Trust in activity tracker measurement and its link to user acceptance

Daniel Trommler[1], Christiane Attig[1], Thomas Franke[2]

Department of Psychology, Cognitive and Engineering Psychology, Chemnitz University of Technology, Chemnitz, Germany[1]
Institute for Multimedia and Interactive Systems, Engineering Psychology and Cognitive Ergonomics, University of Lübeck, Lübeck, Germany[2]

daniel.trommler@s2010.tu-chemnitz.de,
christiane.attig@psychologie.tu-chemnitz.de,
franke@imis.uni-luebeck.de

Abstract

Inaccuracies in data measurement can impair trustworthiness of activity trackers (i.e., wearable fitness devices) and, thus, constitute a usability challenge possibly impairing user acceptance. With the present research, we aim at advancing understanding of perceived trustworthiness of activity tracker measurement and its relevance for user acceptance. N = 79 users of activity trackers were surveyed regarding their daily interaction with the tracker, user experience, and user acceptance. Results indicated a substantial variance in perceived trustworthiness. Many users perceived suboptimal trustworthiness of the tracking of their activity data, indicating potential for optimization. Further, analyses showed that higher perceived trustworthiness of activity tracker measurement was indeed linked to higher user acceptance. The results highlight the potential of enhancing user acceptance of activity trackers by improving perceived trustworthiness of activity measurement.

1 Introduction

Activity trackers (i.e., fitness trackers, smartwatches) have been designed to enhance motivation for physical activity by providing physical activity data (e.g., step count, calorie consumption) as feedback to users (Yang et al., 2015; Attig & Franke, 2018). This automated data collection, processing, and presentation can support users by facilitating the comparison of the current activity level with a specified activity goal. However, what happens when the underlying data is precise but inaccurate?

From the perspective of the continuous feedback loop of self-regulation (Carver & Scheier, 1998) activity trackers constitute a partial automation of self-regulation, particularly automating the input function (i.e., perception of relevant information) as well as assisting the definition of the reference value (i.e., goal setting), the comparator function (i.e., comparison of reference and input), and sometimes even assisting the output function (i.e., providing tips which activity can help to reduce the discrepancy between input and reference).

A key variable that has gained much attention in the broader automation literature is trust in automation (e.g., Hoff & Bashir, 2015). Since automation assumes tasks from the user (e.g., data collection), the user transfers responsibility to the automated system and has to rely on the assumption that the automation works correctly (Hoff & Bashir, 2015). If this assumption is violated, trust can decrease, possibly leading to impaired user acceptance (e.g., Beggiato, & Krems, 2013). Indeed, trackers can have difficulties in recognizing human activities accurately, possibly resulting in inaccurate feedback (Bedogni et al., 2012) and a decrease in trust. Decreased trust has been shown to be connected to a decreased activity tracker usage intention, at least in a short-term experimental setting (Rupp et al., 2016). However, further issues of decreased trust, especially in real-life and long-term use, are unclear.

The objective of the present research was to examine the perceived trustworthiness of activity tracker measurement in everyday usage and the relation of trustworthiness to user acceptance. To this end, an online study with actual users of activity trackers was conducted that examined the following research questions: (Q1) To what extent do users perceive optimal trustworthiness of activity tracker measurement?, and (Q2) To what extent is perceived trustworthiness of activity tracker measurement related to user acceptance of activity trackers? Moreover, we hypothesize (H1) higher ratings of perceived trustworthiness of activity tracker measurement to be related to higher ratings of acceptance of the activity tracker.

2 Method

2.1 Participants

Participants were recruited via relevant Facebook groups on activity tracking and, additionally, in local gyms. Requirement for participation was the regular usage (at least once a week) of a wearable activity tracker (wristband/smartwatch) that at least offered steps and calories measurement. The N = 79 users who completed the online questionnaire had an average age of 34.4 years (SD = 10.5); 62% were female. The day before taking part in the questionnaire, participants' activity data showed on average M = 14,440 steps (SD = 6,984) and a calorie consumption of M = 2,530 kilocalories (SD = 915); 78% stated to wear their tracker for 23-24 hours a day and 87% stated to wear the tracker 7 days a week.

2.2 Scales and measures

The 12-item *trust in automated systems* (TIAS) scale (Jian et al., 2000, German translation by Beggiato & Krems, 2013) and the 5-item *facets of system trustworthiness* (FOST) scale

(Franke et al., 2015) were applied to assess perceived trustworthiness of activity tracker measurement. The TIAS scale is a frequently used general-purpose scale for trust assessment. The FOST scale is speficially tailored for assessing trust in information displays. Both scales were administered for steps and calories separately and instruction and item texts were specified accordingly (e.g., “system” in TIAS specified to “step count measurement”). Reliability (Cronbach’s alpha) was excellent (steps: α_{TIAS} = .92, α_{FOST} = .95; calories: α_{TIAS} = .92, α_{FOST} = .96). On both scales, participants provided answers on a 6-point Likert scale from 1 – *completely disagree* to 6 – *completely agree*. Further, the 9-item Van der Laan acceptance scale (Van der Laan et al., 1997) was used to assess user acceptance (α = .93).

3 Results

Results regarding Q1 showed substantial variance in perceived trustworthiness (see values for the 25th and 75th percentile of trust ratings in the sample in the second row of Table 1) and a substantial share of participants with suboptimal perceived trustworthiness ratings (see last three rows in Table 1). Interestingly, for both scales, the average trustworthiness for calories was significantly lower than for steps ($p_{TIAS} < .001$, $d_{TIAS} = 0.58$, $p_{FOST} < .001$, $d_{FOST} = 0.51$; moderate effects). *T*-tests against the test value 6 showed significant differences ($p < .001$, $d_{FOST\text{-}steps} = 1.30$, $d_{TIAS\text{-}steps} = 1.52$, $d_{FOST\text{-}calories} = 1.56$, $d_{FOST\text{-}calories} = 1.88$; large effects) for both scales and both data types, indicating that participants did not experience the data measurement as perfectly trustable, reliable, and traceable (example items of FOST scale). Also compared to the evaluation of a typical electric vehicle range information interface ($M_{FOST} = 5.01$, $SD_{FOST} = 0.66$; Franke et al., 2015) the ratings in the present study were somewhat lower, especially for calories. Finally, there was a strong relation between the two trust scales TIAS and FOST ($r_{steps} = .89$, $p_{steps} < .001$; $r_{calories} = .91$, $p_{calories} < .001$) and means and standard deviations showed similar values, indicating that the scales have good convergent validity.

	TIAS		FOST	
	Steps	Calories	Steps	Calories
M (*SD*)	4.65 (0.89)	4.05 (1.04)	4.58 (1.09)	3.93 (1.33)
P25 / P75	4.08 / 5.42	3.17 / 4.83	3.80 / 5.40	2.80 / 5.00
Ratings <6	95%	97%	86%	92%
Ratings <5	63%	80%	48%	70%
Ratings <4	20%	43%	27%	39%

Table 1: Descriptive statistics for the two scales assessing trustworthiness of activity tracker measurement.

To examine Q2, we used Pearson correlations (*r*). Due to some violations of normality, we also calculated Spearman (ρ) correlations. There were large significant correlations between trustworthiness of step measurement and acceptance regarding both scales ($r_{FOST} = .59$, $r_{TIAS} = .64$, $\rho_{FOST} = .58$, $\rho_{TIAS} = .60$, $p < .001$ for all effects). For calories, effects were moderate to large and also significant for both scales (r_{FOST} = .49, r_{TIAS} = .52, ρ_{FOST} = .53, ρ_{TIAS} = .52, p < .001 for all effects). Thus, H1 was supported.

4 Discussion

The results show a high variance in the ratings of trustworthiness of activity tracker measurement, indicating potential for optimization of perceived trustworthiness. Moreover, perceived trustworthiness is strongly connected to user acceptance. Thus, to enhance user acceptance of activity trackers, interface designers should focus on improving trustworthiness of activity tracker measurement by improving measurement accuracy. However, this might be challenging given the available sensor technologies. A deeper understanding of subjective factors influencing trustworthiness of activity trackers could facilitate tracker design targeted towards enhancing trustworthiness. Hence, research should examine factors influencing trustworthiness beyond measurement accuracy, for instance, transparency of measurement (and measurement errors).

Moreover, the effects of low perceived trustworthiness and impaired user acceptance on activity tracker abandonment should be examined. Issues in trustworthiness might be one factor contributing the currently high rate of abandonment of activity tracker usage (e.g., Epstein et al., 2016).

However, it should be noted that the examined participants only reflect a certain user group of activity tracker users, namely users with a high usage intensity and a rather high activity level. Further user groups (e.g., novice users, casual users, former users) should be examined in further research to gain insight concerning generalizability of the aforementioned findings.

Finally, the high convergence between the two trust scales (TIAS and FOST) indicates that the shorter FOST scale can be used as a highly economical alternative to measure trustworthiness of information interfaces in the context of automated systems.

References

Attig, C., & Franke, T. (2018). I track, therefore I walk – Exploring the motivational costs of wearing activity trackers in actual users. *International Journal of Human-Computer Studies*.

Bedogni, L., Di Felice, M., & Bononi, L. (2012). By train or by car? Detecting the user's motion type through smartphone sensors data. In *2012 IFIP Wireless Days* (pp. 1-6). Red Hook, NY: Curran Associates.

Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F*, *18*, 47-57.

Carver, C. S., & Scheier, M. F. (2000). On the structure of behavioral self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Ed.), *Handbook of self-regulation* (pp. 41-84). San Diego, CA: Academic Press.

Epstein, D. A., Caraway, M., Johnston, C., Ping, A., Fogarty, J., & Munson, S. A. (2016). Beyond abandonment to next steps: Understanding and designing for life after personal informatics tool use. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing System* (pp. 1109-1113). New York, NY: ACM.

Franke, T., Trantow, M., Günther, M., Krems, J. F., Zott, V., & Keinath, A. (2015). Advancing electric vehicle range displays for enhanced user experience: The relevance of trust and adaptability. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 249-256). New York, NY: ACM.

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*, 407-434.

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*, 53-71.

Rupp, M. A., Michaelis, J. R., McConnell, D. S., & Smither, J. A. (2016). The impact of technological trust and self-determined motivation on intentions to use wearable fitness technology. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 1434-1438). Los Angeles, CA: SAGE.

Van der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C*, *5*, 1-10.

Yang, R., Shin, E., Newman, M. W., & Ackerman, M. S. (2015). When activity trackers don't „fit": End-user difficulties in the assessment of personal tracking device accuracy. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 623–634). New York, NY: ACM.