



# Flexible Minimalist Self-Tracking to Support Individual Reflection

MATTHEW BARKER-CANLER, DANIEL GOOCH, JANET VAN DER LINDEN, and MARIAN PETRE, The Open University, Milton Keynes, UK

Self-tracking technology can help users develop new knowledge about themselves, supporting their health and general wellbeing. Most of these devices inform users about their lives by autonomously generating data about highly constrained topics. Recent research has started to explore more flexible, manual self-tracking systems, which allow users to express themselves through their data and learn by engaging with their experiences in-the-moment. This article describes a field trial using the Chromatize app, designed to explore the value of minimalism and flexibility – including in data histories – in personal tracking systems. Our data indicate that low-burden experience logging systems which focus on the creation of symbolically-rich, self-defined data can support self-awareness, reflection and regulation at opportune moments. The analysis also contributes a formulation of reflection in the context of self-tracking technologies that is more meaningful than the loose use of ‘reflection’ used in much of the literature.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**;

Additional Key Words and Phrases: Self-tracking, wellbeing, reflection, minimalism, personal informatics, quantified self

## ACM Reference format:

Matthew Barker-Canler, Daniel Gooch, Janet van der Linden, and Marian Petre. 2024. Flexible Minimalist Self-Tracking to Support Individual Reflection. *ACM Trans. Comput.-Hum. Interact.* 31, 3, Article 30 (August 2024), 35 pages.

<https://doi.org/10.1145/3660339>

## 1 Introduction

In recent years, technologies for **personal informatics (PI)** have risen in popularity. Today there are a range of self-tracking applications available which can help people to learn about a range of different aspects of their life, many of which have a bearing on their health and wellbeing.

Much of the research into self-tracking has focused on quantified measures of physical health. Rapp and Tirassa have critiqued what they term the “disappearance of the self” in PI research [51]. Review studies have noted that the majority of self-tracking articles rely on the self-improvement hypothesis, with little evidence for the efficacy of designing technology in this manner [19].

With the prevalent focus on quantifying aspects of peoples’ lives, technology is not supporting users to understand their unique perspectives and individual experiences – which is of great

Authors’ Contact Information: Matthew Barker-Canler, The Open University, Milton Keynes, UK; e-mail: matthew.barkercanler@gmail.com; Daniel Gooch (Corresponding author), The Open University, Milton Keynes, UK; e-mail: daniel.gooch@open.ac.uk; Janet van der Linden, The Open University, Milton Keynes, UK; e-mail: janet.vanderlinden@open.ac.uk; Marian Petre, The Open University, Milton Keynes, UK; e-mail: m.petre@open.ac.uk.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike International 4.0 License.

© 2024 Copyright held by the owner/author(s).

ACM 1557-7325/2024/8-ART30

<https://doi.org/10.1145/3660339>

value in supporting wellbeing. Other researchers are starting to explore more flexible modes of self-tracking, where automated data entry is combined or replaced with manual data entry, giving users more control over what form their data take and what it represent for them as individuals (e.g., [3, 11, 39]). Novel manual data creation practices have been explored which move beyond numerical data collection and introduce alternative media (e.g., the creation of clay models [41], personalised tangible data visualisations [59], sketches [38], chocolate treats linked to self-monitored data [36] and paper bullet journals [3]).

This emerging strand of PI research emphasises the inherent value of flexible manual data entry as a meaning-making practice. Systems which are designed to support expression in this way appear to promote richer self-awareness, reflection and self-insights into personal experiences, all of which can be associated with wellbeing. However, such manual tracking systems can be burdensome to use during everyday life as they require users to log their experiences manually, often through effortful, expressive actions.

In this article, we explore the design space for systems that minimise the effort users have to expend to log their experiences in a flexible manner. We discuss the design and evaluation of *Chromatize*, a prototype for a low-burden, experience-logging smartphone application. It is specifically intended to encourage users to have moments where they become aware of their experiences in an open and unrestricted way. By deploying the app with users, we can examine the design features expressed through *Chromatize* and probe the PI design space. We thus stimulate discussion about how we can design tools that are tailored for helping users to enhance their wellbeing through meaningful sense-making. Our work highlights the value in the deliberate, expressive logging of, and reflection on, experiences in-the-moment.

In the PI literature, ‘reflection’ is often characterised as a key process through which users gain value from logging their data. Reflection at the time of logging has been described as either emerging from engagement with real-time feedback from a device during an activity or from the act of manually logging data on the device. Explanations for how these processes unfold are high level and are challenging to apply to data collected through PI studies. Many studies fail to assess whether the designed systems result in data-driven change [19], potentially due to the complexity of relating participant experiences to our understanding of reflection. Based on our findings, we also contribute clarity over two forms of reflection that are meaningful in analysing PI data, hard and soft reflection, which distinguish between acts of reflection which require a change in perspective, and those that do not.

These joint research areas are best expressed through our key research questions:

- RQ1: How can minimalist, flexible self-tracking technology support experiential logging during everyday life?
- RQ2: How do users reflect when they use experiential logging devices and what does this tell us about how reflection should be conceptualised within PI?

## 2 Literature Review

Our literature review is structured to first examine specific design considerations regarding PI technologies before considering how self-tracking research has interpreted the concept of ‘reflection.’ We first consider the importance of minimising the effort embedded within logging technology and the associated visualisations. This leads into a broader consideration of flexible PI systems, covering both expressive and ambiguous media, before examining the importance of topic selection. The literature review concludes with an examination of different formulations of reflection, and how this has been interpreted within the field of PI.

## 2.1 Minimalist Logging with PI Technologies

Researchers have long noted the tendency of users to abandon self-tracking technologies when they grow frustrated with the workload of logging, managing and reviewing data from day-to-day [28, 40]. Expressive self-tracking platforms exacerbate this issue, demanding even more investment into the logging and interpretation process [3, 59].

Recognising this, there has been some explorations of minimalist PI system designs, emphasising simplicity and ease-of-use. For example, SnAPP is a system that uses a physical button attached to the audio jack of the user's phone to solicit input from users [24]. A single press records a positive moment and two clicks record less positive moment. The 4-week field trial with seven participants required users to log their moods. The results indicate that participants appreciated the low-effort nature of the logging tool, while still being able to construct rich narratives based on the minimal data collected.

SleepTight is an application designed to support sleep quality [11]. Choe et al. examined the value of minimising the effort of data collection by developing two versions: one which allows data input directly from the Android home screen and one which requires the user to enter the SleepTight app. Through a 4-week deployment with 22 people, they found that using the minimal effort version led to a significant increase in data entry compliance (93% compared to 72%).

Although these minimalist interface designs support quick data entry due to the simplicity of the input mechanism — which seems to encourage users to log more often — they also only facilitate low-resolution data logging, which compromises their flexibility. Users of SnAPP can only record two different types of input: a single button press for inputting a good mood and a double press for inputting a bad mood. Users of SleepTight, meanwhile, have just five inputs directly available to them.

Chong et al. expand on this approach, developing a system in which users log a timestamp by squeezing a physical ball when an experience happens, and associating richer data with that timestamp afterwards [13]. The week-long trial with eight participants suggests that in addition to the time management benefits, such a design can scaffold elements of reflection by requiring the user to consider the context of the recorded timestamp when trying to add richer data. However, the principal cost is that users can often forget what the timestamp represented.

Beyond the act of logging, some have focused on the benefits of providing minimal visualisations of the recorded data. Elsdon et al. argue that when there is a sentimental attachment to the event, such as a wedding day, there should be less focus on scale because the meaning can get lost within it, calling for smaller scale displays that “succinctly distil some essence of the experience” [20].

Rapp and Tirassa have argued that inexperienced users tend to find large-scale data visualisations, in which multiple streams of data are integrated together, particularly complex and abstract [51]. They suggest that data should be presented in such a way that users can recognise something of themselves in it and that users should be able to revisit individual memories symbolised within it. This implies an approach where there is less focus on enabling users to engage with data as a mass — through trends, quantification and statistics, and more focus on engaging with individual data pieces and their individual meanings.

## 2.2 Flexible PI Systems

A second factor in supporting meaningful self-tracking is the flexibility of the PI technology, which needs to prioritise the lived experience of patients and better integrate the design of technologies into everyday life [9, 16, 44, 58]. Within the subset of PI technologies specifically developed to support self-reflection, this has been achieved through two principal means: utilising expressive media and supporting open topic selection.

*2.2.1 Expressive Media.* We use the phrase ‘expressive media’ for forms of data that support personalisation and customisation by a given user. This covers a wide range of media, both digital and analogue.

A long-standing analogue approach has been the use of bullet journals; a paper-based journaling format for the rapid logging of bullets, short sentences, numbers and other visual data forms in a pre-determined structure. Ayobi et al. examined views on bullet journals from data posted on social media websites, noting the vast range of different topics, styles and creative skills deploying in this form of tracking [3]. They note the capacity of users to express and explore their felt experiences through bullet journals as a way “to keep track of and cope with felt experiences in everyday life.”

These ideas are mirrored in a study by Thudt et al., in which they implemented a toolkit for logging physical personal data [59]. As with bullet journaling, the toolkit allowed users to both choose a topic and choose how they will represent their experiences as data – except that, rather than drawing out the self-tracking structure on paper, users created it with physical crafts objects, such as beads, strings, pins and plasticine. Nine participants used the toolkit for 2 weeks, with the results indicating that the effortful personalisation of the process supported diverse forms of reflection.

An alternative approach has been to support a system whereby users freely self-define the meaning of each individual piece of data in-the-moment when they create the data entry. This approach is exhibited in the Mindtracker platform, an emotion tracking toolkit with which users create clay models representing their emotions and record them in an associated smartphone app by taking a photograph [41]. Based on a 2-week data collection study with 16 college students, Lee and Hong noted that the meanings that users associated with their models were diverse; ranging from the sensation of ‘deliciousness,’ to the feeling that ‘time is passing’ and that the physical form allowed a significant level of personalisation.

Using such toolkits in the long-term are unlikely to be popular given the necessary level of time commitment. However, diverse digital technologies have been developed to support a similar level of expressiveness while logging.

The GoSlow app was designed to “help users slow down, contemplate and be alone” [10]. Logging functions include free-form text, colour logging and photo recording. The findings from a 1-week trial with seven participants suggested that the app may support some level of reflection, with little detail on what form that reflection would take.

The Trackly app supports people in defining and colouring pictorial trackers [2]. Users can choose from six different types of trackers which range from “a time ring tracker that represents 24 hours” to “two mandala trackers, namely a flower and rainbow.” Users select their specific parameters (e.g., ‘tired – nap needed’ when tracking fatigue) and associate colours with the parameters. Participants from the 4-week field study reported a sense of ownership through defining their own form of tracking, noting that the expressiveness of the tool supported them in making sense of their data.

A similarly flexible approach is designed into the LifeMosaic app [45]. The app is a result of co-designing with four teenagers, who expressed a need for self-tracking technology to support self-expression. A user selects two opposing qualities to track (e.g., ‘I slept well’ vs. ‘I slept badly’). These are represented by two coloured buttons with the app, which the user selects to represent to what extent that event has occurred. The data are visualised through a set of daily tiles, with the colour for a given dataset by the combination of button pushes. Such a design supports a level of ambiguity within the visualisation, which is expressive rather than precise, providing the gradient between the two extremes rather than how many times the user pressed each button.

The use of colour as an ambiguous and personal media is a notable feature of both of these apps. As well as being easy to represent digitally, there is a strong relationship between colour

and emotion [29]. While there are broad cultural associations between colour and emotion [61], colour-related interpretations are highly personal and individualistic [35].

These concepts are the foundation of the Lumino device, a wall-mounted circular lamp used to select and visualise moods in a standard colour palette used across graphics programmes. An inner dial is used for selecting a colour to log, viewing the data history and discarding a recorded colour. The design intention was to focus on providing free choice to participants to associate colour selection with meaning. The device was deployed with six participants for a 3-week field trial, during which use of the device both became routine and showed potential in assisting the user with positive reflection. Kim et al. provide little detail on the form of reflection the device engenders, but the physical properties of the device limit the potential of supporting reflection in action.

Colour has also been used as the media in expressive visualisations of automatically collected data, designed such that users can emotionally connect with the data being presented [48].

There have been early trials of using clothing which changes colour as a result of biosensing (specifically, skin conductance) [31, 32]. While the focus has been on the social dynamics of wearing such displays, in both trials participants appear to have valued the expressive ambiguity of the displays and the emotions they publicly convey.

Sanches et al. present a month long trial with 23 participants of their ‘Affective Health’ system, integrating wearable biosensors with a mobile app visualising the users arousal [53]. The open-ended nature of the design led to participants forming individual interpretations of the data which ranged widely across stress management, emotion tracking, general life logging and behaviour change goals. The results demonstrate how the ambiguous form of the biosensing can provide insight into “an ongoing lived emotional experience that can only be understood in a specific socio-historical context.”

**2.2.2 Topic Selection.** Many studies of PI argue that providing users the ability to set their own topics for tracking provides the flexibility necessary to enable users to personalise their self-monitoring practices and thus use the tool to support their own purposes [57, 58].

Such user-led practices of meaning-making require PI tools to be designed to support this level of personalisation [3]. Examples from our ambiguous data tools include LifeMosaic [45], where during the design process, users expressed a clear desire for the tracking topic to be selected by the user. Trackly also supports open-ended tracking, with users defining personally meaningful tracking parameters in their own words [2]. Ayobi et al. argue that this encouraged creative, reflective thinking.

**2.2.3 Summary.** While this review of the literature highlights the importance of minimalist logging, expressive media and flexibility in topic selection, there has been little research into combining these design properties into a single device. This motivates our first research question, which focuses on understanding whether expressive media – if recorded in a minimalist manner – can support experiential logging:

- RQ1: How can minimalist, flexible self-tracking technology support experiential logging during everyday life?

Furthermore, few of the studies we have discussed are explicit in either what form of ‘reflection’ they are designed to support. For those technologies deployed in field trials, little consideration is given to what form of reflection participants state the technology helped to support.

### 2.3 Formulations of Reflection

In his book *How We Think*, the educational philosopher John Dewey established what is perhaps the most commonly cited definition of reflection, one which has become the basis for many subsequent definitions of reflection, an “active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions to which it tends” [17]. Dewey’s interpretation has been the basis for a lot of the subsequent conceptual work on reflection and has spurred debate about the nature of reflection.

Donald Schön approaches the problem from a different perspective. In his book on reflection in the workplace, *The Reflective Practitioner* [54], he has proposed an alternative interpretation of the relationship between reflection and action. He recognises the form of reflection described by Dewey, where reflection is separate from action. However, he introduces another process in which the two are tied together as part of a single process — reflection-in-action. For Schön, reflection-in-action is when, rather than reflecting on their actions after-the-fact, individuals reflect directly through their actions while they are being enacted. As such, reflection and action are intimately bound up together, feeding off one another. Schön’s approach puts greater emphasis on experience and intuition as a source of knowledge but is focused less on emotional and experiential reflection.

### 2.4 HCI-Based Reflection Frameworks

Reflection has emerged as an increasingly important concept in **Human–Computer Interaction (HCI)** research. However, despite being frequently pointed to as a key dynamic in the interaction between the technology and its user, there has been a general lack of clarity in how it has been applied. Baumer et al. [5] and Slovák et al. [56] have both recognised that despite the breadth of its usage, reflection as a concept is rarely well-defined and has not been consistently or effectively applied in HCI. As Baumer concludes, “... little work actually explicitly defines what reflection is, and even less grounds the definition in a conceptual or theoretical framework” [5].

Slovák et al. generate a conceptual framework to underpin the design of technology to support transformative reflection in the context of social-emotional learning [56]. The three key components are *explicit*, *social* and *personal*. While the framework provides a clear link from design decisions to reflective outcomes, unfortunately, the design concepts don’t integrate well into flexible self-tracking practices which aren’t explicitly focused on transformative reflection. Furthermore, creating explicit activities for reflection, and supporting social dynamics, do not match well with the goal of self-directed reflection.

Rather than focusing on frameworks to support reflective technology in specific domains, two key pieces of literature provide conceptual frameworks of how to understand reflection within the context of HCI. The first is Baumer’s paper on reflective informatics, in which they define three conceptual dimensions of reflection [5]. The second is in Fleck and Fitzpatrick’s paper where they set out a framework, describing reflection in HCI in terms of five different levels of reflection [26].

Baumer marks out three broad dimensions of reflection. The first of these — *breakdown* — is what causes users to reflect in the first place, described as when users detect “disturbing anomalies... violating expectations.” The second concept, *inquiry*, is when users examine inconsistencies in their knowledge and seek an explanation, through “iterative hypothesis testing and refinement” and the “re-examination of knowledge and its origins.” The third conceptual dimension is framed as *transformation*. This means “change to the fundamental, basic conceptualisation of a situation.” Baumer suggests that this may be the most challenging design problem of all.

In contrast, rather than employing broad concepts to frame how reflection emerges through interaction, Fleck and Fitzpatrick provide a level-based framing of different types of reflection, operationalising the characteristics of each level of reflection with detailed descriptions [26].



Table 1. Fleck and Fitzpatrick’s Model of Reflection

Level	Description	Explanation
R0	Description	Non-reflective description of events
R1	Reflective description	Reportive description of events with justifications No shift in perspectives Not always considered reflection
R2	Dialogic reflection	Probes relationships in the descriptions Seeing things from multiple perspectives Reach a different level of understanding
R3	Transformative reflection	Asking fundamental questions Challenging personal assumptions Changes in in practice or understanding
R4	Critical reflection	Consider aspects beyond the immediate context, such as moral and ethical issues Very rare

The tiering of Fleck and Fitzpatrick’s model is structured by the depth and scope of the evaluation of the information (see Table 1). At the bottom level is *Description*, which is not a reflective process, and is based around a simple statement of events. The first reflective step is *Reflective description*, which is thinking about a single belief or view and doesn’t involve a shift in perspective. *Dialogic reflection* has a broader scope because the user is exploring relationships between information about different dimensions of their beliefs or perspectives and does involve a shift in perspective. *Transformative reflection* has more depth again, because in addition to exploring relationships, the user is considering their fundamental basis and experiencing a significant shift in perspective. *Critical reflection*, as the top level, additionally requires application of social or ethical issues to the generation of insights.

The two frameworks emphasise different aspects of how users might reflect through their use of technologies. Fleck and Fitzpatrick’s model puts more emphasis on how users reflect by cognitively processing experiences and information about those experiences. Baumer’s model puts more emphasis on the role of the device and how it can provocatively prompt users to reflect. Of the two, Baumer frames reflection in the more extreme way, picturing it as a more demanding, by-definition-transformational process. It seems to imply that the system should invoke reflection by jolting users into a reflective state through its interaction design. Fleck and Fitzpatrick’s model, by comparison, is more flexible; accommodating lighter moments of reflection, which aren’t intentionally provoked by the system, and may not involve a shift in perspective.

## 2.5 Reflection in Self-Tracking Systems

The main aim of most self-tracking technologies is to support behaviour change by helping people to learn about themselves and enhance their self-knowledge [42]. Reflection, as the catalyst for learning, is recognised as being a key contributor to this. As such, there has been substantial discussion in self-tracking research about the dynamics of reflection in technology use and how it can be fostered through interaction design.

**2.5.1 Transformative Reflection.** In their model of self-tracking reflection, Li et al. position reflection as one of their five stages, which includes preparation – collection – integration – reflection – action [42]. Although the model does not provide a clear definition of reflection, the action stage is

described as when “people choose what they are going to do with their newfound understanding of themselves.” This implies that reflection is a transformative process that involves the development of newfound understanding that users can act on.

Other researchers have also focused on reflection as a process that catalyses behavioural change. Cox et al. have argued that users can have digital epiphanies when interacting with data — moments when they have a striking realisation about something and make a conscious decision to change their behaviour [14]. Similarly, Thudt et al. have argued that most instances of reflection detected in their field trial of users physically constructing data representations were transformative because they could be linked to changes in action and beliefs [59]. The highly flexible, low-constraint Echo smartphone app, makes use of multiple expressive media. Over the course of three deployments, there are indications that the scaffolding provided is somewhat successful at supporting transformational reflection [34].

**2.5.2 Transformative Reflection-on-Action.** There is also a question of *when* reflection occurs. Li et al. suggest that there are two different occasions when users can reflect while using PI technologies: (1) when they revisit data that have been displayed back to them a period after it has been logged, which they call ‘long-term reflection’ and (2) when they look at data just after it has been logged, ‘short-term reflection’ [42]. This separation of reflection by time also underpins the way that Epstein et al. model reflection in their Lived Informatics model [23].

In terms of the frameworks, reflecting when revisiting data resonate with Fleck and Fitzpatrick’s description of dialogic reflection as “looking for relationships between different pieces of experience” [26].

This approach is common across many fitness and wellbeing tracking technologies, where significant amounts of data are collected for later review. The Health Mashups system supports users in tracking a range of data (including steps, locations, food, sleep and mood) and then provides text summaries (e.g., “last week you were happier than usual”) and displays any correlations between different streams of data [6]. Similarly, Epstein et al. have examined how to best present the complexity of diverse data streams, proposing ‘cuts’ showing trends and patterns [21]. Cuttone et al. discuss how adding additional information to data dashboards allows different questions to be posed, positioning reflection as a process of data analysis [15].

Given this complexity, some researchers have started considering whether people are best placed to undertake the reflective process. For example, there has been some focus on developing algorithms that assist in processing data to inform users about their ‘future selves’ [30].

The shortcoming of much of this work is that our lives are rarely entirely encapsulated by easily-collected data streams. This has led to some work exploring the value of ambiguous data in PI technology, and how it supports reflection.

**2.5.3 Ambiguity Supporting Reflection.** While the conventional approach is making data visualisations easily interpretable, there is an argument that including a degree of ambiguity in the presentation of the data can encourage reflective thinking. Kim et al. trialed an app that implemented ambiguous data representations [38] and have recommended that designers focus on “unconventional encodings to create an element of surprise and stimulate reflection.” The sense of perplexity, or as Baumer terms it, ‘moment of breakdown,’ when faced with ambiguous data might incite users to reflect.

Whooley et al. argue that abstract representations are ambiguous, encouraging close personal engagement with the data [62]. They go on to argue that while automated collection pulls data together to reveal insights, manual logging allows a close engagement with both the lived experience and the collected data.



Many of the manual tracking methods using ambiguous media detail how they lead to rich reflective practices. Ayobi et al. note that users reflect in the act of adding new data to their bullet journal as they review and add data simultaneously [3]. The authors imply that bullet journaling can give rise to moments of transformative and non-transformative reflection. Participants noted the value of self-expressiveness, while also finding value in reflecting on present experiences, retrospective reflective thinking and identifying thoughts/goals/intentions for future selves. Thudt et al., who provide a more detailed breakdown of the reflective process, have suggested that the physical manipulation of materials to create data contributes to the users' understanding of the experiences it represents [59]. Similarly, in their analysis of the practices of quantified-selves, Choe et al. described how users "feel intimacy with data" and "make sense of it," when they manually collect it [12]. While this implies some form of reflection, it is challenging to map to the frameworks.

*2.5.4 Non-Transformative Reflection.* Not all technology supports transformative reflection. The OmniTrack app is a smartphone application that provides a neutral self-tracking platform which users can adapt around their needs [39]. Users set their topic and then select from a range of data entry techniques (such as text, ratings, location or audio record) which are associated with the topic. Field trials indicate that participants can make use of the expressive form of the app to best suit their tracking needs – be that in supporting self-awareness or self-reflection.

Choe et al. have attempted to frame reflection in a more nuanced way than most [12]. Their study focused on an exploration of personal data visualisations in a lab setting. Using a think-aloud data collection method, users discussed their thoughts as they explored the visualisations. Analysing these data and associating them with the different levels of Fleck and Fitzpatrick's framework, Choe et al. found multiple instances when users exhibited level 1 (reflective description) and level 2 (dialogic reflection), but few examples of level 3 (transformative reflection) and no examples of level 4 (critical reflection).

## 2.6 Summary

Our review of literature demonstrates the lack of clarity over what different researchers mean by 'reflection,' given how imprecisely it is discussed across different articles. One of the challenges of analysing the level of reflection reported in previous studies is the under-reporting of definitions of reflection, and the complexity of trying to apply the different frameworks to published work, where the focus of attention may not be on the nature of the reflective process, but the outcome that results. This is not to critique specific authors but to highlight that existing frameworks are hard to apply to lived experience and real-world data, leaving many relying on a broad sense of 'reflection.' This leads to our second research question:

- RQ2: How do users reflect when they use experiential logging devices and what does this tell us about how reflection should be conceptualised within PI?

## 3 App Design

To gain a deeper understanding of the complex concept of reflection within the self-tracking design space, we developed an app for a mobile phone. The developed app was named *Chromatize*. The codebase for the app can be found at [https://github.com/MetroBeam/Colourize\\_Code/tree/master](https://github.com/MetroBeam/Colourize_Code/tree/master).

In designing this app, we focused on two key design principles. The first was to minimise the burden of manual tracking. As noted in our literature review, there has been some success in the context of constrained logging, that minimal logging increases adherence [11, 24]. Supporting minimalist logging allows the technology to integrate into daily life, with the unobtrusiveness of

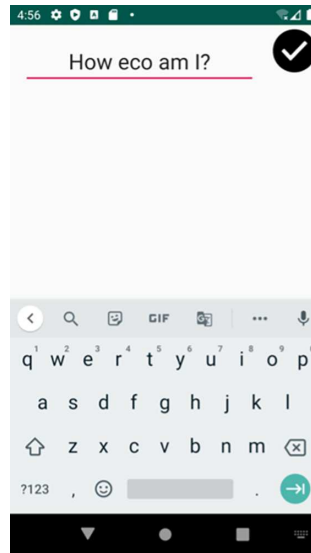


Fig. 1. Chromatize topic setting interface.

the technology supporting logging within the flow of daily life [4]. We took this principle forward and tried to minimise the interaction necessary for users to log an entry.

The second principle was to utilise an ambiguous form of media, to allow users to engage more freely in associating the logged data with their lived experience [38, 62]. In the literature review, we presented a range of devices utilising ambiguous media, before concluding that colour would be the best medium for supporting users to openly express and engage with aspects of their life that matter to them and that they want to track [29, 35, 61]. In many ways we are building on the Lumino device [37] and GoSlow app [10] in foregrounding colour as a logging technique.

This combination of minimalist mobile logging using ambiguous media provides a set of design features ideal for exploring the design space of PI technology and provides a platform for discussing how such technologies support self-reflection. We drew extensively from the literature in developing the key design features embedded within the app.

### 3.1 Topic Selection

We noted in the literature review how providing users the ability to set their own topics for tracking supports users to use the tool to support their own purposes [57, 58]. Previous applications have implemented this through users defining personally meaningful tracking parameters in their own words [2, 3].

We have followed this approach. During the initial setup of the app, users state which topic they want to record data about (see Figure 1), with the knowledge that they can change the topic if they change their mind or they want to refine the topic. Within their chosen topic they then log experiences through picking colours using one of the logging methods discussed below.

At any point, users can update their topic through accessing a settings menu. This returns them to this identical screen. An example of this can be seen in Figure 3(b); while we cannot see the original topic, on Thursday it was set to ‘Parenting’ and 5 weeks later changed to ‘Days out.’ The updated question also appears on the widget in the title bar (see Figure 2(b)).

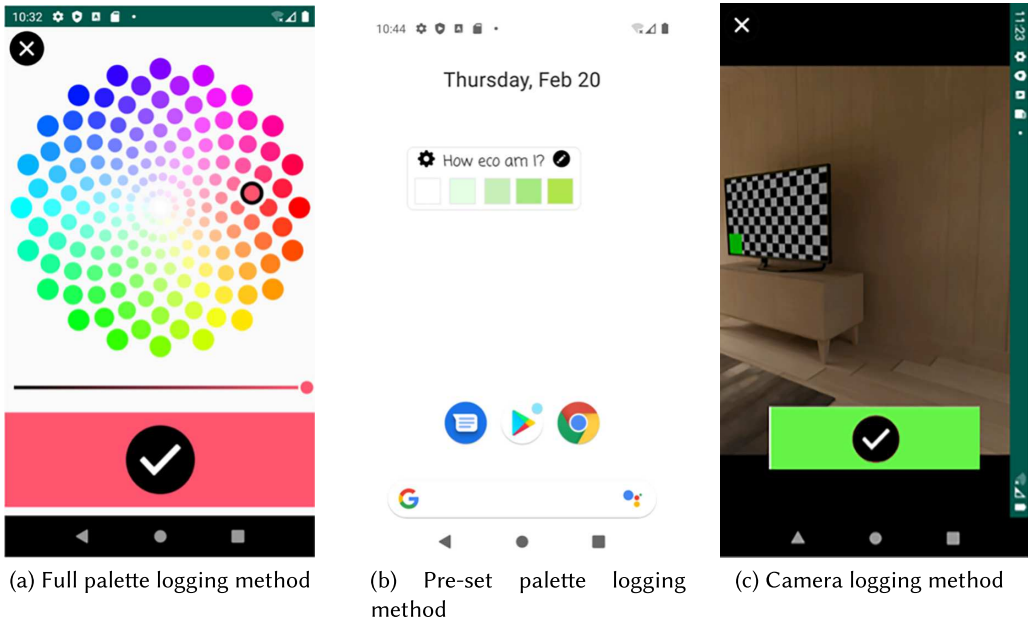


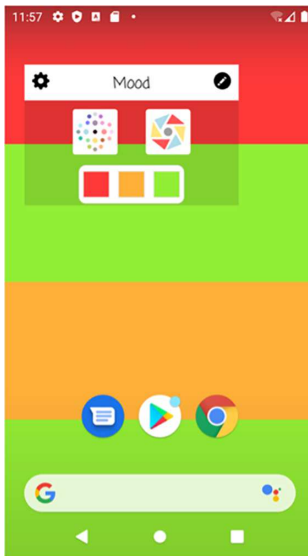
Fig. 2. The three logging methods of the Chromatize app.

Given our focus on exploring how a minimalist, flexible self-tracking system was used to engage with lived experiences, we set no constraints on what topics participants should log about. Participants were given complete freedom to set their own self-tracking topics.

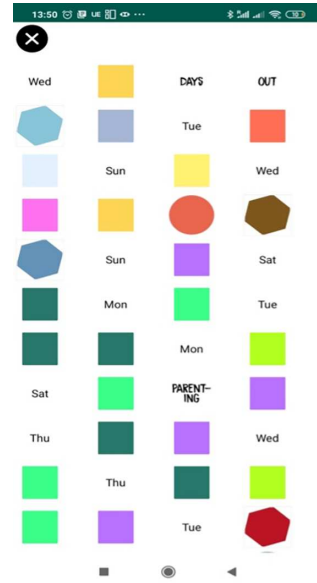
### 3.2 Logging Methods

We developed three different methods for colour data logging, as shown in Figure 2. The idea was that this would enable the investigation of different forms of minimalist data collection. The three methods were as follows:

- (1) Large palette logging method: Users can select any hue from a large palette of colours with three button presses. They tap on the palette icon on their smartphone home screen which takes them to the large palette selection interface, where they then choose a colour by tapping on it, before pressing the check mark to confirm. There is a slider for modifying the brightness of the colour (see Figure 2(a)). This method is based on the colour palette used across many graphics programmes and within the Lumino system [37].
- (2) Pre-set palette logging method: Users can pre-set a small number of colours (e.g., different shades of green) which they can subsequently log with a single button press. Users can change their own pre-set selection of colours available to them on the app at any time (see Figure 2(b)). This method is similar to the quick data capture mechanism within the OmniTrack tool [39].
- (3) Camera capture logging method: Users log a colour as if capturing the colour from their environment, with three button presses. This is done by tapping on the camera shutter icon on their smartphone home screen, which opens a camera view. However, instead of taking a photo, the user taps on the display to capture a colour. When they tap, the colour on the screen, where they tapped is added to the box at the bottom of the interface. Tapping on the check symbol then logs the colour. In the example image (see Figure 2(c)), the user has



(a) The four data-piece display, showing the last four colours logged. The app that allows a user to use any of the three logging methods of the Chromatize app is also visible.



(b) The all-history display, showing all of the colours logged through Chromatize, and which logging method was used to record the colour.

Fig. 3. The different history screens of the Chromatize app.

captured the colour green from their TV screen. This method was inspired by the presence of a camera on most smartphones, with the ambiguity coming from the selection of a colour from within the user's locale.

We also developed an alternative version of the app which brings together the three different logging methods described above. All of the logging methods can be accessed directly from the smartphone home screen. The app has three button icons, one for accessing the large palette, another one for capturing colours and one for directly logging one of the pre-set colours (Figure 3(a)). We refer to this as the 'all' method of logging, as it provides the user with completed access to the three identified logging methods.

### 3.3 Data History

Given our focus on minimalist logging, we also wanted to explore how previously logged data are made available to the user. In particular, we wanted to explore the impact of providing an easily accessible view of the last four data points logged. This was achieved through what we refer to as the four data-piece display, which shows the data on the user's smartphone home screen wallpaper in the form of large horizontal banners. Figure 3(a) shows an example where the last four colours that were logged by this user were green, orange, green and red, with red being the least recently colour logged. This provides an ambient, glanceable access to a minimal history, much as with the Lumino system [37].

The app also features a more traditional history of logged colours which shows all previous colour data entries logged chronologically, similar to many existing tools (Figure 3(b)). This all-history display is read as a timeline from right-to-left and from bottom to top, so that the most recently logged colour is in the top left-hand corner of the interface. As noted earlier, this history also includes the topics set by the user.

Recognising that the mechanism of logging becomes part of the reflective experience, we wanted to ensure that the all-history display recorded the logging technique alongside the colour logged. Given the distinct visual style of the different logging mechanisms, shapes were selected to represent the different form of logging, with the full palette represented by a circle, the pre-set palette represented by a square and the camera represented by the lens-like hexagon.

## 4 Method

Our research interest is in how minimalist, flexible self-tracking technology can support different forms of reflection in users' day-to-day lives. Similar to other work in the area, a field study is the best methodological approach to gather data [2, 37, 41]. However, our key interest is in using the app as an artefact to understand how users respond to the app, the design properties represented through the app and how the app is integrated into daily life [33].

Our approach thus draws on the broad set of approaches are categorised under the umbrella term of 'probes.' Our method is not a cultural or technological probe [27], but our field trial is structured to draw on some of the features. In particular, our consideration of reflection draws on the probe approach to provide rich explanations and detail the progression from data to designs or understanding [7]. Furthermore, we note that while there are benefits of our real-world setting in terms of the context of our data collection, the introduction of the app will change the behaviour of our users [33].

### 4.1 Procedure

To aid the comparison between logging techniques, participants were given access to each of the three logging methods individually in turn during the first 3 weeks of the study (weeks 1–3) and then given access to the 'all' version of the app in the final 3 weeks (weeks 4–6) where they could access all three of the logging methods. The order that users received different versions was counterbalanced across the group.

The study was approved by our university's ethics review board, who reviewed all of the study materials. After receiving informed consent, participants were invited to a study setup meeting. They were asked about their personal logging behaviours (if any), before introducing them to the concept of experiential logging with colour. Participants were instructed that they could log around any topic they felt would be meaningful or useful to them, that they could change topics whenever desired and that they could log as much or as little data as they desired.

To explore the impact of the different logging methods, the study took place in two parts. During each of the 3 weeks in part 1, participants had access to one of the large palette, pre-set palette or camera capture versions. Which version participants had access to changed each week, with the order of use balanced across participants. During part 1, participants only had access to the minimal four data-piece history. In part 2 of the study, participants has access to all three versions of the app for the remaining 3 weeks of the study. During part 2, participants could also access the full history of their logged colours across both parts of the study through the all-history display.

A summary of the procedure can be seen in Table 2 and Figure 4.

Three short interviews of 15–20 minutes were conducted with the participants over the phone at the end of weeks 1, 2 and 3. A longer interview of 45–60 minutes was conducted in-person at the end of 6 weeks. At the end of each interview, the participant was asked to send their usage data and

Table 2. Availability of Logging Methods and Displays in Different Periods of the Study

Study period	Availability of logging methods	Availability of displays
Part 1 (weeks 1–3)	Large palette Pre-set palette Camera capture	Four data-piece display
Part 2 (weeks 4–6)	The ‘all’ version	Four data-piece display All-history display

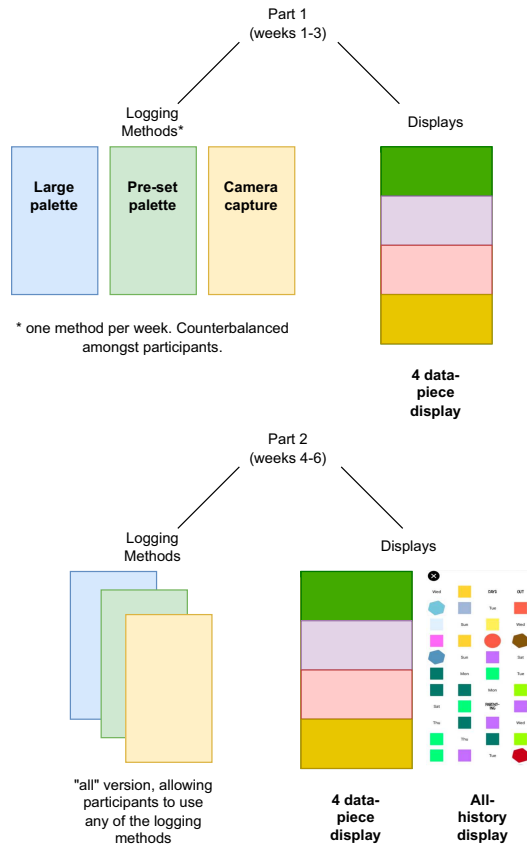


Fig. 4. Outline of the methodology.

also the experience sampling surveys that we conducted from their phone, using a data-transfer function pre-programmed into the app. Through the course of this study, these data were checked over informally between interviews to ensure that the app was functioning correctly.

Participants received a £25 voucher as recompense for taking part in the study.

#### 4.2 Participants

Eleven participants took part in the study, recruited through word of mouth and personal contacts. Participants were asked if they would like to “try out an experience logging device for 6 weeks and discuss your thoughts about it.” Demographic and logging behaviours can be seen in Table 3.



Table 3. Demographics of the Participants

Identifier	Age	Gender	Self-tracking experience	Notable characteristics
P1	25	M	Previous use of MyFitnessPal	Enjoys music
P2	60	F	None	Walking
P3	30	M	Fitbit	Traveling
P4	28	F	None	Horse riding
P5	26	F	Fitbit, Nike Running and MyFitnessPal	Wedding planning
P6	28	F	Tried FitBit and MyFitnessPal	On holiday during trial
P7	33	F	None	Spending time with their young child
P8	28	M	Fitbit, Google Health, Huawei Health	Keen football fan
P9	21	F	None	Moved to stay with parents during Covid-19
P10	21	M	None	Moved to Japan from UK during Covid-19
P11	30	M	None	Enjoys home cooking

F, female; M, male.

### 4.3 Data Collection

There were three different sources of data gathering in the study: interviews with the participants, triggered questionnaires and logging data collected through the app.

**4.3.1 Interviews.** Each participant was interviewed four times. At the beginning of each interview, they were asked to speak openly about their experience of using the app. The interviews also covered users' logging behaviours, their use of the data history, and – where appropriate – compare the version of the app they had been using with previous used versions of the app.

In the final interview in week 6, the focus was on understanding the details of the experiences logged through the app. Participants were asked to describe examples of what experiences they logged and why they chose to log them in the manner that they did. As a prompt for these discussions, participants were asked to get out their phone and discuss specific logged experiences.

**4.3.2 App-Recorded Data.** The app collected data for which colours users logged, when they logged them and which of the three logging methods they used. It also collected data for the topics users set and when they set them.

**4.3.3 Questionnaire Data.** Survey questions were submitted to participants through an experience sampling questionnaire interface. These questionnaires were triggered through specific actions, alongside a rule stating that the question couldn't be asked more than twice a week (see Table 4). This form of live-data capture provided a way of obtaining data about participants' interaction perspectives while they were using the app. As Berkel et al. have suggested, experience sampling facilitates the collection of rich data about micro-interactions that are otherwise difficult to access when users are interacting with mobile phones outside the laboratory [60]. This attribute is particularly valuable given that the focus of this study is on fostering short, fleeting moments of interaction during everyday life, which are likely to be particularly challenging to access.

Table 4. Experience Sampling Trigger Rules

Question type	Schedule condition
Logging a colour	If the user hasn't responded to this question in the previous 3 days
Inactivity	If the user hasn't interacted with the app in the previous 4 days
Topic setting	If the user changes their topic <i>and</i> hasn't responded to this question in the last 3 days
Changing pre-set colours	If the user is using the "Pre-set palette" app <i>or</i> the "All" app <i>and</i> changes their pre-set colours <i>and</i> hasn't responded to these questions in the past 3 days
Comparing logging methods	If it is between week 1 and week 3 in the study <i>and</i> the user logs a colour <i>and</i> the logging a colour question is not active <i>and</i> the user hasn't responded to this question in the previous 4 days
Comparing logging methods	If it is between week 4 and 6 in the study <i>and</i> the user logs a colour <i>and</i> the logging a colour question is not active <i>and</i> the user hasn't responded to this question in the previous 4 days

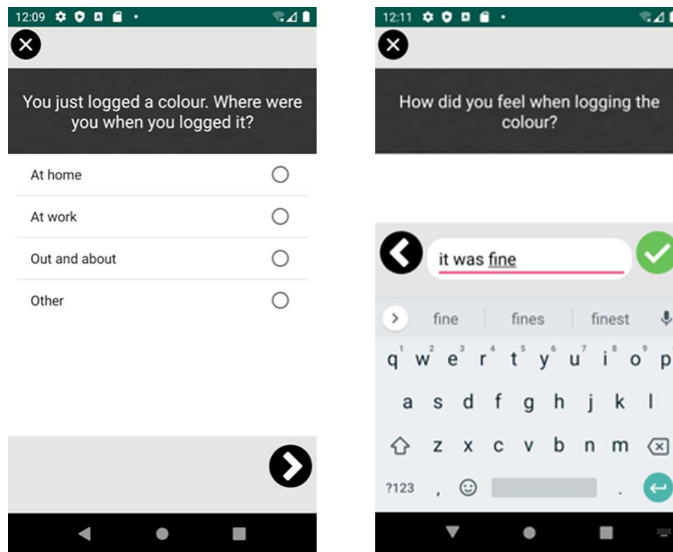


Fig. 5. Experience sampling interface. Multiple-choice question format and free-text entry question format.

The design of the micro-experience sampling system was modeled on the format demonstrated by Ferreira et al. [25]. The app proposes context-sensitive multiple choice survey questions based on a series of pre-set rules (see Table 4). This makes it possible to target the participant with more specialised questions that penetrate deeper into the nuances of user experience than a one-size-fits-all approach would allow. At the same time the system is designed so there is an onus on simplicity and seamlessness. Most of the questions are multiple choice so the participant can respond to them quickly without overly disrupting the flow of their phone-use. The experience sampling survey pops up automatically when participants log colours (see Figure 5).

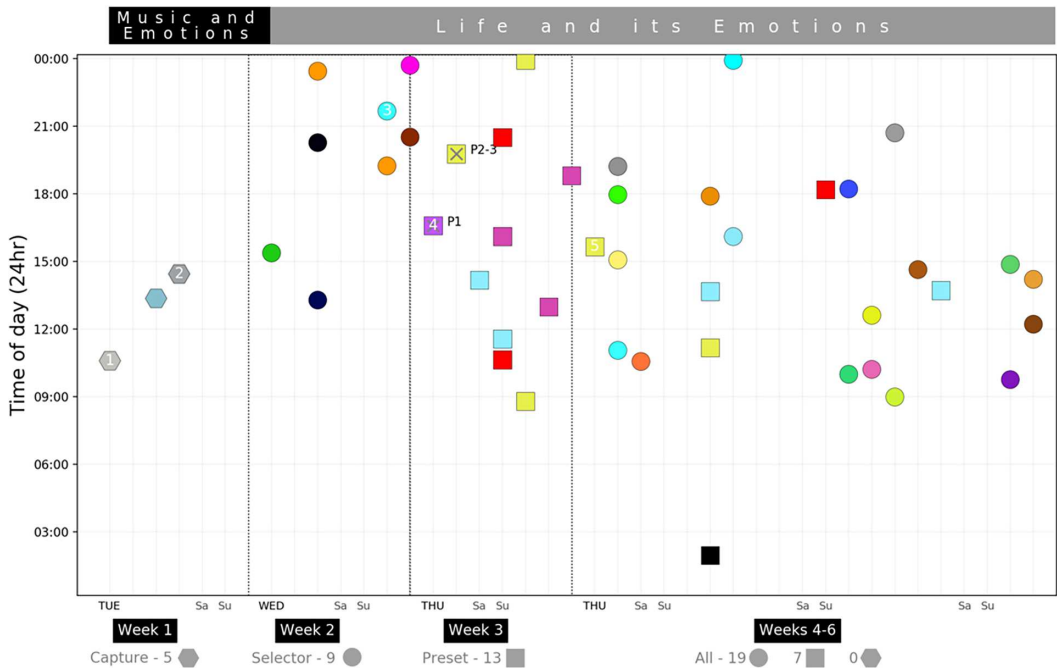


Fig. 6. Timeline chart for P1, showing topic, data logs by time and day of the week and logging method.

#### 4.4 Analysis

Once the participants had completed the study, timeline charts were created (see Figure 6), showing which colours each participant logged, when they logged them, which logging method they used and when they changed topic, providing an overview of their logging behaviour over time. The experience sampling data were compiled into charts depicting the frequency of each survey response. These data were used to inform the thematic analysis of the interview data.

The interview recordings were all transcribed and subjected to a process of thematic analysis [8]. The process was not purely inductive with the coders particularly interested in data about three key themes:

- Reflection – Do users reflect? If so, how do they reflect?
- Logging behaviours – Which topics participants choose to set, when participants log, what prompts them to log and how they decide which colour to log?
- Design characteristics – How the participants compare the different app versions, why they chose to use one version over another in weeks 4–6, how the design characteristics affect the way users log and reflect?

While these themes were a basis for the grouping of the data, our analysis was not constrained by these themes. Themes that didn't fit with these specific areas of interest, but which expressed something important about the way that participants engaged with their experiences, were taken note of. Any chunks of transcribed interview that resonated in this way were initially fitted into a category label 'other' so they could be returned to and unpacked in the next iteration of the analysis process.

Once the data had been sorted into initial themes by the lead author, the other members of the research team considered the themes, also making use of the usage data timeline charts. As

Table 5. Logging Frequency for Our Participants across Time

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Mean	Range
P1	5	9	13	13	7	6	8.3	8
P2	12	12	0	14	2	2	8.4	10
P3	5	6	0	4	5	2	4.4	4
P4	9	10	8	0	0	0	5.4	2
P5	10	5	11	12	4	1	7.2	11
P6	5	4	5	3	2	3	3.6	3
P7	10	8	6	3	2	1	3.8	9
P8	32	7	13	7	7	4	11.6	28
P9	17	13	11	9	4	3	9.5	14
P10	12	7	4	4	3	4	5.6	9
P11	22	0	14	8	9	9	12	6
Mean	12.6	7.4	7.7	7.0	4.1	3.2	9.4	7.3

P11 logged no colours in week 2 because his phone screen broke; P4 logged no colours in weeks 4–6 because she had fallen from her horse, which stopped her wanting to log; P3 logged no colours in week 3 because he was on a busy work-trip in Thailand; P2 logged no colours in week 3. However, this misrepresents her engagement because she was logging in a different way—by changing the pre-set colours on her phone.

a result of these discussions, the themes were adjusted to rectify inconsistencies and the theme titles were refined to increase their accuracy. It was also at this stage that themes in the ‘*other*’ category were subjected to more granular analysis. Where possible, they were accommodated in the newly adjusted theme set or else new themes were created to accommodate them. The full set of themes were then submitted to the secondary researchers once more for a second iteration of the verification step. This prompted discussion about any additional adjustments that might be necessary, and decisions were finalised about the titles and content of each theme.

The analysis of themes relating to reflection was carried out in light of the literature presented. We have argued that reflection has been interpreted as a process that involves both the consideration of, or a change to, knowledge or beliefs. To establish if users reflected when they used the device and how these two processes played into this, it was important to therefore group data according to whether it describes consideration of meaning alone or consideration of meaning and a change in knowledge and beliefs.

## 5 Results

Across the study, participants logged a total of 555 individual colours. The logging data from their phones, paired with their responses in the interviews, indicate that most participants engaged with the study, continuing to log colours throughout (see Table 5). None of the participants dropped out of the study entirely, although some did experience short periods during which they participated less actively, or at least differently.

Across the participants there was an overall reduction in logging frequency over time, with identifiable increases when a new logging method became available. In the final 3 weeks of the study, when there was no experimental intervention, some participants sustained a higher and more consistent logging frequency than others. P1 and P11 are two notable examples of participants who maintained their logging more than others. P2 and P5 are notable for the sharp drop-off in

their logging. Apart from P4, who had a specific reason for having stopped using the app, the participants all logged at least one colour each week.

Colour proved to be an effective medium to support minimalist logging, with participants adopting their own processes to select an appropriate mapping. Some used an extremely limited number of colours (P3/P8), with eight of the participants linking the brightness and hue of the colour to the emotional impact of the experience (P1, P2, P5, P6, P7, P8, P9 and P11). While some participants — notably P1, P2 and P8 — used a systematic mapping of colour to experience (e.g., green is good, red is bad), P2, P7 and P9 described a process where they would select a colour in-the-moment that reflected how they felt without sticking to a particular system:

*“It’s the beginning of the weekend, I’m feeling nice... I’m sat here now. Maybe I would capture the yellow of that throw. But then next weekend I might be feeling similar, but I wouldn’t have that yellow. And it might be that yellow doesn’t feel right that day. You have a range of colours and they all have different tones. So, you’ve got more options.”* (P7, week 6, ‘all’)

Our interviews and survey questions indicate that there were three main dynamics that prompted users to log colours: (1) *seeing the four data-piece display*; (2) *routine logging* and (3) *having an experience*.

In total, participants described 31 instances of logging where the *four data-piece display* acted as a trigger for their decision to log an experience. These examples account for 10 of the participants and cover all 3 of the logging methods.

In their week 6 interview, P8 noted that when his decision to log a colour was prompted by his phone display, he tended to log less-extreme emotional experiences than in other situations. This observation is supported by some of the examples of display-prompted logging that other participants also described. The terminology used implies that our participants did not put much thought into what they were logging when they were prompted by seeing the display:

*“I think I was already on my phone. I think I was in quite a good mood. And so I thought, I might as well log this. I’m not sure I put too much thought into it.”* (P10, week 6, ‘all’)

There is also a sense that in some cases participants felt they *ought* to log something, rather than being intrinsically motivated to do so, when they saw their display.

Two of the participants, P3 and P6 (week 1 only) said that they tried to set themselves *routines*, where they would log colours at a particular time of day. They both said that it made sense for them to log at a particular time of day because it fitted their topics. P3 wanted to rate his day, which made sense to do at the end of the day. P6 wanted to rate her sleep, which made sense to do after she had woken up in the morning.

*“I would then open up the app. And I would kind of, not in depth — wouldn’t sit there for 5-10 minute, but I’d try to do a sort of quick sum up of if something notable happened that would influence my rating. Some days it would be that nothing particularly special happened. It was a pretty routine day at work. And then other days... on the whole that irritating part of the day does not outweigh the other things that were more positive that happened. I’d go, OK, it’s that colour.”* [P3]

It should be noted that unlike the other logging behaviours, there is no indication that routine logging involved mainly mild or strong experiences. Whether or not participants decided to log at the given time-of-day appears to have been independent of the extremity of the experience.

Many of the logging acts from our participants were prompted by an experience, with 10 of our 11 participants discussing such experiences in their interviews (the exception being P3, the routine logger discussed above).

Experience-prompted logging appears to have predominantly occurred when users were experiencing more extreme emotional experiences. P1, P9 and P8 all pointed to this relationship. This forms a contrast with display-prompted logging, which was associated with milder experiences:

*“I normally log intense emotions. So, I think that there have been some times when I don’t think about the app, when it’s a mild emotion. But there have been times when something has been happening, when I’ve been like. Woah! This is great! Or times when I’ve been like, Oh! I want to kill myself! For example, when the code I’m writing just isn’t working. And then I realise, oh I’ve got the app, I’m going to log this.”* (P1, week 6, ‘all’)

P9 noted that she would only log experiences when something out of the ordinary had happened:

*“I think kind of, having a strong emotion or feeling, a bit more than an ordinary feeling. Or if a particular event had happened in the day. I remember playing a board game with my family and that was a nice experience. I logged that. It was when things happened rather than nothing.”* (P9, week 6, ‘all’).

In this second example, P9 is referring to a period of stasis during the Covid-19 lockdown when there was not much happening. The lockdown situation brought about a set of abnormal circumstances, which may have accentuated her tendency to log “*when things happened rather than nothing*,” bringing it into more focus than ordinary life might allow.

## 5.1 Reflection

As discussed, there is substantial debate about what constitutes reflective thinking, in particular how deeply an individual must engage with the meaning of their experiences for the process to be defined as reflective thinking. Some argue that, in addition to considering the meaning of something, the individual must exhibit a fundamental shift in perspectives or beliefs. Others argue for a more subtle conceptualisation of reflection.

To encompass the full spectrum of evidence of reflection from the user study, this section begins by laying out findings associated with *topic setting*, how participants set and changed topics over time, and why they changed them is indicative of how they wanted to engage with meaning through the use of the device in the first place. We then present instances when users *considered the meaning of their experiences* before finally presenting instances when consideration of the meaning of experiences resulted in a *shift in perspectives*.

**5.1.1 Topic Setting.** Participants were able to set the topic they were logging and were free to change it at any time. How and why our participants changed their topics provides insight into how they wanted to express themselves through the app, and as such, which aspects of their life they engaged with.

P2 is an outlier, in that she changed her topic much more frequently than the other participants. While all the other participants used the topic as a broad theme-setter for the group of experiences they were logging, she used it to describe features of the individual experiences she logged. For example, in week 2, she logged a purple colour and simultaneously switched her topic to ‘Frosty Night’ to describe what the colour represented.

Four participants selected a topic at the start of the study and never changed it. Most topic changes were made early on in the study, with only two participants changing their topic after week 3 and no changes from week 5 onwards. This indicates that there was a probing period early on, where participants were working out what they wanted to focus on.



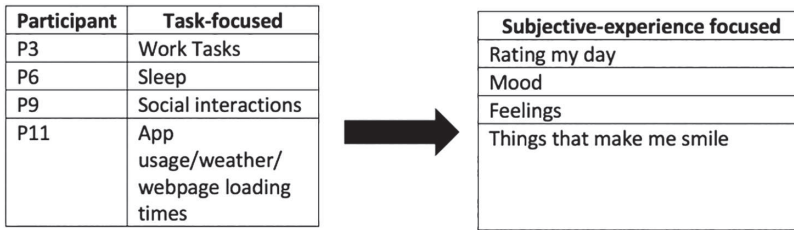


Fig. 7. Transitions from task-focused to subjective experience-focused topics.

Participants who started with topics that had a rigid, functional, task-focus, tended to move towards topics with a broader subjective-experience-focus emphasising emotional aspects, as the study progressed (see Figure 7).

P3 wrote the following on the experience sampling interface in week 1, explaining his own transition to a subjective-experience focus:

*“Early stages, but felt rating work tasks was a bit limited and perhaps focusing on the wrong thing. I’m now having a go and giving a subjective review of how my day has gone, across everything. I feel I often have a tendency to be quite negative, and this could be a way to look across all aspects of the day, and focus on the good things more than I do currently, will see how it goes!”* (P3, experience sampling, day 2)

This shift towards subjective experience-focused topics appears to demonstrate participants trying to find something that felt meaningful for them on a personal level. In interviews, P11, P4 and P2 all spoke about this. For example, in weeks 1 and 2, P11 was struggling to find a purpose for the app and it was only in week 3, once he set the topic to *“things that make me smile”* that he really settled:

*“I think if you’d asked me that a few weeks ago, one of the things I might have asked for would be recommendations of questions that other people have used. But I actually think that would’ve been more restrictive for me. Because the early ones I did didn’t work for me. I needed to explore for a while to find a solution that worked for myself. And it was quite a reflective, personal experience.”* (P11, week 3, camera capture)

**5.1.2 Considering the Meaning of Experiences.** Ten of our participants described 75 situations where they considered the meaning of an experience while logging. After describing the experience, in some cases participants went on to discuss the general impact that the occasions of experiential engagement had on them. One participant (P6) spoke about how using the app made her think more deeply about an experience than she would otherwise:

*“I probably wouldn’t have stopped and thought as deeply about having had a relaxed day if I hadn’t had to pick a colour.”* (P6, week 3, large palette)

Other participants (P4, P7) said that logging colours made them think about their experiences more frequently than they would otherwise. P11 and P1 stated that logging colours generally made them more aware of their experiences. P11 highlighted how it brought his attention to the small pleasures in life:

*“I did definitely think more about the simple pleasures in life and genuinely trying to focus on the positive experiences. Because I think I do have a tendency to overlook them... to be able to think back to them after a day or two and be reflective was good.”* (P11, week 6, ‘all’)

*5.1.3 Shifts in Perspective.* Our data suggest that, in occasional circumstances, using the device affected participants' perspectives or beliefs. In such cases, users were not only considering the meaning of their experiences but were evaluating their deeper implications, leading to a shift in perspective. In some cases, this affected the way they felt about the situation or incited them to take action and change their behaviour.

Most of these instances can be separated into two categories: (a) a change arising from logging a colour and (b) a change arising from looking at the display on the phone.

P2 and P11 both described examples of occasions when logging a colour by using the camera capture logging method made them evaluate the qualities of a current ongoing experience, which resulted in them formulating a perspective on it. For example, P11 reported how the process of logging a colour became stimulus for him to shift his current focus towards the positives of the current moment. At first, he felt that the situation was "gloomy" and "grim." However, having considered the qualities of the moment when logging a colour, he started to focus on the positives and think of the situation as "warm and comfortable."

*"There was a really vibrant orange colour that I recorded once and that was a train ticket. It was really close to Christmas, it was a really, really long week; I'd been working hours and hours and hours. I'd just done like a 12-hour shift and it was pitch black dark. There was no light, and it was really gloomy and everything was sort of grim. But being sat on the train was such a warm and comfortable feeling. And I recorded the colour of the ticket and that put me in a totally different frame of mind, because I was reflecting on it short term, I'm on my way home I am currently warm and comfortable. And I had found a warm and vibrant colour in the gloom as well. You know, it was a bright orange colour of the train ticket."* (P11, week 6, camera capture)

For P3, rather than changing their perspective on a single experience, logging changed his perspective on the series of experiences he had had across a day. As a routine logger, P3 would routinely choose a colour that represented his day as a whole. When considering the broader impact of this process of picking the daily colour, he reported that it had given rise to a general shift in the way he perceived everyday life, affecting his behaviour:

*"I think that coming home at the end of the day and chatting to my fiancé about how the day has been. There would often be a tendency to go straight for the really bad thing and just go straight to talking about that. But I do think that even if I've not always been logging it, now I'd maybe have that conversation differently. And I'd say, oh no, it was pretty good, just these things few things happened that were a bit annoying. Because I do think that it has changed my mindset to an extent, in trying to find the positive instead of going straight to [the] really bad thing that happened."* (P3, week 6, 'all')

Four participants (P11, P1, P6 and P8) spoke about the effects that displayed data could have on their perspectives. One participant, P11, spoke about this generally, saying that he would look at the colours on his display, representing "things that make me smile," which changed his perspective on how his day had been, giving him a more positive outlook on the day:

*"By recording positive experiences, it put me more... It made me more mindful of that positive experience as and when it was happening. And it allowed me to... whenever I opened my phone, I would see those 4 colours and it would put me in, or make me aware of those positive experiences. You know... even just reflecting on positive experiences puts me in a positive mood."* (P11, week 6, all)

There are also examples where, having drawn an insight from the relationships between their data, participants decided to take action and change their behaviour:

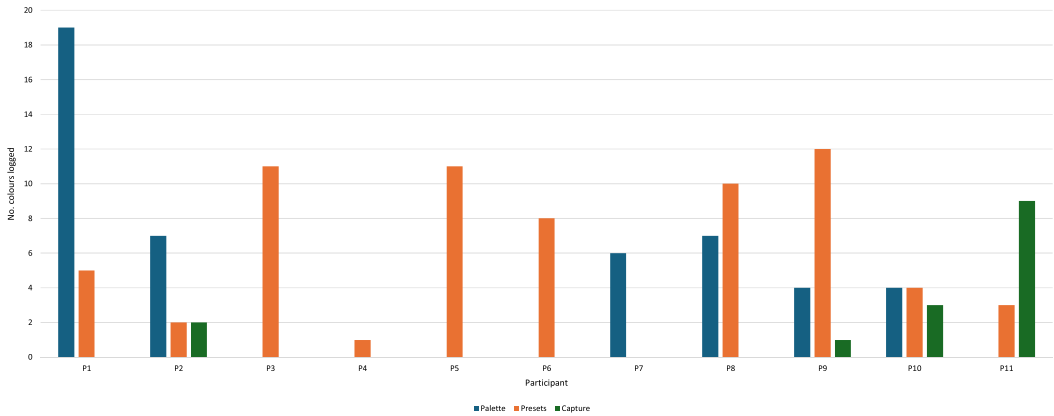


Fig. 8. Number of colours logged, by logging method, in weeks 4–6.

*“Mostly when it was all dark. I thought, OK, maybe I need to do something... I was thinking, what should I do to change to a brighter palette of colours?”* (P1, week 6, all)

*“I had a similar experience this week as last week. I could see the background of my phone and see that it was all red and I thought, ooh my God!... It did instigate me to think, OK, come on! Try to do something positive.”* (P8, weeks 4–6, all).

In both cases, seeing the negative connotations of a series of colours on their display seems to have incited the participant to think about doing something positive and uplifting.

## 5.2 Design Characteristics

Two key design characteristics – the method of logging and the history display – were identified by our participants as affecting the way they associated meaning with their logging behaviours.

**5.2.1 Opinions on Logging Methods.** In the first half of the user study, participants were given access to each logging method (pre-set palette, large palette, camera capture) in turn, in successive weeks. Then, in the final 3 weeks they had access to all of the logging methods at once. Usage data, showing how often they utilised each of the different logging methods in the final 3 weeks of the study, provide an insight into their logging method preferences. As the chart (Figure 8) shows, participants generally had one logging method that they employed more frequently than the others, indicating a preference for this method. There was only one participant, P10 who used the different logging methods with similar frequency.

Four of the participants (P1, P3, P6, P8) said that they liked the simplicity of the pre-set palette because they could choose colours quickly:

*“I’ve tended to use the one where there’s the five pre-set choices, because when you open it, it’s the most straightforward one to select.”* (P3, week 6, ‘all’)

For other participants the small range of colours provided by the pre-set palette was a limitation on their ability to express themselves. For P7 and P11 the simplest logging method was the one that allowed them to express themselves openly as well as easily. However, they had different perceptions on which of the other logging methods fulfilled these criteria the best. For P11, the camera capture logging method was the best, while for P7, the large palette was the best option.

Participants also spoke about how the different logging methods made them engage with their experiences differently. Two of the participants, P9 and P3 spoke about how choosing colours with

the large palette made them analyse what their experiences meant more deeply than when they used the pre-set palette or camera capture methods. P9, compared the large palette she used in week 2 with the camera capture method she had used the previous week, noting that logging with the former meant the colour ‘stuck’ in her mind:

*“Yeah, I picked a blue because I was singing at church because that was the colour of the cassocks. And I liked that I could find that exact blue. So I think that was useful. But then again, I would have been able to do that easily with the last version. But then, because I was having more... I was actually having to find it and do it... and almost make the colour, it stuck in my mind more than if you just log it in an instant and that’s it, maybe.”* (P9, week 2, large palette)

On the contrary, participants sometimes picked colours impulsively without having considered their choice much at all. Two of the participants mentioned that they would tend to log with the pre-set palette when they were in a more extreme situation, when they didn’t feel like making a considered, rational choice.

*“You see yellow is quite a nice colour... But at the time it just made me feel about being sick... Because I was feeling rubbish and there were just a few that were there [points at the pre-set colours on the app] I just thought that will do... Because I couldn’t be bothered.”* (P2, week 6, ‘all’)

**5.2.2 Data Displays.** One of the main design characteristics was how the history was displayed to users; limited to the four data-piece display for the first 3 weeks, and the larger all-history data display during the last 3 weeks of the study.

The four data-piece display was visibly present every time participants looked at the home screen of their phone. P3, P6 and P7 all noted that they didn’t always see the display, despite its presence. P3 and P9 both said that they were more likely to notice the display when it showed unexpected colour combinations. P3, who generally logged similar colours, said he didn’t notice his colours when they formed a single homogeneous block:

*“I guess that because I’ve been putting in the same thing (colours), the background is not as... it doesn’t jump out at me as much as it would if there had been, maybe a red in there or a yellow, or maybe a darker.”* (P3, week 6, ‘all’)

However, when participants did notice the four data-piece display it could sometimes incite strong reactions. P7 and P9 both described occasions where, having encountered colours that were incongruous with how they felt, they had an urge to change them all:

*“There was a day, when I was really, really tired. And I hadn’t felt well at all. And so I changed the whole screen because it was too bright. I kept colours that I liked but I went for a more pastel version of it. So it was like, a calmer screen. Not as intense... it was funny. I was very aware of why I was doing it. I was very aware of feeling quite tired and flat and that the colours were too intense to look at. It was a calmer screen, a calmer thing to look at.”* (P4, week 2, large palette)

There was a clear contrast between how users interacted with the four data-piece display and the all-history display. The usage data show that excluding the time when they first received the all-history display during the week 3 interview, only one participant (P2, 4 times) looked at this display more than once. When asked about their use of the all-history display in the interviews, most of the participants said they generally just forgot to use it or didn’t feel the need to use it in their logging practice. P8 noted how:

*“I didn’t find the overview screen that useful. But I did find this background screen with the few logs [4-colour display] quite useful. It gave me the most recent timeline of what has been going*

*on. And made me think, OK, maybe I need to take action to change things. This was interesting.”*  
(P8, week 1)

Meanwhile, P2, the one participant who did look at the colours on the all-history display, said that she couldn't remember what they represented.

## 6 Discussion

Our study was motivated by our two key research questions:

- RQ1: How can minimalist, flexible self-tracking technology support experiential logging during everyday life?
- RQ2: How do users reflect when they use experiential logging devices and what does this tell us about how reflection should be conceptualised within PI?

Our discussion of the findings reflects these questions, first focusing on RQ2. In our review of the literature we argued that, despite often being cited in self-tracking research, reflection has rarely been well-defined, and there has been little attempt to describe the ways in which different design features affect the way users reflect. In the first part of the discussion, we consider how our findings support a new way of considering different forms of reflection with regards to self-tracking technology. The second half of the discussion focuses on RQ1, considering the success of designing for valued minimalist flexible self-tracking experiences.

### 6.1 Reflection

The Chromatize study has demonstrated how users might use a minimalist experiential self-tracking device during everyday life. We use the term “experiential self-tracking” here, rather than the more classic ‘self-tracking,’ to foreground the design objective of encouraging open tracking of topics relevant to user’s lives. Participants in the study seem to have engaged with this objective, logging data on a diverse range of everyday life experiences, switching focus as they sought to engage with what was most meaningful for them at a given time.

Based on our findings, we argue that the existing frameworks of reflection in HCI do not map well onto people’s experiences of the impact of logging. Our findings highlight a distinction between reflection that leads to general self-awareness, affecting knowledge in subtle ways (which we term *soft* reflection) and reflection that leads to a shift in perspective on something, affecting knowledge in a more significant way (which we term *hard* reflection). The second of these relates more closely to how reflection is generally defined in PI literature, with its emphasis on the role of reflection as a transformational process that incites behavioural change. The advantage of this formulation is its ease of application, over the complexity of the models presented by Baumer [5] and Fleck and Fitzpatrick [26].

Hard and soft reflection do not map onto the occurrence of when the reflection occurs as with Schön’s model of reflection-in-action or reflection-on-action [54]. The key difference is in the nature of the change that occurs as a result of the reflection, be that a more general self-awareness (soft) or something directly leading to change (hard). In unpacking the distinction below, we provide examples from our data of both reflection-in-action and reflection-on-action when discussing hard reflection.

**6.1.1 Soft Reflection.** Our findings contain many cases of participants associating logging acts with emotive experiences. In the process of choosing which colour to log and what to associate it with, they considered the meaning of what they were experiencing. Participants also sometimes considered the meaning of their experiences while revisiting previously logged data, focusing on

the meanings of individual data pieces or the general experiential landscape of the four colours as a group.

Although small, fleeting moments of soft reflection don't involve the kind of significant, transformational shift in perspective that has been a focus in the self-tracking literature, they can still be valuable for users. Participants remarked that they found it useful to check-in with themselves, focusing on the meaning of little moments in everyday life that would usually pass them by. They rarely pointed to an individual insight as being of particular importance but said they liked the general sense of experiential awareness that they gained from manually logging data, which made them feel like they knew more about their everyday lives than they would do otherwise. It appears that the effects of these instances of soft reflection, though small and insignificant individually, accumulate, and have a significant impact on users' knowledge about their experiences as a whole.

This *soft reflection*, bears some resemblance to level 1 reflection from Fleck and Fitzpatrick's conceptual framework (reflective description, see Table 1) [26]. The key similarity is that neither form of reflection requires a shift in perspectives. However, where the conceptualisations differ is that level 1 reflection involves a stage where the individual justifies their interpretation of the experience. This is not part of what we understand by soft reflection, where we see the focus on a more casual awareness of people's understanding of themselves and their emotive state, prompted by the logging experience.

**6.1.2 Hard Reflection.** Our participants also discussed a small number of examples which appear to have involved a shift in perspectives. These instances of *hard reflection* are examples which bear a closer relationship to how reflection is traditionally framed in the self-tracking literature.

There appear to have been two different periods during the interaction with the app when this happened: reflection while revisiting data and reflection while manually logging data. This tallies with how reflection has been framed in the literature [42].

The process of selecting data that summarises a series of experiences over time can give rise to hard reflection. The clearest example is when P3 would sit on his bed at the end of the day and think about which data piece, from his set of five colour options on the pre-set palette, summing up a range of experiences from across the day. The process of deciding where, on-balance, his day fitted best on a scale, hence compressing a range of experiences into a single data-piece, caused him to explore the relationships between the experiences. This is an instance of hard reflection because it resulted in him having a shift in perspective, changing his attitude to how his day had been to a more positive one.

This kind of reflection-on-action [54] process can be associated with level 2 from Fleck and Fitzpatrick's conceptual model (dialogic reflection, see Table 1) [26]. This is because the user is exploring the relationships between multiple experiences, as opposed to explaining a single experience.

The process of logging an individual experience as it is unfolding can also support hard reflection. The key difference between this form of reflection is that instead of reflecting on the relationships between a series of experiences, the user is reflecting on the qualities of the individual ongoing experience that they are having in-the-moment. A key example is when P11 was logging the orange colour of his train ticket and started to think more positively about the situation — the warmth and comfort of the train. In such examples, the colour logging process seems to direct the users' attention to the qualities of the situation, to the extent that it affected their perspective on the situation. The process of logging the colour made the experience explicit in some way, focusing their minds on its discrete features and their broader meaning in-the-moment.



This kind of reflection-in-action [54] process can be associated with level 1 reflective description in Fleck and Fitzpatrick's model. The user logs a colour, which they associate with a quality of an individual experience (e.g., the warmth of the train), explaining their interpretation of the experience back to themselves. It is the process of logging and interpreting the colour that draws the meaning out of the experience. However, in the case of the reflective process described here — distinct to level 1 reflection and from soft reflection — logging the data appears to have brought about a shift in perspective.

*6.1.3 Implications of the Framing of Reflection.* Reframing considerations of reflection to *soft* and *hard* reflection have three main implications: on how (1) reflection-in-action, (2) reflection at the time of manual logging and (3) reflection on displayed data are each perceived in the self-tracking technology field.

In the self-tracking literature, there has been a tendency to explain how reflection emerges when users log data by associating it with Schön's concept of reflection-in-action [3, 21, 59]. However, we argue that users can not only reflect-in-action when they log data but also reflect-on-action. Which of these takes place is dependent on whether they are engaging with their experiences in-the-moment or retrospectively when they log the data. If users log data that represent an ongoing experience, they are reflecting-in-action, but if they are looking back into the past and logging data that represent experiences that have already happened, they are reflecting on-action. This is an important distinction to make because of the different effects that these two different forms of reflection can have on users. When users reflect-in-action while logging data, they are in a position to intuitively act on what they have learned as it is ongoing. Their use of the device can affect the individual experience as it is unfolding by, e.g., helping them regulate their emotions or improve their mindset. When users reflect-on-action while logging or revisiting data, meanwhile, the experiences they are considering have surpassed, so they cannot be directly affected. However, users can take a minute to summarise, analyse and reflect on multiple experiences as a group and therefore explore deeper reflective insights by bringing together different pieces of information.

The way that reflection has been presented above has potential implications for how it should be perceived conceptually in the self-tracking technology community. A first key point to note is that soft reflection emerging from manual data logging appears to play a more important role in the way users engage with and learn about their everyday lives than has perhaps been appreciated in the literature. Although moments of soft reflection have little significance individually, they can emerge frequently and can have a valuable effect on self-knowledge as a whole, raising users' general sense of awareness of the experiences that populate their everyday lives.

By contrast, examples of hard reflection can be very sparse, in particular instances of critical or fundamentally transformative reflection, the kind of reflection that most of the self-tracking literature focuses on. One general factor that might go some way to explaining this is that it is challenging to develop a methodology for pinning down clear, concrete examples of how users reflect. It is possible that there were more instances in the user study when users performed hard reflection than is implied by those reported on, and that participants were unable to recall or describe them all clearly in the interviews. A second explanation is that the particular system designs trialed in this study, given their minimalism, weren't as conducive to transformative reflection as typical self-tracking technologies. A third, and perhaps the most profound possible explanation, is that during everyday life it is actually rare that people have moments of significant, transformative insight where they consider the fundamental basis of their beliefs in light of evidence. This reflects Slovak's claim that transformative reflection is hard and rarely emerges from users simply being exposed to data [56].

Given the focus on data as a stimulus for behaviour change in the literature, it is understandable that there has been a lot of focus on hard, transformative reflection in the past and less attention on the role of manual logging as a stimulus for soft reflection. It is important, therefore, to broaden our scope and consider the value of a range of different forms of reflection emerging from different kinds of interaction, including manual data logging and other forms of interaction, and how they can serve users with different personalities and reflective approaches in different ways, supporting behaviour change but also self-knowledge development more broadly.

## 6.2 Design Qualities

We defined a minimalist self-tracking device as one that makes logging data and revisiting it after it has been logged easy for users, by ensuring that these processes require minimal interaction. The Chromatize apps made use of an ambiguous media – colour – and an open choice of topic, resulting in a minimalist flexible self-tracking tool.

Four different aspects of these design qualities appear to impact how our participants experienced reflection: (1) the flexible media, (2) the physical act of logging; (3) prompting the decision to log and (4) the act of revisiting data.

**6.2.1 Flexible Media.** The open form of topic setting appears to have supported users in tracking a wide range of different types of experiences, moving from task-focused logging to subjective experiences. This aligns with similar observations from other flexible PI systems such as bullet journaling [3] and data physicalisation platforms [59]. However, there is a difference in that while users of these systems tend to have set themselves specific, narrowly defined topics (e.g., sleep [39], mood [3], meditation [59]), users of the Chromatize device set broadly defined open topics, relating to their general subjective experience (e.g., life and its emotions, things that make me smile), topics, which in themselves are more open and offer greater flexibility.

This may be due to selecting colour as a data type, which users developed into both expression-driven and measurement-driven approaches to logging. Expression-driven loggers perceived the colours they logged as individual symbols, representing the qualities of their experiences (e.g., blue as a symbol representing a visit to the seaside), while measurement-driven loggers perceived a set of colours as a quantitative scale for grading their experiences (e.g., a spectrum of colours differentiated by brightness, considered as a rating scale for sleep quality).

The ambiguity of colour data gives users a licence to express themselves freely according to what makes sense for them as an individual; they aren't limited to expressing pre-defined meanings that are prescribed to them by convention and can associate colours with whatever makes sense to them. Such associations may make use of the relationship between colour and emotion [29, 35]. Further work should consider other forms of ambiguous media available. While colour has been used in a variety of systems – due to its relationship with emotion – other media, such as sounds or processed imagery could provide a similar prompt for reflection that may better suit some users.

**6.2.2 Simplifying the Logging Interaction.** As detailed when discussing the design, the Chromatize app logging methods were designed to make the physical act of logging data as simple and direct as possible. Similar to the observations drawn by Ferrario et al. [24] and Choe et al. [11] in their work on systems for minimalist logging, users appear to have found the physical act of logging data easy with the Chromatize app because they could easily log their data while doing their everyday life activities.

The manner in which users log colour data and how this matches up with their intended style of logging can affect the perceived simplicity of the logging act. Participants were asked which of the logging methods they preferred, with all responses emphasising the importance of simplicity. However, users' perceptions of what simplicity meant wasn't determined by the physical simplicity

of the logging interaction, i.e., the one that required the least number of taps or button presses. It was the logging method that allowed them to log in their preferred style most easily. As noted in our results, this could be either expression-driven or measurement-driven logging. Those participants who wanted to express themselves most openly preferred the camera capture and large palette methods of logging and those who wanted to measure their experiences over time preferred the pre-set palette logging method. It appears that in each case their preference was shaped by which of the logging methods let them pursue their favoured logging style in the simplest way. In the order of priorities, their logging-style preference came first before the physical simplicity of the interaction.

This range in preferences indicates there could be benefits in exploring providing personalised and customisable options for logging. This shouldn't be limited to the logging media (e.g., colours, sounds and images), but also the mechanism for recording that log, as different users extracted different value from different logging modalities. This would potentially limit the consistency of what any given log means, actively useful in the context of open-ended self-tracking, but more problematic in other areas of self-tracking where the meaning of a log needs (e.g., pain logging [46]).

**6.2.3 Prompting the Decision to Log.** Our study results showed that users sometimes log data about their experiences because they are self-motivated to do so. They experience something and decide to log it because they feel it is important to do so. Interview data indicated that this kind of self-motivated logging appears to happen when users experience an extreme emotion, one of which they are already aware, but which they want to record and make explicit in data. By contrast, subtler, more mundane, everyday experiences often pass by without it occurring to users that they should log them and without the user necessarily recognising the experience at all. On these occasions, it appears that PI devices can play an important role, by intervening and prompting users to engage with their milder, less-noticeable experiences — clearly related to *soft* reflection.

This underlines the importance of design features that draw the users' attention to what they are experiencing and prompt them to log data. Our findings indicate that, despite the absence of a notification system, users were still prompted to log data, in particular by the four data-piece display, which acted as a passive prompt.

There were a number of occasions when the four data-piece display on the wallpaper of the user's smartphone served as a passive prompt. It functioned somewhat like an ambient display, with a persistent visual presence on the periphery of the user's attention. Whenever the user looked at their phone throughout the day, there was a chance that they would see the display. Because of its constant availability and the simplicity of the data visualisation, it seems to have had what Rogers et al. termed 'glanceability' [52], subtly drawing the user's attention and prompting the user to log an experience. There was a certain level of spontaneity to this kind of prompting. Rather than demanding engagement at a particular point in time, it proffers an ongoing opportunity for engagement, which means users can be prompted to log at any time, rather than just in the moments when the notification was pre-programmed to be triggered.

It should be acknowledged that based on the evidence from our study with the Chromatize app, it is difficult to say how effective these passive prompts would be over the longer term. There is a possibility that after a while the background would completely blend-in, particularly if users stopped logging colours, and therefore stopped renewing its appearance. Likewise, users could become so familiar with the form of the device that it no longer catches their eye. A longer-term study would be required to analyse these effects.

**6.2.4 Supporting Revisiting of Logged Data.** Our findings show that our participants only revisited their data and engaged with its meaning on the four data-piece display, not using the all-history functionality at all. This challenges the notion that revisiting data are integral to the experience

of using self-tracking technology [43]. Our data provide some insight into why our participants engaged with the minimalist four data-piece display instead.

The first aspect was simplicity; participants could both find the four data-piece display more easily, and interpret what the information displayed, compared to the all-history display. This insight supports the argument of Rapp and Cena who noted that inexperienced self-trackers have difficulty in understanding complex data displays [47].

Perhaps more meaningfully, our participants found that the limited history was more relevant during their everyday life. Participants explained that they wanted to focus on data representing their most-recent experiences. This makes sense if we reconceptualise the purpose of most experience tracking to support *soft* over *hard* reflection, given that the focus is no longer on transformation of perspective, but on a better understanding of everyday life.

### 6.3 Beyond Open-Ended Self-Tracking

Through our work we make two main contributions. Firstly, we have provided insight into the benefits of supporting minimalist tracking using flexible media. Secondly, we have used our data to develop a different perspective on reflection which is easier to apply in the context of self-tracking.

Both contributions extend to areas of life outside of open-ended self-tracking. Reducing the burden of tracking while supporting rich reflection applies to virtually any area of life which includes goal setting, be that physical activity and exercise, nutrition, sleep, or mental health [18, 55]. Beyond health and wellbeing, other areas could include workplace productivity, location, personal finances, sustainability and teaching or learning [22].

For example, Rapp and Tirabeni have been working with elite and amateur athletes to understand their tracking needs [49, 50]. While the main focus is on automated tracking to support performance, there are aspects of tracking – such as mood or sleep – which are related to their performance and yet there are difficulties in interpreting and making use of the data. Minimalist flexible tracking – particularly our experiences with minimalist visualisation – could be of use here, particularly framing the form of tracking such that it “closely interacts with the body” [50].

In a completely different context, Schroeder et al. have noted in their investigation into the tracking of migraine symptoms that tracking goals change over time, and that personalisation and customisation are key, as individual experiences differ [55]. While ambiguous media may be helpful to support self-reflection of migraine experiences, the need to share data in a meaningful fashion does mitigate towards a more standard scale. However, our approach to minimising data collection could prove helpful (as it has with post-surgical pain collection [46]), as can our insights into the value of minimalist histories.

This aligns with the arguments made by Ankrah et al. in their work with children with attention deficit hyperactivity disorder, considering smartwatch health data: “Systems that are otherwise excellent may run the risk of infantilizing people and limiting their ability to act on their own health information in the face of inaccessible or incomprehensible data visualizations... Designers and researchers must instead incorporate accessible and engaging data visualizations” [1]. We have contributed to the ongoing development of a clearer understanding of how to make self-recorded data of greater value to the user.

## 7 Limitations

We have already noted that it is challenging to develop a methodology for identifying clear, concrete examples of how users reflect. Relatedly, we acknowledge that it is challenging to determine whether our participants’ accounts of reflective experiences were supported by the app, or triggered either by the experience sampling questionnaire or the interviews.

Our participants highlighted several features of the app as triggers of reflection, including topic-setting (Section 5.1.1), the four data-piece display (Section 5.1.3) and the ability to express themselves openly through colour (Section 5.2.1). This provides some indication that it is the design properties expressed through the *Chromatize* app that were supporting participants to reflect, rather than the data collection techniques themselves.

Furthermore, the data collected through the experience sampling questions were sparse. While sufficient to frame the analysis of the interviews, it did not highlight any specific reflective practices; nor did participants discuss it as a significant property of the app during the interviews. Our participants were selective when describing their reflective experiences, easily identifying specific cases of both hard and soft reflection through their interviews. We are therefore comfortable in arguing that our participants' reflective practices were not a result of the experience sampling.

The usage data of the app demonstrates a gradual decline over time. A limitation of our data is that we have little insight into what design factors had an impact on the level of engagement. The continual use could imply that participants had developed past the novelty of their tracking behaviours and had reached a tracking level suitable for their selected topic. An alternative explanation is that the app was a novelty but was unsuitable for supporting longer-term tracking. Further work is needed to better explore the long-term value of minimalist flexible self-tracking to users.

## 8 Conclusion

The design, evaluation and analysis of the *Chromatize* app has helped to highlight how people might benefit from personal self-tracking systems which are less data-centric than typical self-tracking technologies. It has shown that low-burden devices which let people manually log data that expresses their subjective experiences may serve as powerful meaning-making tools. The results from the field study suggest that in some cases the physical act of manually logging data in this way may be enough for people to learn about themselves and make sense of complex, nuanced subjective experiences. It is perhaps, therefore, not always necessary for systems to manufacture meaning for users by collecting, integrating, analysing and then presenting large sets of personal data to them about their behaviour in the past. Indeed, we argue that distinguishing between hard and soft reflection provides a straightforward conceptualisation for the value that experiential self-tracking technology brings. By highlighting the difference, we also bring softer forms of reflection into focus and place it on an equal standing with critical and transformative reflection. The soft/hard terminology might therefore encourage researchers to see softer forms of reflection as more valuable within the field of self-tracking technology.

## Acknowledgements

The authors thank all of the study participants.

## Authors' Statement

This submission has not been previously submitted. The authors have not previously published any of the work in this manuscript. The contribution of this article does not overlap with the contribution made by any article previously published by the authors.

## References

- [1] Elizabeth A. Ankrah, Franceli L. Cibrian, Lucas M. Silva, Arya Tavakoulnia, Jesus A. Beltran, Sabrina E. B. Schuck, Kimberley D. Lakes, and Gillian R. Hayes. 2023. Me, my health, and my watch: How children with ADHD understand smartwatch health data. *ACM Transactions on Computer-Human Interaction* 30, 4 (Sep. 2023), Article 59, 25 pages. DOI: <https://doi.org/10.1145/3577008>



- [2] Amid Ayobi, Paul Marshall, and Anna L. Cox. 2020. Trackly: A customisable and pictorial self-tracking app to support agency in multiple sclerosis self-care. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, 1–15. DOI : <https://doi.org/10.1145/3313831.3376809>
- [3] Amid Ayobi, Tobias Sonne, Paul Marshall, and Anna L. Cox. 2018. Flexible and mindful self-tracking: Design implications from paper bullet journals. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, 1–14. DOI : <https://doi.org/10.1145/3173574.3173602>
- [4] Matthew Barker. 2021. *Designing Personal Informatics Technologies for Meaningful Engagement with Experiences During Everyday Life*. Open University. DOI : <https://doi.org/10.21954/ou.ro.0001366e>
- [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 Conference on Designing Interactive Systems (DIS '14)*. Association for Computing Machinery, New York, NY, 93–102. DOI : <https://doi.org/10.1145/2598510.2598598>
- [6] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 5, (Nov. 2013), Article 30, 27 pages. DOI : <https://doi.org/10.1145/2503823>
- [7] Kirsten Boehner, Janet Vertesi, Phoebe Sengers, and Paul Dourish. 2007. How HCI interprets the probes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, 1077–1086. DOI : <https://doi.org/10.1145/1240624.1240789>
- [8] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. DOI : <https://doi.org/10.1191/1478088706qp063oa>
- [9] Yunan Chen. 2010. Take it personally: Accounting for individual difference in designing diabetes management systems. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems (DIS '10)*. Association for Computing Machinery, New York, NY, 252–261. DOI : <https://doi.org/10.1145/1858171.1858218>
- [10] Justin Cheng, Akshay Bapat, Gregory Thomas, Kevin Tse, Nikhil Nawathe, Jeremy Crockett, and Gilly Leshed. 2011. GoSlow: Designing for slowness, reflection and solitude. In *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. Association for Computing Machinery, New York, NY, 429–438. DOI : <https://doi.org/10.1145/1979742.1979622>
- [11] Eun Kyoung Choe, Bongshin Lee, Matthew Kay, Wanda Pratt, and Julie A. Kientz. 2015. SleepTight: Low-burden, self-monitoring technology for capturing and reflecting on sleep behaviors. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, 121–132. DOI : <https://doi.org/10.1145/2750858.2804266>
- [12] Eun K. Choe, Bongshin Lee, Haining Zhu, Nathalie H. Riche, and Dominikus Baur. 2017. Understanding self-reflection: How people reflect on personal data through visual data exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '17)*. Association for Computing Machinery, New York, NY, 173–182. DOI : <https://doi.org/10.1145/3154862.3154881>
- [13] Ming K. Chong, Jon Whittle, Umar Rashid, and Chee S. Ang. 2015. Cue now, reflect later: A study of delayed reflection of diary events. In *Human-Computer Interaction – INTERACT 2015*. Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.), Springer International Publishing, Cham, 367–375.
- [14] Anna Cox, Jon Bird, and Rowanne Fleck. 2013. Digital epiphanies: How self-knowledge can change habits and our attitudes towards them. In *Presented at: The 27th International British Computer Society Human Computer Interaction Conference*. BCS, Swindon, UK, 1–4. <https://discovery.ucl.ac.uk/id/eprint/1416686/>
- [15] Andrea Cuttone, Michael K. Petersen, and Jakob E. Larsen. 2014. Four data visualization heuristics to facilitate reflection in personal informatics. In *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*. Constantine Stephanidis and Margherita Antona (Eds.), Springer International Publishing, Cham, 541–552.
- [16] Tabby Davies, Simon L. Jones, and Ryan M. Kelly. 2019. Patient perspectives on self-management technologies for chronic fatigue syndrome. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, 1–13. DOI : <https://doi.org/10.1145/3290605.3300452>
- [17] John Dewey. 1910. *How We Think*. American Psychological Association, Washington, USA. DOI : <https://doi.org/10.1037/10903-000>
- [18] Tina Ekhtiar, Armagan Karahanoglu, Rúben Gouveia, and Geke Ludden. 2023. Goals for goal setting: A scoping review on personal informatics. In *Proceedings of the ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, 2625–2641. DOI : <https://doi.org/10.1145/3563657.3596087>
- [19] Elisabeth T. Kersten-van Dijk, Joyce H. D. M. Westerink, Femke Beute and Wijnand A. IJsselstein. 2017. Personal informatics, self-insight, and behavior change: A critical review of current literature. *Human-Computer Interaction* 32, 5–6 (2017), 268–296. DOI : <https://doi.org/10.1080/07370024.2016.1276456>

- [20] Chris Elsdén, Mark Selby, Abigail Durrant, and David Kirk. 2016. Fitter, happier, more productive: What to ask of a data-driven life. *Interactions* 23, 5 (Aug. 2016), 45. DOI: <https://doi.org/10.1145/2975388>
- [21] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: Exploring visual cuts of personal informatics data. In *Proceedings of the Conference on Designing Interactive Systems (DIS '14)*. Association for Computing Machinery, New York, NY, 667–676. DOI: <https://doi.org/10.1145/2598510.2598558>
- [22] Daniel A. Epstein, Clara Caldeira, Mayara C. Figueiredo, Xi Lu, Lucas M. Silva, Lucretia Williams, Jong H. Lee, Qingyang Li, Simran Ahuja, Qiuer Chen, Payam Dowlatyari, Craig Hilby, Sazedra Sultana, Elizabeth V. Eikey, and Yunan Chen. 2020. Mapping and taking stock of the personal informatics literature. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (Dec. 2020), Article 126, 38 pages. DOI: <https://doi.org/10.1145/3432231>
- [23] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, 731–742. DOI: <https://doi.org/10.1145/2750858.2804250>
- [24] Maria A. Ferrario, Will Simm, Adrian Gradinar, Stephen Forshaw, Marcia T. Smith, Thomas Lee, Ian Smith, and Jon Whittle. 2017. Computing and mental health: Intentionality and reflection at the click of a button. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '17)*. Association for Computing Machinery, New York, NY, 1–10. DOI: <https://doi.org/10.1145/3154862.3154877>
- [25] Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, Louise Barkhuus, and Anind K. Dey. 2014. Contextual experience sampling of mobile application micro-usage. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '14)*. Association for Computing Machinery, New York, NY, 91–100. DOI: <https://doi.org/10.1145/2628363.2628367>
- [26] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on reflection: Framing a design landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction (OZCHI '10)*. Association for Computing Machinery, New York, NY, 216–223. DOI: <https://doi.org/10.1145/1952222.1952269>
- [27] Connor Graham, Mark Rouncefield, Martin Gibbs, Frank Vetere, and Keith Cheverst. 2007. How probes work. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces (OZCHI '07)*. Association for Computing Machinery, New York, NY, 29–37. DOI: <https://doi.org/10.1145/1324892.1324899>
- [28] Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and Jon Bird. 2015. Activity tracking: Barriers, workarounds and customisation. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, 617–621. DOI: <https://doi.org/10.1145/2750858.2805832>
- [29] Michael Hemphill. 1996. A note on adults' color-emotion associations. *The Journal of Genetic Psychology* 157, 3 (1996), 275–280. DOI: <https://doi.org/10.1080/00221325.1996.9914865>
- [30] Victoria Hollis, Artie Konrad, Aaron Springer, Matthew Antoun, Christopher Antoun, Rob Martin, and Steve Whittaker. 2017. What does all this data mean for my future mood? Actionable analytics and targeted reflection for emotional well-being. *Human-Computer Interaction* 32, 5–6 (2017), 208–267. DOI: <https://doi.org/10.1080/07370024.2016.1277724>
- [31] Noura Howell, Laura Devendorf, Rundong (Kevin) Tian, Tomás Vega Galvez, Nan-Wei Gong, Ivan Poupyrev, Eric Paulos, and Kimiko Ryokai. 2016. Biosignals as social cues: Ambiguity and emotional interpretation in social displays of skin conductance. In *Proceedings of the ACM Conference on Designing Interactive Systems (DIS '16)*. Association for Computing Machinery, New York, NY, 865–870. DOI: <https://doi.org/10.1145/2901790.2901850>
- [32] Noura Howell, Laura Devendorf, Tomás A. Vega Gálvez, Rundong Tian, and Kimiko Ryokai. 2018. Tensions of data-driven reflection: A case study of real-time emotional biosensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, 1–13. DOI: <https://doi.org/10.1145/3173574.3174005>
- [33] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. 2003. Technology probes: Inspiring design for and with families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. Association for Computing Machinery, New York, NY, 17–24. DOI: <https://doi.org/10.1145/642611.642616>
- [34] Ellen Isaacs, Artie Konrad, Alan Waiendowski, Thomas Lennig, Victoria Hollis, and Steve Whittaker. 2013. Echoes from the past: How technology mediated reflection improves well-being. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, 1071–1080. DOI: <https://doi.org/10.1145/2470654.2466137>
- [35] Naz Kaya and Helen H. Epps. 2004. Relationship between color and emotion: A study of college students. *College Student Journal* 38, 3 (2004), 396–405.



- [36] Rohit Ashok Khot, Deepti Aggarwal, Ryan Pennings, Larissa Hjorth, and Florian 'F.' Mueller. 2017. EdiPulse: Investigating a playful approach to self-monitoring through 3D printed chocolate treats. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, 6593–6607. DOI: <https://doi.org/10.1145/3025453.3025980>
- [37] Dohee Kim, Sangsu Jang, Beom Kim, and Young-Woo Park. 2022. Design and field trial of lumino in homes: Supporting reflective life by archiving and showing daily moods with light colors. In *Proceedings of the ACM Designing Interactive Systems Conference (DIS '22)*. Association for Computing Machinery, New York, NY, 1715–1728. DOI: <https://doi.org/10.1145/3532106.3533465>
- [38] Nam W. Kim, Hyejin Im, Nathalie Henry Riche, Alicia Wang, Krzysztof Gajos, and Hanspeter Pfister. 2019. DataSelfie: Empowering people to design personalized visuals to represent their data. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, 1–12. DOI: <https://doi.org/10.1145/3290605.3300309>
- [39] Young-Ho Kim, Jae H. Jeon, Bongshin Lee, Eun K. Choe, and Jinwook Seo. 2017. OmniTrack: A Flexible Self-Tracking Approach Leveraging Semi-Automated Tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (Sep. 2017), Article 67, 28 pages. DOI: <https://doi.org/10.1145/3130930>
- [40] Amanda Lazar, Christian Koehler, Theresa J. Tanenbaum, and David H. Nguyen. 2015. Why we use and abandon smart devices. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, 635–646. <https://doi.org/10.1145/2750858.2804288>
- [41] Kwangyoung Lee and Hwajung Hong. 2017. Designing for self-tracking of emotion and experience with tangible modality. In *Proceedings of the Conference on Designing Interactive Systems (DIS '17)*. Association for Computing Machinery, New York, NY, 465–475. DOI: <https://doi.org/10.1145/3064663.3064697>
- [42] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, New York, NY, 557–566. DOI: <https://doi.org/10.1145/1753326.1753409>
- [43] Ian Li, Jodi Forlizzi, and Anind Dey. 2010. Know thyself: Monitoring and reflecting on facets of one's life. In *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. Association for Computing Machinery, New York, NY, 4489–4492. DOI: <https://doi.org/10.1145/1753846.1754181>
- [44] Francisco Nunes, Nervo Verdezoto, Geraldine Fitzpatrick, Morten Kyng, Erik Grönvall, and Cristiano Storni. 2015. Self-care technologies in HCI: Trends, tensions, and opportunities. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 6 (Dec. 2015), Article 33, 45 pages. DOI: <https://doi.org/10.1145/2803173>
- [45] Kyrill Potapov and Paul Marshall. 2020. LifeMosaic: Co-design of a personal informatics tool for youth. In *Proceedings of the Interaction Design and Children Conference (IDC '20)*. Association for Computing Machinery, New York, NY, 519–531. DOI: <https://doi.org/10.1145/3392063.3394429>
- [46] Blaine A. Price, Ryan Kelly, Vikram Mehta, Ciaran McCormick, Hanad Ahmed, and Oliver Pearce. 2018. Feel my pain: Design and evaluation of painpad, a tangible device for supporting inpatient self-logging of pain. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, 1–13. DOI: <https://doi.org/10.1145/3173574.3173743>
- [47] Amon Rapp and Federica Cena. 2016. Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. *International Journal of Human-Computer Studies* 94 (2016), 1–17. DOI: <https://doi.org/10.1016/j.ijhcs.2016.05.006>
- [48] Amon Rapp, Alessandro Marcengo, Luca Buriano, Giancarlo Ruffo, Mirko Lai, and Federica Cena. 2018. Designing a personal informatics system for users without experience in self-tracking: A case study. *Behaviour & Information Technology* 37, 4 (2018), 335–366. DOI: <https://doi.org/10.1080/0144929X.2018.1436592>
- [49] Amon Rapp and Lia Tirabeni. 2018. Personal informatics for sport: Meaning, body, and social relations in amateur and elite athletes. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 3 (Jun. 2018), Article 16, 30 pages. DOI: <https://doi.org/10.1145/3196829>
- [50] Amon Rapp and Lia Tirabeni. 2020. Self-tracking while doing sport: Comfort, motivation, attention and lifestyle of athletes using personal informatics tools. *International Journal of Human-Computer Studies* 140 (2020), 102434. DOI: <https://doi.org/10.1016/j.ijhcs.2020.102434>
- [51] Amon Rapp and Maurizio Tirassa. 2017. Know thyself: A theory of the self for personal informatics. *Human-Computer Interaction* 32, 5–6 (2017), 335–380. DOI: <https://doi.org/10.1080/07370024.2017.1285704>
- [52] Yvonne Rogers, William R. Hazlewood, Paul Marshall, Nick Dalton, and Susanna Hertrich. 2010. Ambient influence: Can twinkly lights lure and abstract representations trigger behavioral change? In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing (UbiComp '10)*. Association for Computing Machinery, New York, NY, 261–270. DOI: <https://doi.org/10.1145/1864349.1864372>
- [53] Pedro Sanches, Kristina Höök, Corina Sas, and Anna Ståhl. 2019. Ambiguity as a resource to inform proto-practices: The case of skin conductance. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 4 (Jul. 2019), Article 21, 32 pages. DOI: <https://doi.org/10.1145/3318143>

- [54] Donald Schon. 1983. *The Reflective Practitioner. How Professionals Think in Action*. Temple Smith, London.
- [55] Jessica Schroeder, Chia-Fang Chung, Daniel A. Epstein, Ravi Karkar, Adele Parsons, Natalia Murinova, James Fogarty, and Sean A. Munson. 2018. Examining self-tracking by people with migraine: Goals, needs, and opportunities in a chronic health condition. In *Proceedings of the Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, 135–148. DOI: <https://doi.org/10.1145/3196709.3196738>
- [56] Petr Slovák, Christopher Frauenberger, and Geraldine Fitzpatrick. 2017. Reflective practicum: A framework of sensitising concepts to design for transformative reflection. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, 2696–2707. DOI: <https://doi.org/10.1145/3025453.3025516>
- [57] Yulia Stepanchuk. 2017. Quantified construction of self: Numbers, narratives and the modern individual. In *Proceedings of the International Conference IMS-2017 (IMS '17)*. Association for Computing Machinery, New York, NY, 28–36. DOI: <https://doi.org/10.1145/3143699.3143720>
- [58] Cristiano Storni. 2014. Design challenges for ubiquitous and personal computing in chronic disease care and patient empowerment: A case study rethinking diabetes self-monitoring. *Personal and Ubiquitous Computing* 18 (2014), 1277–1290. DOI: <https://doi.org/10.1007/s00779-013-0707-6>
- [59] Alice Thudt, Uta Hinrichs, Samuel Huron, and Sheelagh Carpendale. 2018. Self-reflection and personal physicalization construction. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, 1–13. DOI: <https://doi.org/10.1145/3173574.3173728>
- [60] Niels van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The experience sampling method on mobile devices. *ACM Computing Surveys (CSUR)* 50, 6 (Dec. 2017), Article 93, 40 pages. DOI: <https://doi.org/10.1145/3123988>
- [61] Johann W. von Goethe and Charles L. Eastlake. 1840. *Goethe's Theory of Colours*. J. Murray. Retrieved from <https://books.google.co.uk/books?id=OIoIAAAAMAAJ>
- [62] Mark Whooley, Bernd Ploderer, and Kathleen Gray. 2014. On the integration of self-tracking data amongst quantified self members. In *Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014 - Sand, Sea and Sky - Holiday HCI (BCS-HCI '14)*. BCS, Swindon, GBR, 151–160. DOI: <https://doi.org/10.14236/ewic/hci2014.16>

Received 12 December 2022; revised 15 February 2024; accepted 27 February 2024