# Sensemaking Challenges in Personal Informatics and Self-Monitoring Systems

**Simon L. Jones**
University of Bath
Bath, BA2 7AY, UK
s.l.jones@bath.ac.uk

**Ryan Kelly**
University of Bath
Bath, BA2 7AY, UK
r.m.kelly@bath.ac.uk

## Abstract
Personal informatics (PI) systems, which aggregate and analyse personal data from activity tracking devices and lifelogging services, have been shown to provide benefits in health and wellbeing settings. In this workshop paper we report a preliminary analysis of interviews with users of a personal informatics system and discuss the challenges that these users encounter in making sense of their data. We identify four challenges that may have implications for the use of PI systems in a health context, which we propose to discuss at the WISH workshop with other researchers who have considered self-monitoring from health and sensemaking perspectives.

## Author Keywords
Quantified Self; Personal Informatics; Interactive Health Systems; Sensemaking.

## ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

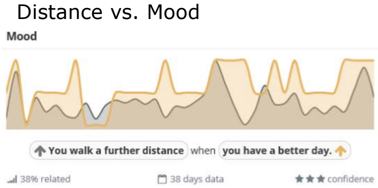
## Introduction
Health is a multi-faceted phenomenon, affected by myriad factors such as diet, exercise, environment, stress, and daily routine. A growing number of

***Sidebar 1***

**1.1 Data Collected:**
18 participants (9 males, 9 females) provided Exist with data that included daily measurements of: physical activity and sleep (both recorded by a wearable Fitbit sensor); productivity and distracting time (recorded by RescueTime logging software); mood (self-reported likert-scale scores by daily emails); events (automatically retrieved from online calendars); social media interactions (from Twitter and Instagram); music listening (recorded by Last.fm 'scrobbling' from music players such as Spotify and iTunes); and local weather conditions (from Forecast.io).

**1.2 Participant Information:**
N = 18, Mean age = 28.3 years, Age range = 21-60 years, Gender: Male = 9, Female = 9, Previous tracking experience: Yes = 6, No = 12, Level of education (UK): Secondary School = 2, College = 3, Bachelor = 4, Master = 6, Doctorate = 3

**1.3 Example Exist Correlation:**
Distance vs. Mood




“You walk a further distance when you have a better day”

interactive systems, which aim to support independent health monitoring and management, are attending to this multifaceted configuration by providing users with the ability to aggregate and analyse data from many aspects of their lives. The Health Mashups system [1], for example, captures diverse data about users and reveals statistical associations between health-related metrics (e.g. quality of sleep, levels of pain) and other contextual factors. These statistical patterns are presented to users in the form of natural language statements, such as “You experience less pain on days when you get more exercise”.

A growing number of mainstream personal informatics tools are adopting a similar multifaceted approach to personal tracking and attempt to present insights to users in an understandable way. Whilst it has been shown that these multifaceted systems provide users with insights that they could not easily derive themselves [1,2], few studies have explored the challenges that users face in *making sense* of the information provided by this type of system.

Broadly, sensemaking pertains to finding meaning in a situation [6]. In HCI, it refers to the cognitive act of understanding information [8]. Supporting sensemaking is a known challenge in HCI, especially when users are confronted with the task of interpreting complex information spaces. For example, Kelly & Payne found that users struggled to make sense of search returns due to the sheer volume of pages gathered and the presence of large amounts of irrelevant content [3]. In a qualitative study of various information visualisations, Lee et al. [4] found that users sometimes ‘floundered’ when trying to make sense of visual representations that were unfamiliar.

We believe that effective self-management of health-related conditions requires users to be able to make sense of their data. Mamykina et al. [5] identify three essential sensemaking activities for health self-management: 1) perception of new information related to health, 2) development of inferences that inform selection of actions, and 3) carrying out actions in response to new information. Problems with these activities are likely to result in users either missing out on the benefits provided by a system or, worse still, engaging in courses of inappropriate or harmful action. Yet little is known about sensemaking in the context of multifaceted personal informatics systems, whether it is problematic, and how these problems could be resolved.

In this paper we highlight some of the initial sensemaking challenges that were uncovered from preliminary analysis of interviews with 18 users of *Exist*, a multi-faceted personal informatics system. Exist bears a close resemblance to the Health Mashups system [1]. It aggregates data from numerous distinct self-tracking services and discovers statistical correlations present within the data. The service presents correlational information to its users as graphical visualisations and natural language statements (see Sidebar 1.3), e.g. ‘You sleep better on days when you are more physically active’, or ‘You have a better mood when you listen to more classical music’. While Exist does not focus explicitly on health-related data, it has the potential to do so. It exemplifies a growing number of personal informatics systems that process diverse personal data. Therefore, we contend that the challenges we uncover are relevant to health informatics systems of this kind, and are unlikely to be unique to Exist.

***Sidebar 2**:*

*Interview Quotes 1-9*

*Q1) There's so many. Yay! That's really cool. [P9]*

*Q2) I've been really excited to find out how much it has to display about me. [P13]*

*Q3) You've really got to have your brain in gear to go through these! I'm not used to this sort of thing. [P17]*

*Q4) ...Is that right? Oh I can't figure it out, I'm getting analysis paralysis. I've got it big time. I've seen so many… I'm struggling to figure out what these things mean now. [P1]*

*Q5) There's so many it's hard to reflect on what's missing. The only one I noticed was one I was particularly interested in, which is music and mood. [P1]*

*Q6) I would like to select which ones I find most useful, and take away some of the repetition stuff. That would be nice. [P9]*

*Q7) Cloud, wind and rain are not really of interest to me. It goes into more depth than is needed. I would say rain, whether it's dry, is the only one I would need. [P3]*

*Q8) There's quite a lot, I'd take out all the "it's rainier when it's cloudier", "you sleep more when you're in bed more"...all of that obvious stuff... Too much to deal with. [P4]*

*Q9) Obvious. That's just science! It just is! [P2]*

**Method**

18 participants were recruited via advertisements on University noticeboards. Each participant provided data to the Exist system for 1-3 months (information about participants and the data collected is shown in Sidebar 1). At the end of the data collection period we conducted face-to-face semi-structured interviews with all of the participants, lasting 45–111 minutes (M=62). In this session, participants were shown printed screenshots of the correlations revealed by the Exist system and were asked to think aloud whilst reviewing the output. The outputs contained a mean average of 76 reported correlations per participant (range=24-109). The interviewer probed for deeper explanations of the output until no new information seemed to come from participants' responses. Interview transcripts were analysed inductively, using phases of open coding to identify concepts within the data, and axial coding to identifying relationships among the concepts. Four themes associated with sensemaking that emerged from our preliminary analysis are discussed in this paper. We discuss their relevance to health-related settings at the end of the paper.

**Sensemaking Challenge 1: Quantity and unfamiliarity of information can overwhelm users**

Our analysis revealed that some participants initially considered the presentation of numerous correlations to be a positive characteristic of the system (see Sidebar 2; Quotes 1 & 2). These participants valued the prospect of having many outputs to explore, seemingly because they believed that this would correspond to the insights they could derive from the system. However, it was apparent that for at least five of the participants, the initial satisfaction of receiving many correlations gave way to frustration with regards to the cognitive effort required to review all of the outputs (e.g. Quotes 3, 4, 5). Some participants expressed difficulty in accurately comprehending the information that they were presented with, due to the sheer volume of analysis overwhelming the sense-making process itself. One participant referred to his feeling of being overwhelmed and unable to form rational conclusions as "analysis paralysis" (Quote 4). Difficulties interpreting the outputs of the system were compounded by the unfamiliarity of correlational information of this kind (e.g. Quote 3). Thus there is a challenge associated with condensing or filtering the information that is shown to the user, and presenting this information in an intuitive and familiar way.

Many participants felt that one approach to mitigate the overload of information would be to remove insights that appeared in duplicate (e.g. Quote 6). These duplicates appeared in part due to the design decision of Exist to present correlations between variables A and B twice, once showing that A correlates with B, and again showing that B correlates with A. In addition, some insights were deemed to be duplicates of one another because of the level of granularity at which they were interpreted, and because they co-varied with other variables to some extent. For example, variables capturing cloud cover, wind speed and precipitation levels were simply viewed as redundant and overly detailed measures of good or bad weather by several participants (e.g. Quote 7). For health related variables these redundancy issues might arise where similar measures are being recorded (e.g. blood glucose levels and urine glucose levels).

An additional source of frustration for participants was the presence of many insights that were considered "obvious" (e.g. Quotes 8,9,10). These correlations offered little value and added to the feeling of being

***Sidebar 3**:*

*Interview Quotes 10-14*

*Q10) There are some silly things, like you're more active when you have more steps. That's kinda obvious. [P10]*

*Q11) These (statements) only really show the effects of 'more of this', or 'less of that'. For some things that's fine...like, a better mood is always a good thing, and a worse mood is always a bad thing, but for something other than mood there might be a sort of sweet spot. I guess, sleep for example... too little is bad, but too much is bad too … I kinda want to know what leads to just the right amount of sleep, not just more or less. Yeah, these sentences don't really do that. [P5]*

*Q12) You're more productive when you log more distracting time. That's a strange correlation. Really. So, it's telling me I should really spend more time being distracted, it'll help me get more done. What's it saying? Has something gone wrong there? [P8]*

*Q13) Quite interesting that I'm more productive when I listen to more music. I always thought that I worked best without music. [P8]*

*Q14) You're more productive when you spend more time in bed (laughs). I'm laughing because that statement seems odd… [P1]*

overloaded with information. Bentley et al. reported similarly negative reactions to "obvious" insights in Health Mashups [1], noting that their presence did not make sense to some users and created a tension between telling people what they already know and educating them with new insights.

**Sensemaking Challenge 2: Poor information presentation can lead to misinterpretation**

A second sensemaking challenge related to the misinterpretation of information. While systems such as Exist and Health Mashups aim to improve the understandability of insights by presenting them as natural language statements, we found numerous examples of the phraseology of these statements (see Sidebar 1.3) leading to misinterpretation of information and over-simplification of complex relationships (e.g. Quotes 11,12). This meant that users of the system developed flawed inferences. For example, a statement revealing an inverse correlation between amount of exercise and levels of pain, e.g. 'You have less pain when you do more exercise', might imply that more exercise is *always* beneficial, when in fact over-exercise could be detrimental to the user's health.

A common theme emerged relating to the overly literal interpretation of natural language statements by participants. Because statements in Exist were constructed in the form "*You are/do/get more X when you are/do/get more Y*" people understood this to mean that activities X and Y were simultaneous (e.g. Quotes 13,14). However, many correlations relate to factors that are temporally disconnected. For example, Participant 8 had previously stated that he avoided listening to music whilst working. Despite this disconnection, the phrasing of a statement showing a correlation between these two variables led him to believe that the data advocated listening to music whilst he worked (Quote 13). Statements that more explicitly indicate the possibility for two activities to be correlated, without necessarily occurring simultaneously, may help to resolve this misinterpretation.

We also found that many of the participants were inclined to reflect on some correlational relationships in terms of causes and effects (e.g. Quote 15) and that they sometimes thought that the configuration of the natural language statements implied which was which (despite this not being knowable from the data). Furthermore, participants who were unaccustomed to considering statistical measures such as significance and goodness-of-fit often viewed these statements as definitive assertions of fact, even when the correlations were weak. Participants that were familiar with analysing statistical relationships questioned the definitive nature of these statements (e.g. Quote 16). These problems point to broader challenges associated with information presentation that is misleading due to poor design or oversimplification.

**Sensemaking Challenge 3: Inferences and actions are hindered by a lack of transparency in outputs**

Many of the difficulties that participants experienced whilst reviewing the insights provided by the system were a consequence of their distrust of the data and the methods by which it was analysed. For example, several participants complained that self-reported measures (e.g. mood) overlooked fluctuations that could not be captured by a single rating for an entire day (e.g. Quote 17). Similarly, automated tracking measures, such as those gathered by wearable activity sensors, were occasionally viewed as unreliable (e.g. because they failed to accurately capture activities such

**Sidebar 4**:

*Interview Quotes 15-20*

*Q15) You spend more time active when you have a better day...can I reverse those? Because actually I think it's like, I had a good day because I spent more time being active. I knew that already. [P14]*

*Q16) They've made this statement as if its definitive. … But it's not true. I was inclined to believe it because it said it, but then I look at it and go, what does it actually mean there? This is coming from someone who knows about R-values and the significance of correlations, whereas other people might not. I'm guessing that this 39% is...it's the R-value basically. [P1]*

*Q17) Measuring how I feel at the end of the day...I feel great at some points of the day and terrible at others, so asking once at the end of the day might not be accurate. Even once a day is not a good snapshot. [P9]*

*Q18) You tweet more when the night is cooler. Just rubbish! Complete coincidence! [P2]*

*Q19) This one's only got one star confidence, so I don't know what to make of it really. [P16]*

*Q20) I don't know if that's a coincidence because its warm the whole month, but I wouldn't have guessed. I'd do it again over another period of time, a colder period. See what it says then. [P4]*

as swimming and cycling). Participants therefore deliberated over the true nature of the trends being shown, whilst reflecting on the caveats and limitations of the data involved. This meant participants were sometimes reluctant to carry out actions in response to the information that they received.

With regards to the reliability of the analysis, rather than the data itself, some unexpected results were plainly dismissed as being incorrect or coincidental (e.g. Quote 18). These insights were considered improbable, and participants assumed, therefore, that the analysis was erroneous. Participants occasionally questioned the validity of drawing definitive conclusions from the data, given the potential for bias and confounding factors (e.g. Quotes 19, 20). The lack of transparency about which data was or was not included in the correlation analysis occasionally resulted in difficulties interpreting and trusting results (e.g. Quote 21). Participant 18 admired the potential for the system to provide "life changing" information, but was hesitant about acting on this information without greater transparency in the data analysis (Quote 22).

Some participants encountered counter-intuitive information, for example discovering that their productivity (in terms of time spent productively on their computer) correlated with their distracting time (also on their computer). One participant struggled to arrive at a logical interpretation of this correlation (Quote 12), questioning whether the system was advocating distraction as a mechanism for being more productive, or whether something had "gone wrong" in the analysis. In fact, this correlation appeared because both productivity and distracting time were correlated with overall time spent on a computer. Data representing productivity and distracting time as *percentages* of overall time spent on a computer, rather than *absolute* duration values were argued to be more appropriate by some participants. Several participants struggled to interpret correlations where the units of measurement and the possibilities for co-linearity were not apparent to them.

**Sensemaking Challenge 4: Holistic insights are difficult to obtain from disjointed outputs**

A final theme that surfaced was that participants sought to associate findings related to multiple, discrete correlations, e.g. seeking a chain of association between three or more facets of their life to gain a more holistic understanding. Participant 3 illustrated this by trying to understand the associations between the quality of her sleep, the temperature at night, and her mood (Quote 23). The Exist system provided few explicit mechanisms to support this linking activity, beyond grouping pairwise correlations that involved a particular attribute on a single screen. While this enabled participants to identify multiple correlates of a particular variable (e.g. Quote 24), further exploration was required to understand the factors associated with each of these correlates. We observed participants struggling to locate and pool relevant information in order to obtain the holistic insights that they were seeking. A challenge for systems of this kind is therefore not only to aggregate many data inputs, but also to provide mechanisms for combining their discrete outputs to produce more integrative, holistic insights about their users.

**Implications for Interactive Health Systems**

Personal informatics systems that process health-related data offer the potential for users to understand and manage aspects of their health at home, independent of their interactions with clinicians.

***Sidebar 5**:*

*Interview Quotes 21-25*

*Q21) I think this correlation might be thrown off by a few outliers. There's one day there where I got really little sleep, and uh, I wasn't productive…I don't know how much those extreme days would throw off this correlation. Maybe over time it would. [P1]*

*Q22) To really trust it, and base some life changing decision on it, I'd want to see how it does these correlations. [P18]*

*Q23) There's something about my sleep when the night is cooler, and my mood when I get good sleep, so this is mood when the night is cooler, so are they like in a circle? You know...connected. [P3]*

*Q24) So what makes me more active? Sunnier days, the weekend and when I have a better night's sleep. Which one affects it the most though? [P4]*

*Q25) I'd rather just take to this to my doctor and get him to make sense of it [P12]*

However, the challenges associated with making sense of outputs from these systems affect their potential for situated use (e.g. Quote 25). The issues of information overload, misinterpretation from poorly designed and overly simplified outputs, lack of transparency, and difficulty in combining disparate outputs all contribute to the potential for users to make incorrect inferences about their health and take incorrect courses of action that could be ineffective or harmful. We consider the need to address sensemaking challenges to be particularly acute in a health context, due to the causal relationship between poor education and poor health [7]. That is to say, those who stand to benefit significantly from access to interactive health systems may also experience reduced information literacy, affecting their ability to identify, assess, and effectively use information. Hence, improving the ease with which non-expert analysts can make sense of the information output of these systems is imperative. At the WISH workshop we hope to discuss the challenges of using interactive health systems from a sensemaking perspective and explore a research agenda to provide solutions to the problems we have identified.

## REFERENCES


1. 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 Trans. Comput.-Hum. Interact. 20, 5, Article 30 (November 2013), 27 pages. http://dx.doi.org/10.1145/2503823
2. Simon L. Jones. 2015. Exploring correlational information in aggregated quantified self data dashboards. In Proc. of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct). ACM, New York, NY, USA, 1075-1080. http://dx.doi.org/10.1145/2800835.2800963
3. Ryan Kelly and Stephen J. Payne. 2014. Collaborative web search in context: a study of tool use in everyday tasks. In Proc. of the 17th ACM conference on Computer supported cooperative work & social computing (CSCW '14). ACM, New York, NY, USA, 807-819. http://dx.doi.org/10.1145/2531602.2531617
4. Sukwon Lee, Sung-Hee Kim, Ya-Hsin Hung, Heidi Lam, Youn-ah Kang, and Ji Soo Yi. "How do People Make Sense of Unfamiliar Visualizations?: A Grounded Model of Novice's Information Visualization Sensemaking." Visualization and Computer Graphics, IEEE Transactions on 22, no. 1 (2016): 499-508. http://dx.doi.org/10.1109/TVCG.2015.2467195
5. Lena Mamykina, Arlene M. Smaldone, and Suzanne R. Bakken. "Adopting the sensemaking perspective for chronic disease self-management." Journal of Biomedical Informatics 56 (2015): 406-417. http://dx.doi.org/10.1016/j.jbi.2015.06.006
6. Sharoda A. Paul and Meredith Ringel Morris. 2009. CoSense: enhancing sensemaking for collaborative web search. In Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA, 1771-1780. http://dx.doi.org/10.1145/1518701.1518974
7. Mary A. Silles. 2009. "The causal effect of education on health: Evidence from the United Kingdom." Economics of Education review 28.1:122-128. http://dx.doi.org/10.1016/j.jbi.2015.06.006
8. Steve Whittaker. 2008. Making sense of sensemaking..In *HCI remixed*: Reflections on works that have influenced the HCI community, MIT Press, Boston, MA, USA.