Non-judgmental Interfaces: A New Design Space for Personal Informatics

Romain Toebosch
romain.toebosch@uni.lu
University of Luxembourg
Esch-sur-Alzette, Luxembourg

Arne Berger
arne.berger@hs-anhalt.de
Anhalt University of Applied Sciences
Koethen, Germany

Carine Lallemand
carine.lallemand@uni.lu
University of Luxembourg
Esch-sur-Alzette, Luxembourg

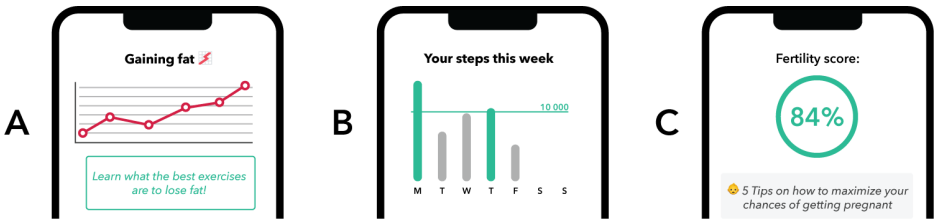

Figure 1: Examples of Personal Informatics (PI) interfaces judging the user or pushing toward normative goals: (A) a smart scale judging the user for gaining weight, (B) a fitness tracker assessing 10,027 steps a day as more desirable than 9,987 steps, (C) a menstrual cycle tracker displaying a "fertility score" and giving tips on how to get pregnant, judging pregnancy as the desired outcome.

ABSTRACT

Personal Informatics (PI) systems like self-trackers implicitly or explicitly judge user behaviour based on the data they collect. In a behaviour change context, this may help individuals recognise corrective actions to take. However, these value judgments are unnecessary or harmful in some situations, for instance when (a) the user does not use the self-tracker with an intent to change, (b) the user needs to reflect more deeply on their data to achieve their desired change, or (c) the user's behaviour or intended use case is non-normative. In this provocation, we aim to challenge the normative status quo by offering reflections on alternative, "non-judgmental" ways to represent self-tracking systems interfaces. We present three design concepts to support these reflections. We discuss what makes PI systems judgmental by detangling the elements embedding judgment, and the relations between judgmental and normative. Is it even possible to remove all value judgments from PI systems after all?

CCS CONCEPTS

• **Human-centered computing → Interaction design theory, concepts and paradigms**.

KEYWORDS

personal informatics, self-tracking, normativity, non-judgmental interfaces, interface design

DIS Companion '24, July 01–05, 2024, IT University of Copenhagen, Denmark

ACM ISBN 979-8-4007-0632-5/24/07
https://doi.org/10.1145/3656156.3663706

ACM Reference Format:
Romain Toebosch, Arne Berger, and Carine Lallemand. 2024. Non-judgmental Interfaces: A New Design Space for Personal Informatics. In *Designing Interactive Systems Conference (DIS Companion '24), July 01–05, 2024, IT University of Copenhagen, Denmark.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3656156.3663706

1 INTRODUCTION

What is the difference between a smart scale displaying *"You are severely overweight"* or *"Good job, you lost 1.3 kg over the last month"*? In this provocation, we argue there is none.

Personal Informatics (PI) systems - defined as systems that "*help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge*" [9, p. 1] - are ubiquitous nowadays. Mobile apps, wearables, and Internet of Things (IoT) devices enable users to track various aspects of their lives. In Human-Computer Interaction (HCI), publications on PI have risen steadily over the past decade [5]. Most commonly investigated are PI systems to support behaviour change [5], and many commercial devices are also marketed for this purpose. In an attempt to make collected data easy to reflect on for users, these systems interpret the data and use messages, scores, graphs, or icons to recommend what to change on the user's behalf — e.g., setting more steps or adhering to a sleep routine. While giving such recommendations, PI systems are, implicitly or explicitly, judging users' behaviors to prescribe to them what to do. Figure 1.A conveys that losing 0.9 kg is more desirable than gaining 0.7 kg. In Figure 1.B, setting 10,027 steps a day would be "good", whereas 9,987 steps are "not enough". The claimed benefit is that users do not need to think for themselves, but understand at a glance what corrective actions to take to reach set goals. Yet this is more problematic when, for instance, the user 1) does not use a self-tracker with an intent to change, 2) needs to reflect more deeply on their data, or 3) their behavior

or intended use is non-normative. In the following sections, we elaborate on these scenarios.

1.1 Assumed Desire for Change

We argue that the diversity of users and their goals warrant more nuanced approaches to personal informatics. Many users come in contact with PI systems by chance: phones include pre-installed health or screen time trackers, smartwatches track bio-metrics, and music streaming services track listening behavior. Behaviour change suggestions prevail, while users may not have had specific goals in mind when picking up these technologies. There are also many "imposed" instances of self-tracking [10], e.g., at the workplace. This again challenges the assumption that users selected a self-tracker with intention to change their behaviour.

Even if users deliberately and voluntarily engage with a self-tracker, the diversity in tracking motivations is rarely reflected in the designs of these systems. Of the five different profiles of self-trackers [11] — directive, documentary, diagnostic, collecting rewards, and fetishized — only one motivation fits behavior change. As Spiel et al. [12, p. 5] state, fitness trackers assume "everybody requires improvement," reflecting this implicit assumption that every user engaging in fitness self-tracking should aim to change their behaviours and/or body and that "the self – as it is – is somehow never adequate or enough" [14]. Furthermore, "failure to meet the tracker's terms is the individuals' problem" [12, p. 5], meaning users are also being judged for abstaining from pursuing these goals.

1.2 Lack of Support for Reflection

PI models link PI to self-reflection and the pursuit of gaining self-knowledge [6, 9]. A 2022 review of 123 commercial tracking apps [2] shows that most PI applications do not go past the lowest levels of reflection as described in Fleck and Fitzpatrick [7]'s model. Their findings suggest that many PI apps are just another form of persuasive technology, constraining the meaning-making process of users and ultimately limiting their ability to acquire self-insight [2]. Furthermore, PI design can potentially lead to rumination instead of reflection, to which the representation of goal-discrepancy — the difference between the users actual and desired state — can play a role [4]. Therefore, even if users engage with self-trackers seeking change, they may not always manage to do so through reflection due to judgmental elements in PI interfaces.

1.3 Inflexibility towards Non-Normative Cases

Assuming the user deliberately engages with a self-tracker to change their behavior, they might implicitly still be judged if they do not follow a normative lifestyle or pursue normative goals. As Spiel et al. [12, p. 1] wrote, self-trackers "merely affirm the fitness of already fit people and can have an adversarial effect on others." They argue that by measuring steps, trackers redefine what a step is for the user based on normative assumptions. Van Dijk et al. [13] also discuss the problem of standardized feedback, which forgoes the personal circumstances from an individual user. They suggest that the value interpretation of a result as being "high" or "low" in a standardized way can potentially mislead users.

In these three cases, PI systems fail to provide users with what they need: instead of helping users gain insights and reflect on their data, a judgment is made and a corrective behavior is suggested to step closer to the predefined, normative goal embedded in the system. But is it possible to design usable PI systems without **any** judgment or normative goals?

Before this background we ask: How can we support reflection in non-judgmental yet system-driven ways? In this provocation, we explore system-driven ways to extract insights from PI data without judgment or normative goals. We present three "*non-judgmental interfaces*" as redesigns of self-tracking apps. Each embodies a different strategy to avoid judgment and normative goals in PI systems. We discuss where judgment lies in interfaces, and whether it is possible at all to make completely Non-judgmental Interfaces

2 DESIGN

Reflecting on prior work in self-tracking, we wanted to avoid making a "data-dump" of seemingly non-judgmental "raw" data. Even if data is only minimally processed, "innocent" or non-judgmental sensors do not exist. Additionally, raw visualizations do not help users gain meaningful insights. Instead, our conceptualization of non-judgmental interfaces, where no outcome is valued over another, relies on (a) an analysis of existing PI systems, reflecting on how these can avoid embedding judgmental (and to some extent normative) views, (b) a research-through-design approach, exploring how different algorithms can extract features from raw data, or how the user interface can convey this data in non-judgmental ways.

2.1 (Un)assuming Menstrual Cycle Tracker

Many menstrual cycle trackers using basal body temperature embed the assumed goal of getting pregnant or not. Although opposite, both are focused on the same phenomenon, for which some use "fertility scores." But what if a user has different aims, e.g., adapting their sports program to their cycle? Can a menstrual cycle tracker be designed without predefined goals in mind? Figure 2 shows how such a tracker could look like. By letting users define what each time period in their cycle corresponds to, it supports a wide range of scenarios, including less and non-normative ones. Most of the processing beneath the interface remained unchanged: basal body temperature is still used to determine where in the cycle the user is. However, by letting users assign labels to each part of their cycle, we become indifferent toward the goal and allow for non-normative goals and uses of the application.

2.2 Curious Calendar

"The more activity, the better" is the rationale behind most activity trackers, which prescribe goals and display scores on various metrics. The Curious Calendar uses unsupervised clustering to detect patterns in activity data and reports these to the user in a calendar view (Figure 3). If the tracked data of two time periods share similarities, the algorithm clusters and visualizes them similarly, prompting the user to reflect on them (Figure 3B). Users may add descriptions of what they were doing, and review what they reported during similarly clustered moments, allowing them to understand patterns of activity on their own terms (Figure 3C). This opens up possibilities to track for motivations beyond normative behaviour

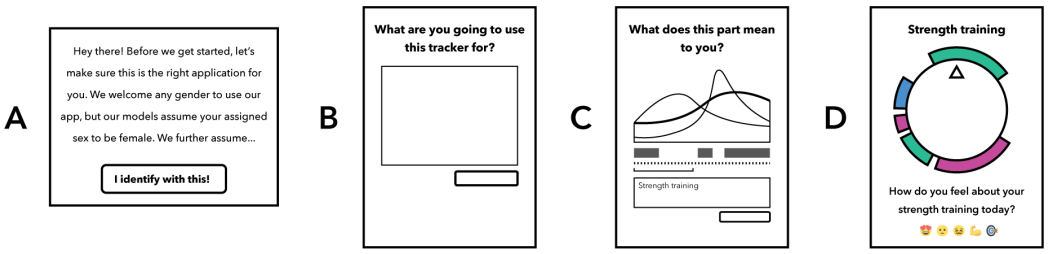

Figure 2: An Unassuming Menstrual Cycle Tracker. A welcome message states the underlying assumptions of the processing in the app (A) and allows the user to describe their motivation to use the tracker (B). The user defines what a specific part of their menstrual cycle means to them, e.g., "strength training" (C). The app visualizes where the user is in their cycle using their self-assigned label (D). Beneath, the user can report their feelings using an emoji, and see an overview of how they have felt previously at this stage (E)

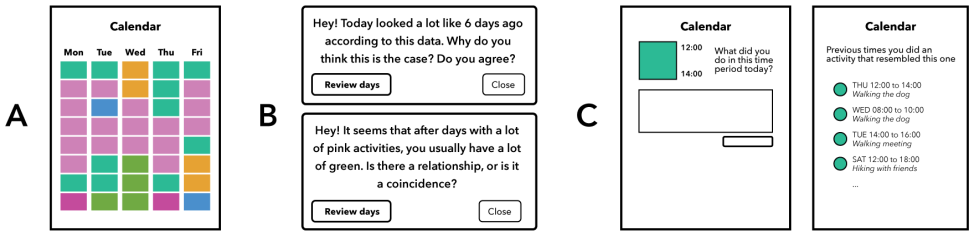

Figure 3: A Curious Calendar. The Calendar visualizes the user's day into 2h blocks (A). The cell colors are based on the accelerometer data clustering from the activity band (similar activities having the same color). Popup messages prompt the user to reflect on relationships in their data (B). They can annotate a specific period and compare it with the descriptions of similarly clustered periods.

change, such as documentary and fetishised tracking [11] and allows for individual and situated interpretation [8] of tracking data. It also supports tracking of non-normative behaviours: as clustering is based on similarity in data (and does not rely on classification models trained with normative assumptions), any pattern in the user's data should allow for classification.

2.3 A Glimpse into the Future

Most sleep trackers calculate "sleep scores" to compare with other users and your past self. Instead of using other users as predictors for "individual sleep performance", what if users are supported to explore their future? A Glimpse into the Future allows users to track their sleep based on self-selected, self-reported variables, e.g., stress, energy, or body feel (Figure 4A). Before sleeping, the user can use these to predict a self-assigned variable the next morning, based on parameters they fill in now (Figure 4B). In this prediction, the user can choose which variables they think are important to their prediction. Instead of focusing on optimizing sleep only, such a tracker could be used e.g., for diagnostic tracking [11], or goal-oriented tracking beyond behavior change (e.g., I want to predict how stiff my back will feel tomorrow to decide whether to book a sports session in the morning or not). In some ways, this approach shares similarities with self-experimentation, where users aim to find knowledge that is individually meaningful to them [3].

3 DISCUSSION

In this provocation, we argued why judgment in PI systems can be problematic and exemplified a "non-judgmental" approach. We use these examples to discuss the nature of judgment in PI systems, take a critical look at whether it is possible at all to avoid judgment, and discuss how non-judgmental interfaces can be used by researchers and designers.

3.1 The Loci of Judgment in PI Systems

The design process for each of the applications shown in this provocation was at times similar to peeling away the layers of an onion. For the (Un)assuming Menstrual Cycle Tracker, removing the first goal-oriented layer of feature extraction (fertility score) and labeling (pregnancy) found in many of its commercial counterparts, seemed to already change it for the better. For the Curious Calendar we stripped away some more layers, and replaced the common feature extraction, usually an algorithm extracting normative metrics from the data such as steps, with an unsupervised clustering method. Lastly, for Glimpse into the Future we went one step further and allowed users to choose what data streams were taken into account and to experiment with these.

Through the design of these concepts, we therefore identified multiple places in PI systems where judgment can occur (Figure 5). These layers, especially the feature extraction, resonate with the levels of abstraction in derived metrics by Bentvelzen et al. [1]. Normative assumptions and value judgments can be present in each

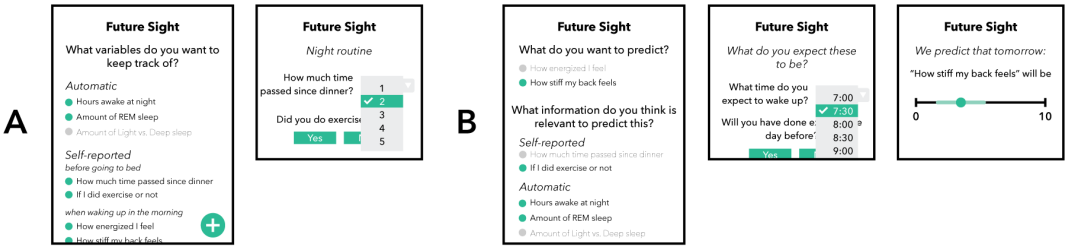

Figure 4: A Glimpse into the Future. The user selects what variable(s) to self-report in addition to sleep tracking (A). They select a variable to predict, choose which input variables are taken into account, and provide the system with an estimate of what values they will likely have. Based on these parameters, the system predicts what the variable will be tomorrow (B)

of these layers. Sometimes, a simple change in labels and interface design makes a significant difference — such as in the (Un)assuming Menstrual Cycle Tracker — but other times judgment is inevitably embedded in the way data is collected and/or processed. PI systems designers should be aware of these layers to make informed design choices.

A commonality between our three concepts is that they rely as much as possible on individual data, rather than on comparisons and models of "general" or "desired" user profiles. Especially in the case of the unsupervised clustering of the Curious Calendar, where there is not even any trained model based on data from other users. We aim for designers to consider whether involving data from other users is necessary or not, because they might embed norms.

3.2 Are Non-judgmental Interfaces in PI even Possible?

3.2.1 Is non-judgmental always non-normative and vice-versa? One challenge in the conceptualization of non-judgemental interfaces relates to the relation between judgment and normativity. Judgmentality arises when one outcome is valued over another (e.g., losing weight is better than gaining weight). Norms concern the social acceptability of behaviours, and implementing these can therefore also lead to judgment. Although it is important to allow for non-normative use scenarios when designing non-judgmental interfaces, the opposite does not necessarily hold. A self-tracking app aimed at increasing one's body fat would be a rather non-normative use case. Yet users would be judged when losing body fat, making this app non-normative yet judgmental.

Figure 5: The different loci of judgment in PI systems. While the user interface (UI) is the most visible layer where judgment happens (including how metrics are labeled, what data visualization or messages are conveyed to users, and how), judgment can occur as soon as the selection of sensors or their calibration to measure a specific phenomenon. Lower level of features extraction (e.g., making "steps" out of accelerometer data) and higher level ones (computing a score) are equally strategic to consider.

In the case of a smart scale, even if the system is not judging its users for gaining fat, societal norms could still deem this undesirable, and the users would probably judge themselves according to those norms. Is the pursuit of creating non-judgmental interfaces futile in the context of judgmental societies that over-accentuate normativity? Isn't there always someone judging (even through internalized norms only)? While our brief examples of non-judgmental interfaces do not address this issue, we argue the presented designs leave more room for cases where users actively disagree with the norms in question. Further work could look into self-tracking designs that help users actively challenge norms, but this does leave us with the question whether value judgement can be removed from PI systems at all.

3.2.2 There is place for judgment. We do not deny that the more common, goal-oriented systems also have their place in the landscape of PI tools. They fulfill the motivation of specific users and we would not practice what we preach if we suggest to rule out a specific design strategy. Admittedly, defined use cases and promises of change are also powerful selling points for self-trackers, and it is no surprise that the industry is marketing these systems as such. Yet, we call for more diverse design approaches, expanding the use scenarios beyond the judgmental or the normative. Could users decide what is being judged, or customize how? Can we sensitize or empower users? Can the systems be more transparent, e.g., on the derived metrics used? Could they perhaps disclose the positionality and values they embed?

4 CONCLUSION

In this paper, we argued for the need for non-judgmental systems and interfaces in the PI domain. We reflected on three concepts of non-judgmental interfaces, employing different techniques to avoid judging its users. We call on the design community to challenge the status quo and explore the opportunities in the suggested design space opened around "non-judgemental interfaces".

ACKNOWLEDGMENTS

This research has been supported by the Luxembourg National Research Fund (FNR) IPBG2020/IS/14839977/C21. The authors would like to warmly thank Pepijn Verburg, Lorena Sánchez Chamorro, and the Industrial Design students at the Eindhoven University of Technology for their reflections on this topic.

REFERENCES

[1] Marit Bentvelzen, Jasmin Niess, and Paweł W. Woźniak. 2023. Designing Reflective Derived Metrics for Fitness Trackers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (Jan. 2023), 158:1–158:19. https://doi.org/10.1145/3569475

[2] Janghee Cho, Tian Xu, Abigail Zimmermann-Niefield, and Stephen Voida. 2022. Reflection in Theory and Reflection in Practice: An Exploration of the Gaps in Reflection Support among Personal Informatics Apps. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–23. https://doi.org/10.1145/3491102.3501991

[3] Nediyana Daskalova, Eindra Kyi, Kevin Ouyang, Arthur Borem, Sally Chen, Sung Hyun Park, Nicole Nugent, and Jeff Huang. 2021. Self-e: Smartphone-supported guidance for customizable self-experimentation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[4] Elizabeth Victoria Eikey, Clara Marques Caldeira, Mayara Costa Figueiredo, Yunan Chen, Jessica L. Borelli, Melissa Mazmanian, and Kai Zheng. 2021. Beyond self-reflection: introducing the concept of rumination in personal informatics. *Personal and Ubiquitous Computing* 25, 3 (June 2021), 601–616. https://doi.org/10.1007/s00779-021-01573-w

[5] Daniel A. Epstein, Clara Caldeira, Mayara Costa Figueiredo, Xi Lu, Lucas M. Silva, Lucretia Williams, Jong Ho Lee, Qingyang Li, Simran Ahuja, Qiuer Chen, Payam Dowlatyari, Craig Hilby, Sazeda Sultana, Elizabeth V. Eikey, and Yunan Chen. 2020. Mapping and Taking Stock of the Personal Informatics Literature. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (Dec. 2020), 1–38. https://doi.org/10.1145/3432231

[6] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 731–742. https://doi.org/10.1145/2750858.2804250

[7] 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, USA, 216–223. https://doi.org/10.1145/1952222.1952269

[8] Albrecht Kurze, Andreas Bischof, Sören Totzauer, Michael Storz, Maximilian Eibl, Margot Brereton, and Arne Berger. 2020. Guess the Data: Data Work to Understand How People Make Sense of and Use Simple Sensor Data from Homes. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–12. https://doi.org/10.1145/3313831.3376273

[9] 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, USA, 557–566. https://doi.org/10.1145/1753326.1753409

[10] Deborah Lupton. 2016. The diverse domains of quantified selves: self-tracking modes and dataveillance. *Economy and Society* 45, 1 (Jan. 2016), 101–122. https://doi.org/10.1080/03085147.2016.1143726 Publisher: Routledge _eprint: https://doi.org/10.1080/03085147.2016.1143726.

[11] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal tracking as lived informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 1163–1172. https://doi.org/10.1145/2556288.2557039

[12] Katta Spiel, Fares Kayali, Louise Horvath, Michael Penkler, Sabine Harrer, Miguel Sicart, and Jessica Hammer. 2018. Fitter, Happier, More Productive?: The Normative Ontology of Fitness Trackers. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–10. https://doi.org/10.1145/3170427.3188401

[13] Elisabeth Kersten Van Dijk, Wijnand IJsselsteijn, and Joyce Westerink. 2016. Deceptive visualizations and user bias: a case for personalization and ambiguity in PI visualizations. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, Heidelberg Germany, 588–593. https://doi.org/10.1145/2968219.2968326

[14] Emily Winter, Bran Knowles, Daniel Richards, Kim Snooks, and Chris Speed. 2022. Multitudes: Widening the research agenda for personal informatics design. In *DRS2022: Bilbao*. Bilbao, Spain. https://doi.org/10.21606/drs.2022.415