Evaluating the Effectiveness of Three Physical Activity Visualizations — How People Perform vs. Perceive

Xin Tong, Diane Gromala, Lyn Bartram, Fateme Rajabiyazdi, Sheelagh Carpendale

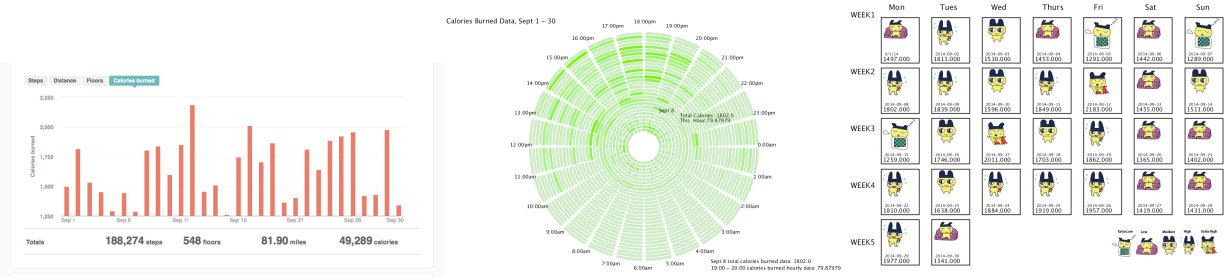

Fig. 1. Three Visualizations Representing Physical Activity in the Study, Showing One-month Data Visualized in a *FitBit* Bar Chart (left); a Circular Ringmap (center); and a Virtual Pet Calendar (right).

Abstract—In this paper, two different types of personal physical activity data visualizations – a Circular Ringmap and a Virtual Pet visualization – are proposed and compared to a widely used commercial approach, the *FitBit* bar chart. A study was conducted to evaluate the effectiveness of the three visualizations based on participants' task performance and self-reported experience. The results showed that in objective terms, there were not significant differences in participants' time spent on the task and their error rates across the three visualizations. However, in subjective terms, participants did perceive the visualizations differently. The pragmatic bar chart won on several attributes (readability, comprehension, and awareness), and the Virtual Pet visualization triggered an emotional response (empathy). Although the Circular Ringmap visualization helped participants with identifying activity patterns, it was not perceived to be intuitive or helpful overall. From this study, first, we learned that how people perform a task does not equal how participants perceive. Second, decisions in designing specific visualizations related to personal data should depend each individual goals and contexts. Last, the role of emotion, which was triggered by the design of visualizations, is discussed.

Index Terms—Personal Informatics, data visualization, effectiveness, physical activity, emotion feeling

1 Introduction

Visualizations are crucial in personal data analytics to help people gain insights, set goals, and self-reflect on their physical activity [1]. By using visualization tools, one's awareness of physical behaviours increases [3], which can help them to make smarter decisions and to support them in changing their behaviour in regards to their personal health.

Visualizations designed for personal information that relate to physical activity should facilitate the process of understanding data and finding activity patterns. New methods of visualizing personal information [2] have been designed in number of research projects and activity systems. For example, *Data in Everyday Life* [7] represents data in a form of a digital calendar; *QS Spiral* shows the top of a clock dial metaphor with a circle representing a time span [3]; *Fish'n'Steps* [4] uses a social computer game to promote physical activity; and the virtual plant in *UbiFit* [2] visualizes a garden that grows according to owners' physical activity.

Many commercial wearable devices are lowering the effort that individuals need to exert to collect and understand their data. To better engage people, *FitBit* visualizes data in multiple ways – including a bar chart (Figure 1, left) and an animated growing flower. But it is still unclear whether people can effectively identify and understand data presented in different ways, what people subjectively prefer in regards to such visualizations and the degree to which subjective preference can play a role. Further, few studies compare the objective findings with subjective experience.

While it may be sufficient for visualizations designed for use in work environments to focus on studying comparative time and efficiency, this may not be the only important factors when creating visualizations for our everyday lives. For example, while being accurate and efficient may be both necessary and sufficient for work, activities in one's personal life may be complex and affected by many other factors, such as emotional response. Therefore, in this study, we investigated people's performances in a visualization task and then explored their self-reported evaluations of visualization attributes – Readability, Comprehension, Awareness, Attractiveness and Empathy.

2 Visualization Design

In this section, we developed two approaches of visualizing personal physical activity data, the Ringmap and the Virtual Pet visualizations. In this study, we then compared these two visualizations to the *FitBit* bar chart visualization.

FitBit has a commercial dashboard for people to access their personal health data online. Most of the personal data is displayed by bar charts (Figure 1, left). People can switch to different data tabs in the dashboard to assess their steps taken, distance, floors or calories burned. The interface is interactive in order to display the daily summary, and adjusted to shoe various ranges of time or specific periods of time.

- *Xin Tong, Simon Fraser University. E-mail: tongxint@sfu.ca.*
- *Diane Gromala, Simon Fraser University. E-mail: gromala@sfu.ca.*
- *Lyn Bartram, Simon Fraser University. E-mail: lyn@sfu.ca.*
- *Fateme Rajabiyazdi, University of Calgary. E-mail: f.rajabiyazdi@ucalgary.ca.*
- *Sheelagh Carpendale, University of Calgary. E-mail: sheelagh @ucalgary.ca.*

Manuscript received 31 Aug. 2015; accepted 01 Aug. 2015. Date of Publication xx Aug. 2015; date of current version 25 Oct. 2015.
For information on obtaining reprints of this article, please send e-mail to: *reprints@ieee.org*.

The Ringmap visualization is a circular structure that represents one month of time (Figure 1, middle). It traces and tracks variable processes of activity data and is based on our prior visualization design – *Life Cycle* [5]. Each ring represents a day starting on the top at 12:00 am. The time is arranged like a clock and reads clockwise; the only difference is that this visualization covers an entire day rather than 12 hours. Each ring is divided into equal slots of 5 minutes. The physical activity levels are mapped onto each slot according to brightness. The brighter the slot, the more active the person was in those five minutes. When the mouse is moved over a time slot in this visualization, the exact physical activity data and time is displayed.

The particular approach that influenced the Virtual Pet visualization was motivating behaviour changes by cultivating a strong internal awareness of one's physical activity through focusing on the care of pets or plants. Figure 1 (right) shows different physical conditions by date in a calendar view. It categorized physical activity data into five levels: extra high, high, medium, low, and extra low. Those categories were then mapped onto the virtual pet's healthy condition or the virtual plant's growth as depicted in Figure 2. The actual five figures depicted in this visualization were designed by mapping a person's approximate activity level to possible real life scenarios. For instance, being a couch potato stands for a low level of activity, whereas being a super hero suggests an extremely high or "extra high" level of activity.

Fig. 2. The actual five figures used in the Virtual Pet visualization

3 The Study

The goal of this study was to compare the effectiveness of the bar charts, the Circular Ringmap, and the Virtual Pet calendar visualizations, and to understand how may they influence peoples' experience.

Eighteen participants (8 males and 10 females, aged from 19 to 29) were recruited among university students. Six participants each were randomly assigned to three groups; each group was provided with one visualization type. The same data set was used for all three visualization types. Next, a quantitative experiment was conducted, measuring between-groups. The definitions of five attributes were given to participants as a reference for their ratings of the visualization they interacted with in the task. The five attributes were:

(1) Readability: to what extent could they read the visualization data.

(2) Comprehension: to what extent did they understand what the visualization data represented in order to identify different activity levels.

(3) Attractiveness: to what extent were they attracted by the graphic design in terms of formal qualities (color, hierarchy, data organization methods and so on), as well as the metaphors and the data mapping.

(4) Awareness: how easy or difficult was the visualization for them to self-reflect on the activity conditions (to distinguish whether the activity level was at a high or low level in certain time periods) represented by the data visualization, and in answering the questions posed by the task.

(5) Empathy: to what extent did the visualization evoke an emotion or feeling.

The study consisted of 3 sections: the pre-test, visualization task and post-test. Each section had a questionnaire, and the data was collected from these questionnaires. The main types of questions that were asked in the task were to: (1) Identify the data's range in a specified day in the month; (2) Identify the activity levels of a specified day in the month (from five categories: extra high/high/medium/low/extra low); (3) Identify if the activity level of the specified day or week is considered enough to maintain health (the definition of "healthy" in this task: over half of the days in a month has an activity data that is above average); (4) Explore the activity pattern presented in the visualization and when it occurs.

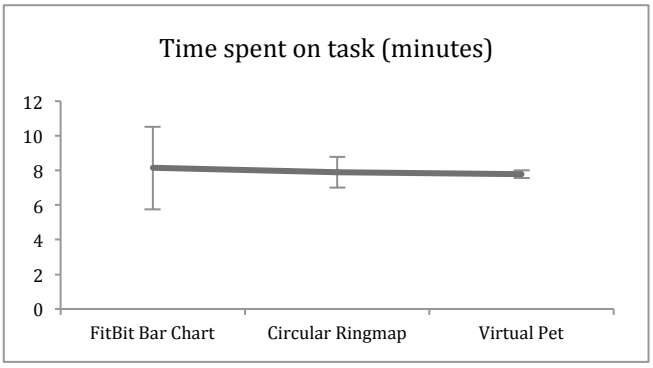

Fig. 3. Time spent on task of the three visualization groups.

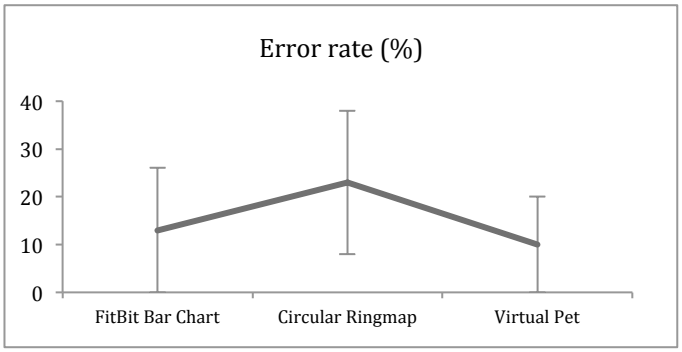

Fig. 4. Error rates of the three visualization groups.

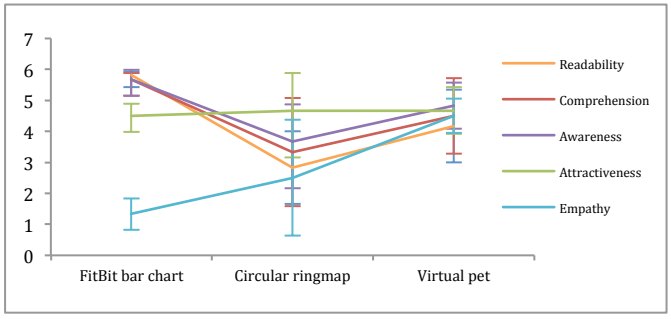

Fig. 5. Self-reported evaluations of the three visualization groups.

For the three groups, as shown in Figure 3, the time spent on task (in minutes) for the *FitBit* bar chart was (M = 8.17, SD = 2.43); for the Circular Ringmap (M =7.89, SD = .88); and for the Virtual Pet (M = 7.77, SD = .22). No significant differences were found among the three groups, $F(2, 17) = .04$, $p = .96$. The error rates were: the bar chart (M = .13, SD = .13), the Circular Ringmap (M = .23, SD = .15), and the Virtual Pet (M = .10, SD = .10). As shown in Figure 4, no significant differences were found among the three groups either, $F(2, 17) = .89$, $p = .46$,.

The self-reported evaluations, however, revealed differences. For the post-test self-reported evaluation, as shown in Figure 5, there were strong significant differences among three visualizations in Readability, $F(2, 17) = 14.02$, $p < .01$, and Empathy $F(2, 17) = 7.61$, $p < .01$. General significant differences were also identified in

participants' ratings of Awareness, $F(2, 17) = 5.86$, $p < .05$, and Comprehension, $F(2, 17) = 5.07$, $p < .05$. However, no significant difference appears among all in terms of Attractiveness, $F(2, 17) = .02$, $p = .98$.

Tukey-HSD tests were run to figure out exactly which pairs have significant differences. From the results, the *FitBit* bar chart and Circular Ringmap showed a significant difference in Comprehension and Awareness. The Virtual Pet had a strong significant difference compared with the *FitBit* bar chart in Empathy, and it also had a significant difference compared with the Circular Ringmap visualization. In terms of Readability, there were significant differences among all three groups – the *FitBit* bar chart had a strong significant difference from both the Circular Ringmap and the Virtual Pet visualizations; the Circular Ringmap and the Virtual Pet visualizations also had a significant difference.

From the results of this study, non-significant differences of task performance were found among the three visualization groups. Interestingly, however, participants' self-reported evaluations of visualization attributes from the post-test questionnaires tell us a different story. Participants perceived the *FitBit* bar chart to be the visualization that was easiest to read, comprehend, and become aware of – how physically active the data represents. The Ringmap visualization ranked last in all attributes. Although the Virtual Pet visualization was rated in the middle for Comprehension and Awareness attributes, there were no significant differences between it and the *FitBit* bar chart. For Empathy, however, strong effects were found between the Virtual Pet and the *FitBit* bar chart, and between the Virtual Pet and the Ringmap.

4 Discussion

The main study results revealed that the Ringmap visualization was not helpful or intuitive enough for the participants to reinterpret. It is unclear whether the novelty of the visualization format or if representing 24 hours instead of 12 hours in a circular format played a part in the results. In comparison, the Virtual Pet visualization was considered to be the visualization that evoked the most empathy.

The Empathy attribute matters in this study, because it tells us there is some relationship between the visualization and the individual's emotions. Emotional engagement may have a role to play in visualizations that are designed for personal informatics, especially physical activity for a number of reasons. First, people are invested in attending to these visualizations over time, for personal reasons, often involving improving their health. Second, numerous studies show that to promote behaviour changes implicit in physical activity data, extrinsic motivation has a short term boost but it is not sufficient enough to sustain engagement. In contrast, intrinsic motivation engages people at an emotional level and can sustain a longer time of motivation and engagement [6]. Emotional engagement is one of the critical factors that personal data visualizations use to evoke potential intrinsic motivation for lifestyle improvement; in this case, improvement of physical activity.

Moreover, interesting results were discovered when we compared the two measurements gathered from the study – the objective efficiency of performances in a visualization task and the subjective self-reported evaluations of the visualization type. In objective terms, the three groups achieved the same task performance, indicating that participants did not see differences across the three visualizations. However, in subjective terms, participants did perceive the five attributes of the visualizations differently, especially for the Readability and Empathy attributes. This is interesting, especially because the approaches to how the data was visualized varied considerably.

From this study, we learned that how well people perform does not represent how they perceive the visualizations. Therefore, we may conclude that the efficiency and errors are independent of the subjective qualities. Although this fact, the result from the study, is normally realized by the community as a common sense, which has not been demonstrated or actually investigated with concrete result findings. The valuable meaning of this independence between objective efficiency and subjective evaluation is that there can be a favourite visualization perceived by specific population of people that is really inefficient. Also, it indicates that the new visualization method – the Virtual Pet visualization – might be achievable, and that such a method does both — achieves good results for time and errors in task performance, as well as for engaging emotions.

Furthermore, according to the open questions in the post-test questionnaire regarding participants' preferences of visualizations, we summarized four important findings as design implications for future personal data visualization research:

First, to make quantified personal health data more beneficial for identifying activity patterns and for self-reflection, a large volume of data needs to be easily accessible, so that the patterns, trends and progress can be easily comprehended and monitored. Although this seems obvious, there are nuances worth mention, especially in terms of detail and pattern recognition. For instance, in the Circular Ringmap, participants reported it was easier to identify patterns when large amount of data was structured in the ringmap shape with the detailed information that the five-minute intervals provided. In the Virtual Pet visualization, participants said they were only able to find the day-level physical activity pattern, but not on the hourly level. However, for the *FitBit* bar chart visualizations, participants reported that it was difficult for them to figure out any physical activity pattern.

Second, different visualization types afford understanding the data in diverse ways, and having multiple visualizations on offer may address the diversity of needs that people have in understanding the different things that their personal data can reveal. From the open questions in this study, participants independently also suggested that it would be helpful to combine two types of visualizations together to get a deeper and better understanding of the physical activity data over the entire month. One participant suggested using bar or line graphs for establishing an initial awareness, and for allowing him to search for specific information. The participant then suggested that the Circular Ringmap would be more useful for reflecting on activity patterns. Similarly, another participant proposed combining the *FitBit* bar chart with the Virtual Pet visualizations because it could provide both accurate information and with more engagement and enjoyment. By including more than one type of visualization, the data may be presented in more meaningful (and potentially more specific) ways, which may, in turn, increase people's engagement with the data over time.

Third, when the visualizations are in the domain of personal analytics and in the sub-domain of physical activity, exploring and incorporating factors related to emotional engagement into efficient visualizations may aid people with maintaining their motivation for changing their behaviours over time. From the study findings, an example is that the Virtual Pet visualization was deemed to be the most empathetic. This is an example of how design metaphors, such as that represented by the Virtual Pet visualization, could be useful in this regard. Moreover, in the *FitBit* bar chart visualization, new interactions can be extended and implemented in order to provide a faster exploration of activity patterns for long-term and large-scale data sets. Color-coded mechanisms for representing different levels of activity intensities is also recommended, because it could benefit people in identifying and counting active/inactive days more easily.

5 Conclusion

In this paper, we first introduced two data visualizations of people's physical activity levels in different structures, and compared them with popular and commercially used bar chart visualization. Next, we compared the effectiveness of the three visualizations in a visualization task and asked participants for self-reported evaluations of five attributes (Readability, Comprehension, Awareness, Attractiveness and Empathy).

The findings from this study proved that the efficiency and error rates of the visualization are independent from its subjective qualities

and participants' affective experiences. The significance of this study lies in the fact that it provides concrete evidence that the commonly tested for factors of time and errors (or efficiency) are independent of the more subjective factors of emotional response and awareness. It opens new visualization design challenges around how to create visualizations that combine these factors, and that these factors may have an important role to play in personal data visualizations.

In fact, participants suggested approaching this challenge from both directions: (1) to incorporate more emotive factors into efficient visualizations; and (2) to increase the readability of attractive and more emotionally resonant visualizations. As visualizations move increasingly into people's everyday lives, these factors may become increasingly important.

REFERENCES

[1] I. Li, A. K. Dey, and J. Forlizzi, "Understanding My Data, Myself: Supporting Self-reflection with Ubicomp Technologies," in *Proceedings of the 13th International Conference on Ubiquitous Computing*, New York, NY, USA, 2011, pp. 405–414.

[2] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. Smith, and J. A. Landay, "Activity Sensing in the Wild: A Field Trial of Ubifit Garden," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2008, pp. 1797–1806.

[3] J. E. Larsen, A. Cuttone, S. L. Jørgensen, J. E. Larsen, A. Cuttone, and S. L. Jørgensen, "QS Spiral: Visualizing Periodic Quantified Self Data," *Proc. CHI 2013 Workshop Pers. Inform. Wild Hacking Habits Health Happiness*, 2013.

[4] J. J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, and H. B. Strub, "Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game," in *UbiComp 2006: Ubiquitous Computing*, P. Dourish and A. Friday, Eds. Springer Berlin Heidelberg, 2006, pp. 261–278.

[5] X. Tong, D. Gromala, C. Shaw, and P. Clarke, "Transformation between Electronic Arts and Chronic Pain: Long-term Body Activity Data Visualisation and Pain Animation Expression," 2014.

[6] B. Burke, *Gamify: How Gamification Motivates People to Do Extraordinary Things*. Bibliomotion, 2014.

[7] D. Huang, M. Tory, and L. Bartram. "Data in everyday life: Visualizing time-varying data on a calendar." Proc. IEEE VIS Poster (2014).