months	Train for running event	Bachelors, Not provided
P10	Male	24	Forerunner 55	7 months	Monitor sports performance	Bachelors, Not provided
P11	Male	42	Vivosmart 4	2 weeks	Track steps, Maintain & improve fitness	Bachelors, Not provided
P12	Female	26	Forerunner 55	4 months	Train for running event	Bachelors, Reservist
P13	Male	58	Fenix 6 Pro	5 years	Maintain & improve fitness	Bachelors, Military
P14	Female	57	Forerunner 965	5 years	Monitor sports performance	Bachelors, Secretary and Pedicurist
P15	Male	30	Forerunner 935	4 years	Maintain & improve fitness, Monitor performance	Professional Degree, PhD Student
P16	Female	31	Forerunner 45	2.5 years	Maintain & improve fitness, Monitor performance	Bachelors, Not provided
P17	Female	20	Forerunner 165	2 weeks	Train for running event	Secondary School, Student
P18	Male	25	Fenix 7	2 years	Track Sleep	Vocational School, Construction
P19	Male	25	Venu 2S	2 years	Maintain and improve well-being	Professional Degree, Software Consultant
P20	Female	54	Venu 4s	1.5 years	Track Fitness Age, Maintain & improve fitness	Vocational School, Office clerk
P21	Female	28	Vivoactive 5	5.5 months	Monitor sports performance, Gain health insights	Professional Degree, Data Scientist

3. STUDY METHOD

To examine how fitness tracker users define health and well-being and which metrics they use when reflecting on these constructs, we conducted gaze-cued retrospective think-aloud (RTA) sessions.

Gaze-cued RTA, enabled by eye-tracking, allows participants first to complete tasks in silence and then later reflect on their interactions while viewing their gaze data (Hertzum 2024). By removing the need to articulate thoughts in real time, this method minimises cognitive load and preserves the natural flow of interaction (Hertzum 2024). Furthermore, compared to concurrent think-aloud (CTA), gaze-cued RTA prompts reflections that extend beyond immediate task performance, offering deeper insights into participant reasoning (Page and Rahimi 1995). We note, gaze data was used strictly as a retrospective cue and for qualitative analysis; no quantitative eye-tracking metrics were processed.

3.1. Participants

We recruited $n = 21$ participants through purposive and voluntary sampling. Specifically, we distributed a recruitment poster both physically on our University's campus and digitally via social media. The study included only participants who had used a Garmin fitness tracker for at least two weeks, either currently or in the past, to ensure sufficient activity data. Full familiarity with all device features was not required. Participants ranged in age from 20–58 years (M = 32.90, SD = 12.23); twelve identified as male and nine as female. Table 1 summarises participant details. We compensated participants by entering them into a raffle for one €50 gift card, redeemable at an online retailer in our country.

3.2. Apparatus

3.2.1. Garmin Fitness Tracker & Garmin Connect. Participants tracked data on their own Garmin fitness trackers and used Garmin Connect[1], a web platform that provides access to the data collected, within the study. Garmin Connect offers a customisable dashboard that displays up to eight fitness metrics. The web interface also allows people to navigate to three pre-defined categories of data: *Activities*, *Health*, and *Performance*. Together, alongside visualisations enabling filter of different time periods (day, week, month, year), participants had multiple ways to review and compare aspects of their health and well-being. Table 2 summarises all available metrics and their descriptions.

3.2.2. Eye-tracking Device.
We used the iMotions Smart Eye AI-X[2], a video-based infrared eye tracker (60 Hz), to capture participants' gaze. Although the sampling rate is relatively modest, it was sufficient for producing gaze replay videos, which served as cues in the RTA sessions (Carter and Luke 2020). As our eye tracker is incompatible with smartphones, participants accessed Garmin Connect via its web interface. Other platforms (e.g., Fitbit and Apple) only support mobile app access and were therefore incompatible with our setup. The eye tracker was mounted below a 24-inch (1920 × 1080) monitor, as recommended by the manufacturer, and connected to a Windows 10 desktop. This unobtrusive configuration required no wearable equipment (e.g., glasses-based eye trackers), supporting a natural browsing experience.

[1] `https://connect.garmin.com`
[2] `https://imotions.com/products/hardware/smart-eye-ai-x/`

Table 2: *An overview of the metrics available in Garmin Connect, grouped into activities, health, and performance as aggregated by Garmin. In addition, we provide a short description for each metric.*

Metric	Description
Activities	
All activities	Summary of all recorded physical activities, including type, duration, and calories burned
Floors	Counts number of floors climbed based on elevation gain
Intensity minutes	Minutes spent in moderate or vigorous physical activity per week
Steps	Counts daily steps taken
Health Statistics	
Blood pressure	Manual logging of blood pressure data or integration with compatible cuffs
Body battery	Estimates body energy levels using heart rate, stress, and activity data
Calories	Estimates calories burned throughout the day based on activity and basal metabolism
Fitness age	Estimates a person's fitness age based on VO2 max and activity
Health snapshot	A quick scan of key health stats like heart rate, respiration, stress, and HRV
Heart rate	Continuous heart rate monitoring from the wrist
Hydration	Tracks water intake manually or through reminders
Menstrual cycle	Manual logging of menstrual cycle phases, including symptoms, moods, and physical changes
Pulse ox	Measures blood oxygen saturation levels (SpO2)
Respiration	Tracks breaths per minute, typically during sleep or yoga activities
Sleep	Tracks total sleep, sleep stages (light, deep, REM), and wake times
Stress	Estimates stress levels based on heart rate variability
Weight	Allows manual logging or sync from smart scales
Performance Statistics	
HRV status	Analyse heart rate variability during sleep to assess recovery and stress
Race prediction	Estimates finish times for common race distances (5K, 10K, half marathon, marathon) based on VO2 max and fitness data trends
Training load	Measures the total training effort over the past 7 days, based on EPOC (excess post-exercise oxygen consumption)
Training status	Evaluates recent exercise history and performance indicators (e.g., VO2 max) to determine if a person is training productively, maintaining, or overreaching
VO2 max	Estimates maximum oxygen uptake during exercise, indicating cardiorespiratory fitness

3.3. Procedure

To start each session, participants were welcomed into the laboratory, provided informed consent, and completed a pre-study survey (available in Appendix A). Participants then logged into their Garmin Connect accounts on the Garmin Connect website.

The researcher then introduced the eye-tracking setup and calibrated the system using iMotions' built-in calibration procedure. Participants sat approximately 59–70 cm from the screen in accordance with manufacturer recommendations. Calibration continued until an "excellent" rating was achieved; when this was not possible due to factors such as thick glasses or contact lenses, a lower but sufficiently accurate calibration was accepted.

Participants were asked to explore their entire Garmin Connect interface to build an overview of their health and well-being. During this exploration, participants were not required to include all or even any metrics within either ecology if they did not want. Participants completed two separate exploration tasks: one in which they explored their data from a health perspective and one in which they explored it from a well-being perspective. We counterbalanced construct order across participants, such that half explored health first and half explored well-being first. For each construct, participants explored their data for three minutes while gaze data was recorded. We selected a three-minute duration based on pilot feedback indicating it reduced participant burden compared to longer exploration periods, while still offering enough time to explore desired metrics.

Immediately after exploring each construct, participants watched a screen recording of their session with gaze overlaid (a screenshot for reference is available in Appendix B) and then verbally reflected on their reasoning and sense-making processes. Participants could pause, rewind, and adjust playback speed during this RTA session. Using guiding questions (available in Appendix A), we focused on how participants interpreted and combined metrics in relation to either concept, why certain metrics were prioritised, and which potentially relevant metrics they felt were missing. Because the study was exploratory, the interview protocol remained flexible and included follow-up questions based on participant behaviour and responses during the session.

Once both concepts of health and well-being had been explored, participants compared the metrics they associated with health and well-being and discussed how they distinguished between the two constructs. For reference, the complete protocol is available in Appendix A.

3.4. Data Analysis

In total, we collected 14 hours and 57 minutes of audio recordings, transcribing them verbatim. After familiarisation with the data, the first author coded a 10% sample of the data using an open coding procedure. Afterwards, through iterative discussions among the first and last author, an initial coding tree was developed and refined, which was then used by the first author to code the remaining transcripts. After the full dataset was coded, we first revisited the coded material to examine how participants defined health and well-being and how they interpreted metrics. We then conducted a broader thematic analysis (Blandford, Furniss, and Makri 2016) to identify recurrent themes that captured participants' perspectives on health and well-being in the context of their collected metrics.

4. FINDINGS

In this section, we first describe how participants distinguished and related health and well-being as overarching concepts. We then highlight incorporation of individual metrics and how metrics gain meaning within an ecology. Finally, we share perceived metric gaps within current fitness trackers and the conceptual mismatches that disrupted participants' inclusion of metrics and thus their sense-making of their health and/or well-being.

4.1. Relationship Between Health and Well-being

4.1.1. General Interpretations of Health

Participants consistently distinguished between health and well-being, constructing different conceptualisations around each. Health was typically framed as physical, measurable, and objective, whereas well-being was described as emotional, contextual, and subjective.

A common idea was that health is quantifiable, allowing it to be expressed and directly inferred from *hard* values. Because it could be captured in numbers, health becomes something measurable. Since it is possible to define specific measurable values for health, our participants considered health to be a more objective concept, something that would apply equally to two different people. They described health in terms of physical performance (P13), fitness (P18), resilience to illness (P16), and recovery (P3), but also with being able to do daily activities without restrictions (P19). As P21 explained: *"I think health is generally more factual, so it's something that's more easily measurable. Like, you have a healthy weight or not, you have a certain level of fitness or not... those kinds of things are just very measurable and clearly either healthy or not."*

While factual for all (i.e., similar data can inform health outcomes), health was seen as something individual in the sense that it focuses on the condition of one's own body. Specifically, one's health is a physical state that does not include others, and cannot be shared. As P14 described: *"Maybe health is kind of selfish - it's all about you. You can't say, "My ear hurts, can you carry some of that pain for me?" That doesn't work. Health is very individual."*

4.1.2. General Interpretations of Well-being

In contrast to health, which was defined as fact-based and physical, well-being was viewed as nuanced and more difficult to define through metrics alone. Participants commonly explained it as a *"general feeling"* (P19), a *"mental state"* (P6), or *"feeling comfortable in my own skin"*. Unlike health, which participants felt could be inferred more directly from physiological data, understanding well-being required additional interpretation and contextualisation of lived experiences and emotions. As P8 explained: *"With health, I think it's easier to define exactly what counts as health, while well-being is a much more open concept. With well-being, in general, you can't measure it directly; you have to interpret it a bit more."* Participants therefore described understanding well-being as requiring additional sense-making to connect their feelings and experiences to their data.

Similar to health, participants also viewed well-being as highly personal and individualised. While they believed the same factors could generally determine whether two people were healthy, they felt that well-being depended on different factors for different people. As P8 further reflected: *"For well-being, there are more factors that come into play, and different ones for different people."*

Furthermore, participants often framed well-being as situated and relational. Participants described factors such as social relationships, weather, and time spent outdoors as important contributors to feeling well. Consequently, many questioned whether fitness trackers could fully capture well-being using physiological metrics alone, with many participants having smaller ecologies for well-being than health. As P10 explained: *"A few [metrics] can affect my well-being, and I look at those, but I know it's not the full picture because well-being depends on so many other things that aren't found here."*

4.1.3. Health and Well-being are Distinct but Interconnected

Although participants clearly distinguished health and well-being, they also described them as interconnected and having an affect on one another. As P12 explained: *"Health is physical, and well-being is mental. And if I'm not doing well mentally, I can improve that by taking care of myself physically. And if I take care of myself physically, then that helps me mentally, too. It goes both ways..."* Moreover, participants explained that being physically unwell for a long time can lead to feeling mentally worse, as it might limit activities in daily life. On the other hand, they noted that struggling with well-being, such as feeling stressed for a long time, can also lead to physical issues, such as feeling tired (P4).

Some participants did share that good well-being does not necessarily require living according to what is considered healthy. As P19 explained, despite engaging in behaviours considered unhealthy, a person might still experience a high level of well-being: *"Being unhealthy is more like eating a lot, not moving - that's what people call an unhealthy lifestyle. And eventually, that has a negative impact on your health. But during an unhealthy lifestyle, you could still have a good level of well-being. You could be happy, for example."*

Although the relationship between health and well-being was seen as dynamic and reciprocal, a clear direction of influence was apparent overall; most often, participants described health as a contributing condition for well-being, with well-being emerging as an outcome of being healthy. This idea was best put by P20, quoting Arnold Schwarzenegger: *"Only in a healthy body lives a healthy mind."*

4.2. Metrics Utilised Within Ecologies

Participants constructed highly individualised metric ecologies that differed in the metrics included, how those metrics were interpreted, and how they were related to health and well-being. To support exploration of all ecologies collected within our study, we developed an interactive visualisation[3] showing how participants associated metrics with health and well-being. Screenshots of the interactive tool are available in Appendix C.

4.2.1. Metric Association

Across participants, 17 of the 22 metrics appeared in both health and well-being ecologies, meaning that each of these metrics was associated with health by at least one participant and with well-being by at least one participant. However, and importantly, the same metric often took on different meanings depending on the conceptual framing participants applied. For example, participants associated physical activity metrics with health when interpreted as indicators of bodily functioning or physiological improvement, but associated those same metrics with well-being when interpreted through emotional or experiential outcomes. For example, P4 connected exercise to health through blood sugar regulation and physical functioning, whereas P3 connected exercise to well-being through improved mood and emotional state.

Participants generally associated physiological measures such as heart rate, blood pressure, VO2 max, and steps with health because these metrics were interpreted as objective indicators of bodily functioning. In contrast, increasingly derived metrics such as stress, sleep, and body battery were more frequently associated with well-being. Participants connected these to energy levels, mood, anxiety, and day-to-day functioning. Importantly, these metrics were rarely interpreted in isolation. Instead, participants related multiple metrics to one another when reasoning about their experiences. For example, participants frequently connected sleep, stress, and body battery together to explain fluctuations in mood, energy, or emotional resilience throughout the day.

4.2.2. Metric Interpretation

Metrics became a meaningful part of an ecology when participants could connect them to lived experiences such as illness, stress, overtraining, recovery, or emotional events. Several participants described how changes in metrics only became interpretable when they could remember what had happened on a particular day. Conversely, metrics often lost meaning when participants could not relate them back to specific experiences or contexts.

[3] `https://healthandwellbeing.github.io`

Participants questioned metrics that failed to align with their lived experiences or personal understandings of health and/or well-being. Composite metrics such as body battery and fitness age were often viewed as helpful, yet participants simultaneously questioned whether complex experiences such as stress, recovery, or ageing could be reduced to single numerical scores, and thus appropriately incorporated within their ecologies. Similarly, body mass index was frequently criticised for failing to represent participants' lived bodily experiences, particularly for those with greater muscle mass. These critiques demonstrate that metrics become a meaningful part of one's metric ecology not simply because they are available, but because people feel they meaningfully represent their bodies, experiences, and goals.

4.3. Expanding Metric Ecologies for Well-being Reflection

While participants felt that available metrics largely supported their health metric ecologies, participants consistently described current metrics as offering incomplete representations of well-being and thus an incomplete well-being metric ecology. While physiological metrics were seen as relatively effective for reflecting aspects of physical health, participants felt that important emotional, social, and contextual dimensions of well-being were poorly represented or entirely absent.

Participants suggested several additional metrics that they believed would better support reflection on well-being. Many desired metrics focused on emotional or psychological experiences, including mood tracking, dream activity, and measures capable of distinguishing psychological stress from physical exertion. Others focused on contextual or relational factors such as time spent outdoors, sunlight exposure, social interaction, nutrition, and daily habits. Participants believed these factors strongly influenced how they felt, yet remained largely invisible within current tracking systems.

Importantly, participants did not necessarily expect trackers to produce definitive measures of well-being. Instead, they viewed these additional metrics as prompts for reflection and contextual interpretation. Several participants described how combining subjective inputs such as mood with existing physiological metrics could help them better understand patterns in their energy, stress, or emotional state. These findings suggest that participants conceptualised well-being through broader ecologies extending beyond physiological sensing alone.

4.4. Metric Exclusions Within Ecologies

Participants did not automatically find all metrics meaningful, and thus some metrics were not

included within formed ecologies. Instead, metrics became irrelevant when they failed to support participants' sense-making around health and well-being. These conceptual mismatches emerged when metrics did not align with participants' goals, lived experiences, interpretations, or expectations.

One common source of mismatch occurred when metrics lacked personal relevance. Even accurate data was often dismissed if it did not relate to participants' goals or experiences. For example, participants ignored calorie metrics when they were not actively managing diet or exercise goals, and some viewed metrics such as menstrual cycle as unimportant when they did not perceive meaningful variation in the data. Metrics also lost relevance when they conflicted with participants' own conceptualisations of health and well-being. For instance, P21 rejected health snapshot because they believed health should be understood longitudinally rather than through isolated measurements: *"I don't link health to a single moment... It's just one moment. So I think, what am I supposed to do with this?"*

Participants additionally struggled to interpret metrics when they could not connect them to lived experiences or contextual events. This was more impactful for metrics considered for a well-being ecology. Unlike health metrics, which were often viewed as more directly measurable, well-being metrics required participants to remember and contextualise what had happened on a given day. When participants could not reconstruct these contexts, the data often became meaningless. Similarly, metrics lost relevance when participants' expectations did not match the data presented. For example, P14 expected their body battery to reflect predictable weekly routines, but found the data difficult to interpret when these anticipated patterns were unclear.

Another source of mismatch involved trust and transparency. Participants questioned metrics when data was incomplete, inaccurate, or generated through opaque calculations. Missing sleep data due to inconsistent watch-wearing, manually logged metrics such as hydration, or stress values that appeared overly sensitive all reduced trust in the resulting interpretations. Participants were particularly sceptical of derived metrics such as body battery and fitness age when they did not understand how these values were calculated. As P8 explained regarding Fitness Age: *"I have no idea how it's calculated, the metric means nothing to me."*

Finally, some metrics became irrelevant because participants lacked familiarity with them or found their meaning too ambiguous. For example, participants questioned whether metrics such as floors climbed or steps taken could reliably indicate health or well-being when identical values could emerge from very different activities and experiences. This highlights that metrics only became meaningful to one's health and/or well-being when participants could connect them to their goals and lived experiences.

5. DISCUSSION

5.1. Health and Well-being are Distinct Reflective Constructs

A central contribution of this study is showing that participants did not treat health and well-being as interchangeable constructs for reflection, despite their frequent conjunction in HCI and PI research (Epstein, Caldeira, et al. 2020). Participants described health as bodily, factual, and measurable, often grounding it in objective values such as resting heart rate, weight, steps taken, and VO2 max. This aligns with definitions of health as a person's physical condition (Oxford English Dictionary 2025a) and with PI work that frames health as primarily physical (Stiglbauer, Weber, and Batinic 2019). By contrast, well-being was overall described as experiential, emotional, and context-dependent, involving metrics such as stress, sleep score, and body battery. This resonates with definitions of well-being as comfort or happiness (Oxford English Dictionary 2025b) and with PI work that frames well-being as psychological and situated (Stiglbauer, Weber, and Batinic 2019).

Through these findings, we can further support an operational distinction for future PI research and design. Specifically, health can be understood as the perceived physical condition and capability of the body, inferred through physiological and behavioural metrics. Well-being, in contrast, can be understood as a broader feeling of functioning well across emotional, social, and environmental contexts, informed by physiological signals but not necessarily reducible to them. While this departs from broad and conjoined definitions (Epstein, Caldeira, et al. 2020; WHO 2025), our findings show that participants clearly delineated the boundary between health and well-being in practice.

Importantly, this distinction was not only reflected in which metrics participants included in each ecology, but also in how they interpreted the metrics. When the same metric appeared in both health and well-being ecologies, its meaning shifted depending on the construct being reflected on. This suggests that fitness tracker metrics are not inherently health- or well-being-oriented; rather, their reflective value is construct- and ecology-sensitive, emerging through the role they play within a user's metric ecology.

5.2. Supporting Health as a Precursor to Well-being

Although participants distinguished health from well-being, they did not treat them as unrelated. A notable pattern in our findings is the directional relationship from health to well-being. Participants consistently described health as foundational to broader well-being. This directional nature indicates that PI systems may benefit from supporting users in reflecting on health as a precursor or contributor to well-being, rather than treating these constructs as either equivalent or entirely unique. For example, tracking metrics like all activities, sleep, steps, or heart rate may facilitate reflection on physical health, which can then be the first step to informing an understanding of subjective well-being, such as increased body battery or reduced stress.

Current dashboard designs within mobile health applications do little to support this kind of directional reasoning. Garmin Connect's "Health Statistics," for example, presents metrics such as steps, heart rate, calories, stress, sleep, body battery, and many others (see Table 2) as equal and adjacent to one another, rather than as supporting separate constructs or as elements in a structural relation. Similarly, Apple Health[4] allows users to explore "Mental Well-Being", recognising that this is shaped by an ecology of metrics. However, Apple predefines this ecology through exercise minutes, time in daylight, state of mind, and sleep. Importantly, this data is presented to the user based on when the data was collected and not presented as a structural relationship from health-driven metrics to well-being.

These findings suggest an opportunity to extend PI models and interfaces with the notion of *reflective hierarchy*. Rather than treating reflection as a singular process, with many metrics presented at once, systems could scaffold users from health focused metrics toward well-being focused metrics. For example, steps and heart rate could support reflection on health, which could then be connected to downstream reflection on body battery, stress, and overall mood. Importantly, this hierarchy should not be implemented as rigid or entail metric exclusivity. A useful design would therefore communicate dependency and progression while still allowing users to decide how particular metrics contribute to their own health and well-being ecologies.

5.3. Health is Represented More Adequately than Well-being

Our findings highlight a current challenge with the metrics offered to fitness tracker users. Participants described health as more factual and quantifiable while well-being was seen as more emotional, contextual, and difficult to define. This created an imbalance in metric inclusion within ecologies, as seen in the majority of metrics being used more often in health ecologies than well-being ecologies. Given currently available metrics, trackers are well suited to capturing bodily signals and activity patterns, but less suited to capturing the everyday experiences that shape well-being. Well-being can be informed by data, but current devices provide stronger evidence about perceived health than well-being.

This, in part, helps explain why participants found well-being harder to interpret. Metrics such as sleep score, stress, and body battery can be useful, but they often capture physiological responses rather than their causes or meanings. For example, a spike in captured stress could be due to many factors. Without additional context, users must infer what the data means. As participants developed their ecologies, we found that they were not looking for a definitive well-being score, but for support in connecting physiological signals to their well-being.

Future work in PI literature needs to further explore how to connect tracked data with subjective and contextual information. Systems could invite lightweight self-reports when unusual patterns appear, or help users relate changes in physiology to automatically tracked calendar events, locations, activities, or social situations. The challenge, however remains, in that we need to elicit useful context with minimal burden. This is important because adding more data is not always beneficial if poorly framed to the user or collected.

5.4. Metric Ecologies are Lived and Personal

Our findings also highlight the personalised nature of health and well-being metric ecologies. Participants' interpretations were shaped by individual experiences, goals, routine, trust in the data, and situational contexts. While some broader patterns were visible, no single definition of health or well-being fully captured how participants made sense of these constructs. Metrics became meaningful, and thus included within an ecology, when they were familiar, had accurate and transparent data collection, were easy to interpret, aligned with personal goals and lived experiences, and/or resonated with concepts of either health or well-being. This suggests that metric relevance is not a property of the metric alone, but emerges based on other factors.

Current dashboards do not always offer means for customisation. For example, as introduced earlier, Apple Health provides a metric ecology for "mental well-being", however, this is predefined with no customisation afforded. While the use of an ecology

[4] As seen in iOS 26.4.2, May 2026

is a appropriate, our findings suggest that forcing a predefined ecology risks misaligning with fitness tracker users' own interpretations. Simply put, customised dashboards that recognise and enable the creation of personal metric ecologies is a must within PI systems. We must afford flexible interfaces that let users define, revise, and connect metrics, thus creating metric ecologies, that matter for their own understandings. This aligns with a recent study (Ruvimova, McGrenere, and Fritz 2026), which compared two interfaces for holistic focus-tracking in the knowledge workplace and argued that focus dashboards should offer multiple visualisations to suit different user preferences and goals.

Meaningful reflection on health and well-being thus requires interfaces that can accommodate individual preferences, changing goals, and subjective experience. In practice, this could involve allowing people to annotate metrics with situational context or to reflect on subjective feelings alongside data. This is especially important as metric ecologies may shift over time alongside changing goals, life circumstances, and health concerns. PI systems should therefore support metrics moving in and out of importance, rather than assuming that the same fixed set of measures will remain meaningful for all users.

5.5. Limitations & Directions for Future Work

Although our study was carefully designed, we recognise some limitations. First, we focused on how people distinguish between health and well-being and how this distinction shapes their inclusion of metrics within their ecologies. However, we did not examine all factors that may influence metric selection and tracking practices. Prior work shows that factors such as metric knowledge (Choe et al. 2014; Ding et al. 2021; Karahanoğlu, Gouveia, et al. 2021; Karahanoğlu, Coskun, et al. 2024), motivation (Gouveia and Epstein 2023; Karahanoğlu, Coskun, et al. 2024), and age (Goodyear and Armour 2018) can shape engagement with personal data and reflection. Future work could therefore examine how these factors affect people's perceived relevance, trust, and interpretation of health and well-being metrics.

Second, our study focused on Garmin users because the eye-tracking hardware and software available in our lab required desktop-based interaction with a web dashboard. This excluded fitness tracker brands without comparable web-based platforms, such as Fitbit and Apple. As a result, we recognise our findings are shaped by Garmin's specific metrics and visualisations. Notably, however, many of the metrics offered by Garmin are also common across other fitness tracker ecosystems. In addition, most participants used their Garmin trackers to support general fitness goals, as well as to track steps and sleep (see Table 1, Goal column), suggesting that our findings still speak to common forms of fitness tracker use seen by users of other brands. Future work could extend this study by comparing how users construct health and well-being metric ecologies across different tracker ecosystems.

Finally, our sample may not represent the full diversity of fitness tracker users. Most participants had used their tracker for more than two years and had health or fitness-related goals. Since novice and experienced users can reflect on data differently (Rapp et al. 2018), our findings may not capture all user perspectives, although our sample did include some beginners. In addition, all participants were able-bodied, which may have shaped how they conceptualised health and well-being. Prior work shows that people with disabilities may understand health, well-being, and life satisfaction differently, and that these understandings can be misaligned with objective health measures or levels of disability—a phenomenon known as the disability paradox (Albrecht and Devlieger 1999; Ubel et al. 2005; Drum, Horner-Johnson, and Krahn 2008). This is particularly relevant to our finding that participants often treated health as a prerequisite for well-being, which may not generalise beyond able-bodied fitness tracker users. Future research should therefore examine how more diverse populations conceptualise health and well-being in relation to fitness tracker metrics.

6. CONCLUSION

Fitness trackers increasingly aim to support reflection of both health *and* well-being. These constructs are often treated as interchangeable in PI literature and commercial application design, without clear direction as to what metrics should be incorporated within either construct. Through a retrospective think-aloud study with 21 Garmin users, we examined how people define health and well-being and how they integrate available metrics into metric ecologies for these constructs. Our findings show that participants understood health and well-being as related but distinct: health was commonly described as bodily, factual, and measurable, while well-being was described as experiential, emotional, contextual, and harder to capture through metrics alone. The collected ecologies further reveal that metric meanings are not fixed; the same metric may support different interpretations if being considered for health or well-being. Finally, we highlight that designing for meaningful reflection of health and well-being requires supporting reflective hierarchies between the two constructs, an increased set of metrics for well-being reflection, and personalisation of ecologies that reflect peoples individual understanding of their own health and well-being.

REFERENCES

Albrecht, Gary L and Patrick J Devlieger (1999). "The disability paradox: high quality of life against all odds". In: *Social science & medicine* 48.8, pp. 977–988.

Ayobi, Amid, Paul Marshall, and Anna L. Cox (Apr. 2020). "Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: Association for Computing Machinery, pp. 1–15. ISBN: 978-1-4503-6708-0.

Bentvelzen, Marit, Jasmin Niess, and Paweł W. Woźniak (May 2021). "The Technology-Mediated Reflection Model: Barriers and Assistance in Data-Driven Reflection". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: Association for Computing Machinery, pp. 1–12. ISBN: 978-1-4503-8096-6.

– (Dec. 2022). "Designing Reflective Derived Metrics for Fitness Trackers". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6.4, pp. 1–19. ISSN: 2474-9567.

Blandford, Ann, Dominic Furniss, and Stephann Makri (2016). *Qualitative HCI Research: Going Behind the Scenes*. Synthesis Lectures on Human-Centered Informatics. Cham: Springer International Publishing. ISBN: 978-3-031-01089-7 978-3-031-02217-3.

Carter, Benjamin T. and Steven G. Luke (Sept. 2020). "Best Practices in Eye Tracking Research". In: *International Journal of Psychophysiology* 155, pp. 49–62. ISSN: 01678760.

Choe, Eun Kyoung et al. (Apr. 2014). "Understanding Quantified-Selfers' Practices in Collecting and Exploring Personal Data". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. New York, NY, USA: Association for Computing Machinery, pp. 1143–1152. ISBN: 978-1-4503-2473-1.

Consolvo, Sunny et al. (2008). "Activity Sensing in the Wild: A Field Trial of Ubifit Garden". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. New York, NY, USA: Association for Computing Machinery, pp. 1797–1806. ISBN: 9781605580111.

Consult, Morning (2024). *Use of Health Apps and Wearables Is Up. But Digital Health Still Has Room to Grow*. https://pro.morningconsult.com/trend-setters/use-of-digital-health-apps-wearables-growing. Accessed: 2024-12-31.

Costa Figueiredo, Mayara et al. (Nov. 2018). "Engaging with Health Data: The Interplay Between Self-Tracking Activities and Emotions in Fertility Struggles". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW, pp. 1–20. ISSN: 2573-0142.

Dang, Ting et al. (Oct. 2024). "WellComp 2024: Seventh International Workshop on Computing for Well-Being". In: *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Melbourne VIC Australia: ACM, pp. 980–983. ISBN: 979-8-4007-1058-2.

Dijk, Elisabeth T. Kersten-van et al. (2017). "Personal Informatics, Self-Insight, and Behavior Change: A Critical Review of Current Literature". In: *Human–Computer Interaction* 32.5-6, pp. 268–296. eprint: https://doi.org/10.1080/07370024.2016.1276456.

Ding, Xianghua (Sharon) et al. (May 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*. CHI '21. New York, NY, USA: Association for Computing Machinery, pp. 1–13. ISBN: 978-1-4503-8096-6.

Drum, Charles E., Willi Horner-Johnson, and Gloria L. Krahn (2008). "Self-rated health and healthy days: Examining the "disability paradox"". In: *Disability and Health Journal* 1.2, pp. 71–78. ISSN: 1936-6574.

Eikey, Elizabeth V. and Madhu C. Reddy (2017). ""It's Definitely Been a Journey": A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, pp. 642–654. ISBN: 9781450346559.

Epstein, Daniel A., Clara Caldeira, et al. (Dec. 2020). "Mapping and Taking Stock of the Personal Informatics Literature". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4.4, pp. 1–38. ISSN: 2474-9567.

Epstein, Daniel A., An Ping, et al. (Sept. 2015). "A Lived Informatics Model of Personal Informatics". In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '15. New York, NY, USA: Association for Computing Machinery, pp. 731–742. ISBN: 978-1-4503-3574-4.

Goodyear, Victoria A. and Kathleen M. Armour (Aug. 2018). "Young People's Perspectives on and Experiences of Health-Related Social Media, Apps, and Wearable Health Devices". In: *Social Sciences* 7.8, p. 137. ISSN: 2076-0760.

Gouveia, Rúben and Daniel A. Epstein (Apr. 2023). "This Watchface Fits with My Tattoos: Investigating Customisation Needs and Preferences in Personal Tracking". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.

CHI '23. New York, NY, USA: Association for Computing Machinery, pp. 1–15. ISBN: 978-1-4503-9421-5.

Han, Teng et al. (2015). "Balancing Accuracy and Fun: Designing Camera Based Mobile Games for Implicit Heart Rate Monitoring". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, pp. 847–856. ISBN: 9781450331456.

Hertzum, Morten (Aug. 2024). "Concurrent or Retrospective Thinking Aloud in Usability Tests: A Meta-Analytic Review". In: *ACM Transactions on Computer-Human Interaction* 31.3, pp. 1–29. ISSN: 1073-0516.

Karahanoğlu, Armağan, Aykut Coskun, et al. (May 2024). "Is It Just a Score? Understanding Training Load Management Practices Beyond Sports Tracking". In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. CHI '24. New York, NY, USA: Association for Computing Machinery, pp. 1–18. ISBN: 979-8-4007-0330-0.

Karahanoğlu, Armağan, Rúben Gouveia, et al. (May 2021). "How Are Sports-Trackers Used by Runners? Running-Related Data, Personal Goals, and Self-Tracking in Running". In: *Sensors* 21.11, p. 3687. ISSN: 1424-8220.

Li, Ian, Anind Dey, and Jodi Forlizzi (Apr. 2010). "A Stage-Based Model of Personal Informatics Systems". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. New York, NY, USA: Association for Computing Machinery, pp. 557–566. ISBN: 978-1-60558-929-9.

Liang, Zilu, Bernd Ploderer, and Mario Alberto Chapa-Martell (May 2017). "Is Fitbit Fit for Sleep-Tracking? Sources of Measurement Errors and Proposed Countermeasures". In: *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*. PervasiveHealth '17. New York, NY, USA: Association for Computing Machinery, pp. 476–479. ISBN: 978-1-4503-6363-1.

Lin, James J. et al. (2006). "Fish'N'Steps: Encouraging Physical Activity with an Interactive Computer Game". In: *Proceedings of the 8th International Conference on Ubiquitous Computing*. UbiComp'06. Berlin, Heidelberg: Springer-Verlag, pp. 261–278. ISBN: 978-3-540-39634-5.

Liu, Fannie, Laura Dabbish, and Geoff Kaufman (Sept. 2017). "Supporting Social Interactions with an Expressive Heart Rate Sharing Application". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.3, pp. 1–26.

Loerakker, Meagan B. et al. (2025). "The Framework of the Lived Experience of Metrics: Understanding the Purposes and Activities of Self-Tracking Metrics". In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. CHI '25. Association for Computing Machinery. ISBN: 9798400713941.

Lupton, Deborah (Jan. 2017). "Self-Tracking, Health and Medicine". In: *Health Sociology Review* 26.1, pp. 1–5. ISSN: 1446-1242, 1839-3551.

Merriam-Webster (Dec. 2024a). *Definition of Health*.

– (Nov. 2024b). *Definition of Well-being*.

Niess, Jasmin and Paweł W. Woźniak (Apr. 2018). "Supporting Meaningful Personal Fitness: The Tracker Goal Evolution Model". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: Association for Computing Machinery, pp. 1–12. ISBN: 978-1-4503-5620-6.

Oxford English Dictionary (2025a). *Health—Quick Search Results*. Online. Retrieved January 10, 2025, from `https://www.oed.com/search/dictionary/?scope=Entries&q=health`.

– (2025b). *Well-being—Quick Search Results*. Online. Retrieved January 10, 2025, from `https://www.oed.com/search/dictionary/?scope=Entries&q=well-being`.

Page, Colleen and Mansour Rahimi (Oct. 1995). "Concurrent and Retrospective Verbal Protocols in Usability Testing: Is There Value Added in Collecting Both?" In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 39.4, pp. 223–227. ISSN: 1071-1813.

Paré, Guy, Chad Leaver, and Claire Bourget (May 2018). "Diffusion of the Digital Health Self-Tracking Movement in Canada: Results of a National Survey". In: *Journal of Medical Internet Research* 20.5, e9388.

Rapp, Amon et al. (Apr. 2018). "Designing a Personal Informatics System for Users without Experience in Self-Tracking: A Case Study". In: *Behaviour & Information Technology* 37.4, pp. 335–366. ISSN: 0144-929X.

Rooksby, John et al. (Apr. 2014). "Personal Tracking as Lived Informatics". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Toronto Ontario Canada: ACM, pp. 1163–1172. ISBN: 978-1-4503-2473-1.

Ruvimova, Anastasia, Joanna McGrenere, and Thomas Fritz (2026). "TimeMarbles: A More Holistic Approach to Self-Reflecting on Focus in the Knowledge Workplace". In: *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. CHI '26. Association for Computing Machinery. ISBN: 9798400722783.

Ryan, Kathleen, Conor Linehan, and Samantha Dockray (2021). "Appropriation of Digital Tracking Tools in an Online Weight Loss Community: Individual and Shared Experiences". In: *Designing Interactive Systems Conference 2021*. New York,

NY, USA: Association for Computing Machinery, pp. 999–1014. ISBN: 9781450384766.

Schramme, Thomas (Nov. 2023). "Health as Complete Well-Being: The WHO Definition and Beyond". In: *Public Health Ethics* 16.3, pp. 210–218. ISSN: 1754-9981.

Searle, Benjamin Lucas, Kyrill Potapov, and Amid Ayobi (2025). "Using self-tracking technologies to manage energy for health and wellbeing". In: *International Journal of Human-Computer Studies* 203, p. 103581. ISSN: 1071-5819.

Sefidgar, Yasaman S. et al. (2024). "Migraine-Tracker: Examining Patient Experiences with Goal-Directed Self-Tracking for a Chronic Health Condition". In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. CHI '24. Honolulu, HI, USA: Association for Computing Machinery. ISBN: 9798400703300.

Seligman, Martin E. P. (2011). *Flourish: A Visionary New Understanding of Happiness and Well-being*. Simon and Schuster. ISBN: 978-1-4391-9076-0.

Smits, Merlijn et al. (Apr. 2022). "From Digital Health to Digital Well-being: Systematic Scoping Review". In: *J Med Internet Res* 24.4, e33787. ISSN: 1438-8871.

Stiglbauer, Barbara, Silvana Weber, and Bernad Batinic (May 2019). "Does Your Health Really Benefit from Using a Self-Tracking Device? Evidence from a Longitudinal Randomized Control Trial". In: *Computers in Human Behavior* 94, pp. 131–139. ISSN: 0747-5632.

Ubel, Peter A. et al. (2005). "Misimagining the unimaginable: The disability paradox and health care decision making." In: *Health Psychology* 24.4, Suppl. Place: US Publisher: American Psychological Association, S57–S62. ISSN: 1930-7810(Electronic),0278-6133(Print).

WHO (2025). *Health and Well-Being*. Accessed: 2025-01-07.

A. STUDY PROTOCOL

A.1. Pre-Study

- Setup devices for video and audio recording.
- Open survey.
- Open Garmin Connect: “sso.garmin.com”.
- Start iMotions eye-tracking software.

A.2. Introduction

First of all, thank you for participating in this study. This study is aimed at understanding more about how users of fitness trackers monitor their health and well-being using their tracker. We will explore whether there is a difference between health and well-being, and if so, what that difference is. You can already start thinking about what these terms mean to you and whether and how they might differ.

We will use eye-tracking in this study. I will first familiarise you with the eye-tracking device. Then you will work in silence on the computer for a moment. Afterwards we will watch a video and I will ask you some questions about it. This interview consists of open questions - we are merely interested in your thoughts, opinions and preferences. There are no right or wrongs, so please feel free to take your time and share your thoughts. We will repeat this procedure once more. After I ask you a few final questions, we will conclude the session.

The study will take approximately 1 hour. Please know that participation is voluntary and that you are allowed to stop the experiment at any time. Also, if you have any questions or if something is unclear, please let me know.

During the study we will use eye-tracking to record your gaze. I would also like to take video and audio recordings of the interviews for transcription purposes. All data will be anonymised and the recordings will be deleted after transcribing. Please read this consent form carefully and sign it if I have your consent for this. Do you have any questions before we start?

Let participant read and sign consent form. Open pre-study survey.

A.3. Pre-Study Survey

You will start by completing a brief survey. This survey is simply to collect some basic demographic information and information on your fitness tracker use. Please let me know when you are done.

Survey questions:

- What kind of fitness tracker do you use? (Brand and model, e.g. Garmin Vivoactive 4)
- How long have you been using your fitness tracker?
- How often do you wear your fitness tracker?
- What do you use your fitness tracker for?
- What goals do you have when using your fitness tracker?

A.4. Condition 1: Health/Well-being

- Open Garmin Connect.
- Please log in to your Garmin Connect account.
- Open iMotions.
- Make sure participant sits at right height and distance to screen.

I would like to start by showing you the eye-tracking device we will be using. During the following tasks, the eye-tracking device will record your gaze on the screen. First, we will go through a quick calibration process. During this process, you will be asked to follow a series of points on the screen with your eyes. This step is important because it ensures the system tracks your eye movements precisely. We will continue the calibration until the program indicates that we have an ”excellent” rating. Once we have achieved that, click on “continue” and the eye-tracking will start automatically. Following this, please open the Garmin Connect website in the browser.

Your task is to gain an overview of your health status. You will have three minutes to freely explore your tracker data. I will let you know once the time is over. If you have any questions, please let me know. OR Your task is to gain an overview of your well-being status. You will have three minutes to freely explore your tracker data. I will let you know once the time is over. If you have any questions, please let me know.

- Start 3-minute timer once participant is on Garmin Connect website
- After 2 minutes, tell participant they have 1 minute left
- After 3 minutes, stop eye-tracking - open gaze recording
- Start video and audio recording

Next, we will proceed to the interview part of the study. We will now watch a screen recording of your eye-tracking session together. In this recording, you will see your eye movements on the screen, with

circles representing the points you looked at and lines the transitions between those points.

As you view the recording, please take a moment to reflect on your thought process while you were collecting information on your health/ well-being status. I would like you to describe what you were thinking during the exploration. Specifically, consider why you focused on certain metrics, what information caught your attention, and how you interpreted the data displayed. There are no right or wrong answers in this part of the study; your honest reflections are what we are most interested in. Please feel encouraged to take as much time as you need to share your thoughts. There is no need to rush through the recording. You can pause or rewind the playback at any time if you wish to review a specific moment or reflect further on your thoughts.

Guiding questions during the RTA:

- What did the data you collected with your Garmin tell you about your health/ well-being?
- Which metrics were most relevant to monitor your health/ well-being?
- What made you decide to focus on these particular metrics?
- Could you describe what metric X means to you in the context of your overall health/ well-being?
- How does an increase/ decrease in values of metric X influence your health/ well-being?
- If you could add any metric to improve your overall understanding of your health/ well-being, what would it be and why?

Stop video and audio recording.

A.5. Condition 2 Well-being/Health

Make sure participant sits at right height and distance to screen.

Now we will begin the second part of the study. As before, we will use eye-tracking to record your gaze on the screen. You will first go through the calibration process again until the program indicates that we have an ”excellent” rating. Once we have achieved that, please click on “continue” and the eye-tracking will start automatically. Following this, please open the Garmin Connect website in the browser again.

Your task is to gain an overview of your well-being status. You will have three minutes to freely explore your tracker data. I will let you know once the time is over. If you have any questions, please let me know.

OR Your task is to gain an overview of your health status. You will have three minutes to freely explore your tracker data. I will let you know once the time is over. If you have any questions, please let me know.

- Start 3-minute timer once participant is on Garmin Connect website
- After 2 minutes, tell participant they have 1 minute left
- After 3 minutes, stop eye-tracking - open gaze recording
- Start video and audio recording

Now that we are moving on to the second interview portion of the study, I will again show you the screen recording of your eye-tracking session. While you watch this recording, please reflect on your thought process as you explored your data regarding your well-being/ health. Again, I would like you to describe what you were thinking during the session. Please describe what was on your mind at different moments, why you chose to focus on certain metrics, what information captured your attention, and how you interpreted the data displayed.

Although you have already done this once before, remember that we are now looking at well-being/ health, while we looked at health/ well-being before. We are interested in any new insights you might have this time, but please do not be afraid to repeat yourself if necessary. Again, there are no right or wrong answers here and you are welcome to take as much time as you need to share your thoughts. Please remember that you can pause or rewind the playback at any point if you want to revisit a specific detail or reflect further.

Guiding questions during the RTA:

- What did the data you collected with your Garmin tell you about your wellbeing/ health?
- Which metrics were most relevant to monitor your well-being/ health?
- What made you decide to focus on these particular metrics?
- Could you describe what metric X means to you in the context of your overall well-being/ health?
- How does an increase/ decrease in values of metric X influence your wellbeing/ health?
- If you could add any metric to improve your overall understanding of your well-being/ health, what would it be and why?

A.6. Final Questions

Now that we have completed both sessions - one focusing on health/ well-being and the other on well-being/ health - I would like to ask you a few last questions about these concepts.

Questions comparing health and well-being:

- How do you personally define the concepts “health” and “well-being?” What do these terms mean to you? How do they differ?
- Apart from the metrics measured by your Garmin, what are other factors relevant to your health and why?
- Apart from the metrics measured by your Garmin, what are other factors relevant to your well-being and why?
- Would you like to track these metrics using your fitness tracker as well?
- We are now at the end of this study. Do you have any comments or thoughts you would like to add?

I would like to thank you for taking the time to participate in this study. Stop video and audio recording. Let participant log out of Garmin Connect.

B. GARMIN CONNECT INTERFACE

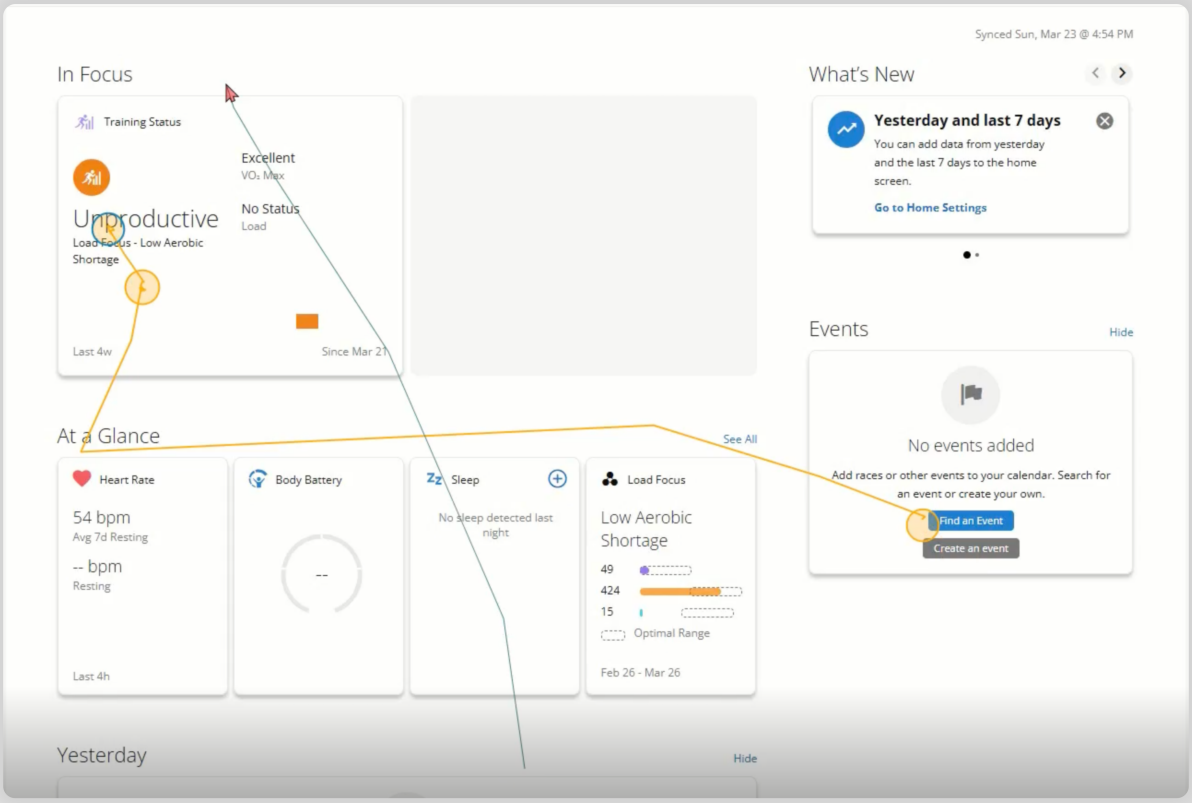

Figure 1: *A frame from a gaze replay video, showing a participant viewing the Garmin dashboard. Yellow lines and circles represent eye movements and fixations. The movement of the mouse cursor is also visible.*

C. INTERACTIVE VISUALISATION TOOL

The interactive visualisation tool can be found here: `https://healthandwellbeing.github.io`.

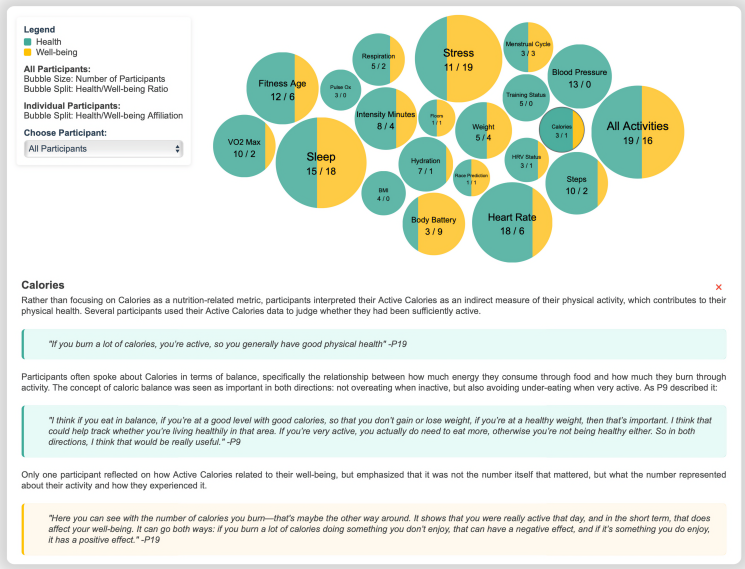

Figure 2: *The interactive visualisation supports exploration of participants' metric ecologies. In the All Participants view, bubbles represent metrics used to interpret health and/or well-being, with bubble size indicating the number of participants who included each metric. In this example, Calories was associated with health by three participants and with well-being by one participant. Selecting a metric reveals illustrative participant quotes and summary information. The visualisation also supports exploration of individual participants' unique metric ecologies, as shown in Figure 3.*

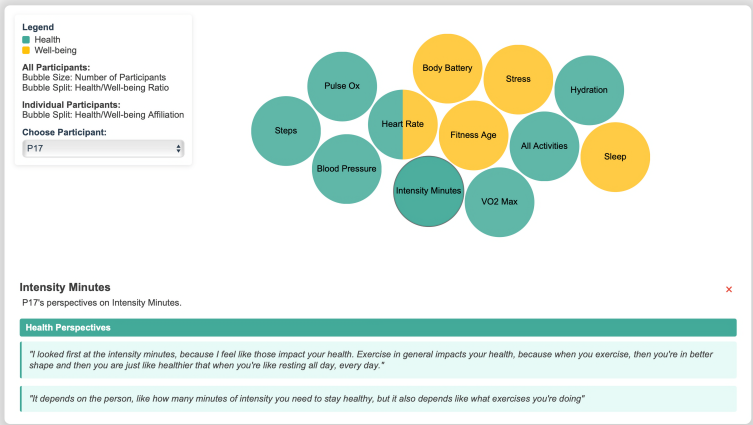

Figure 3: *The interactive visualisation also enables exploration of each participant's unique metric ecology—in this case,* P17*'s. It displays the metrics that this participant associated with their health and/or well-being. Selecting a specific metric (e.g., Intensity Minutes) shows* P17*'s corresponding quotes about that metric.*