Proceedings of the 53rd Hawaii International Conference on System Sciences | 2020

Identifying Quality Factors for Self-Tracking Solutions: A Systematic Literature Review

Barbara Schiller
University of Duisburg-Essen
barbara.schiller@uni-due.de

Tobias Brogt
University of Duisburg-Essen
tobias.brogt@uni-due.de

J. Peter M. Schuler
University of Duisburg-Essen
j.peter.m.schuler@uni-due.de

Gero Strobel
University of Duisburg-Essen
gero.strobel@uni-due.de

Stefan Eicker
University of Duisburg-Essen
stefan.eicker@uni-due.de

Abstract

Self-tracking solutions have become globally widespread, as they promise numerous advantages (e.g. improving health) to their users. Despite their benefits, such solutions are often abandoned due to quality issues. This phenomenon can also be observed for digitized products in general. As self-tracking solutions are hybrid products, combining digital and physical components, traditional domain-independent and abstract quality models like the prominent ISO 25000 standard seem to not cover quality in an appropriate way. We address these issues by answering the research question of which factors affect quality perceptions of different stakeholder groups when interacting in a wearable ecosystem. We use a systematic literature review based on a research protocol to identify and analyze 98 quality-influencing factors from 19 studies that we cluster in a map. The identified factors are compared to the ISO 25000 standard, showing that certain factors like hedonic motivation are unconsidered thoroughly in the existing standard.

1. Introduction

Self-tracking, also known as the notion of the quantified-self, lifelogging, or, in its extreme form, self-hacking [13], has become globally widespread, as it promises users the ability to improve their health, become more athletic, and change their behavior [38]. Self-trackers use multi-sensor devices (e.g., wearables, smartphones) and corresponding software (e.g., mobile applications, web platforms) to track a variety of exercise and health parameters, such as calories, water consumption, blood pressure, steps, and sleep time. In fact, self-trackers "track up to 39 parameters of their daily life" [16] to reach goals, document and analyze data, or collect rewards [39]. For this purpose, some self-trackers use multiple self-tracking solutions simultaneously, favoring wearable solutions (e.g., wristband fitness trackers) over smartphones with tracking capabilities to reduce the possibilities of forgetting, losing, or even damaging an expensive smartphone while exercising [38, 39].

However, despite their potential benefits, many people use their tracking devices for a short time only before neglecting them [38, 41]. As self-tracking solutions are used by a heterogeneous community (different genders, ages, and health and fitness conditions) [16, 39], the abandonment seems to be related to issues with the products rather than an overall lack of appeal to certain demographics. Many factors have already been identified as challenges, such as physical design issues [41], privacy concerns, a lack of technical customer support, functional constraints [3], interoperability issues, and low usability [9]. However, these challenges are not specific to self-tracking solutions, but are also characteristics of digitized products in general [32].

A product's ability to satisfy customer needs and expectations through functionality and performance, as well as the perceived value and benefits of an organization's products and services, is traditionally captured by the concept of quality [10]. The International Standard ISO 25000 for Systems and Software Quality Requirements and Evaluation (SQuaRE) differentiates between and proposes models for software and system product quality, quality in use, data quality [22], and IT service quality [23]. However, self-tracking solutions differ in important characteristics from traditional software products or IT services, making the application of the standard questionable.

In accordance with the similar concepts of digitized products, the Internet of Things (IoT), smart, connected products (SCP) and Ubiquitous Computing [35, 44], self-tracking solutions are hybrid products, combining

URI: https://hdl.handle.net/10125/64194
978-0-9981331-3-3

digital (software) and physical components [32], and using sensors, actuators, and analytical components to provide a digital service [14]. In some cases, they operate in digitized product ecosystems that allow users to add other digitized products into the ecosystem and can connect with other ecosystems (systems of systems), as in the case of the smart home and smart city concepts [32, 35]. Ubiquitous systems in general are used in human-centered, personalized physical environments to the extent that they are part of them, can adapt to them, act on them, or even control them [37]. Self-tracking solutions, especially those connected to an external tracker, manifest many of these characteristics. The wearable tracker and the corresponding mobile application and web platform build an ecosystem, which collectively delivers value to the customer, who perceives the ecosystem elements as one product.

However, unlike, for example, smart home or smart industry solutions, self-tracking solutions are cheap(er) and have been more widely adopted, providing researchers “an accessible domain for experimenting with IoT problems” [13]. Different research fields, such as computer science, information systems, and medicine, already study the technical aspects, adoption, benefits, and threats of self-tracking solutions regarding healthcare [9], because the dissemination and the possibilities of technology provide individuals, medical scientists, and other researchers with objective, high-quality data [13, 24]. However, the quantified-self movement is still “an immature domain of research” [9].

Hence, in this paper, we aim to add to the research field of self-tracking by answering the question of ***which factors affect quality perceptions of different stakeholder groups when interacting in a wearable ecosystem***.

We believe the answer to this question does support developers of self-tracking solutions in building better products and giving researchers a better understanding of possible factors influencing the quality of this type of IT product, which might not yet be covered by existing standards. For this purpose, we used a systematic literature review (SLR) to identify quality factors that influence the different entities of a wearable ecosystem (wearable, app/smartphone, web platform). However, we focused on the overview of relevant quality factors rather than the relationships between them, which is a topic for future research. Additionally, we also identified the stakeholder groups that are affected by the different factors, because quality perception can differ between these groups and might influence decisions in the development process. Further, we compared our findings with the ISO 25000 standard to identify factors that are not covered by the standard or the analyzed literature in order to evaluate our findings.

The paper is structured as follows: In the next section, the research method and process are explained in more detail. In doing so, important pieces of information like the search string and study selection procedure are explained. Afterwards, the results are presented and discussed. The paper closes with a reflection on the limitations of this study, a conclusion, and an outlook on future research.

2. Research method

Literature reviews in general “help to identify research problems and gaps and justify the relevance and timeliness of addressing them” [45], while an SLR in particular helps to identify, evaluate, and interpret “all available research relevant to a particular research question, or topic area, or phenomenon of interest” [25]. However, SLRs suffer from certain limitations that make them applicable only under certain conditions: when answering a narrowly defined, summative research question (no “how” or “why” questions) or conducting a bibliometric analysis [4]. Thus, in our case, the SLR is a suitable approach, because our research question is narrow as our focus is on identifying quality factors of self-tracking solutions (not their relationships or possible measurements) and can be answered in a summative form. Using the methodology of [25], we developed a pre-defined protocol specifying the data sources, search terms, selection procedure, exclusion criteria, and methods of data extraction and synthesis.

2.1. Search terms and databases

The search terms were derived from the research question, relevant literature from previous research, the entities of a wearable ecosystem (e.g., its app), and the main areas of research (e.g., mobile Health, mHealth). The terms and the search string were tested against different literature databases and adapted multiple times due to the resulting output or certain restrictions of the databases. The following represents the resulting string:

("quality model" OR "quality requirements" OR "quality understanding" OR "quality perception" OR "quality assessment" OR "quality of experience" OR "service quality" OR "product quality") AND ("fitness tracker" OR "wearable" OR "app") AND ("mhealth" OR "fitness" OR "internet of things" OR "wellness")

Six well-known databases in information systems research were used as primary data sources: AIS Electronic Library (http://aisel.aisnet.org), SpringerLink (http://link.springer.com), Science Direct (http://www.sciencedirect.com), Emerald Insight (http://www.emeraldinsight.com), and Wiley Online Library (http://onlinelibrary.wiley.com). Google

Scholar was not included due to its low precision and overlapping results.

2.2. Study selection procedure

The study selection process used in this study consists of five stages based on the procedures of [1] and [12]:

(1) *Identification of Records:* The database search based on the pre-defined search string resulted in 1701 studies in total. After removing duplicates, 1289 studies remained for further analysis.

(2) *Screening of Studies:* During the screening process, each study was evaluated separately to decide whether it should be included in the final review set. Studies were first screened based on their titles, afterwards on their abstracts and full-text availability, and finally, on their introductions and conclusions. Three researchers screened the studies in discourse to reduce biases, while applying the following exclusion criteria:

Study Type: Result is an anthology of conference proceedings, synopsis, poster presentation, paper session, handbook, interview, discussion, introduction, book chapter, or overview only

Study Findings: Result presents only a concept (e.g., an mHealth app) that was not tested by real users

Study Context: Result has the right context (e.g., mHealth) but does not address (perceived) quality, or result has a different context and its results do not seem to be transferable to the context of wearables/mHealth

Quality and Language: Result is not peer reviewed and not written in English

Regarding the study type, we focused on conference and journal articles, as we did not expect valuable insights from synopses and similar documents, and excluded book chapters, as the research topic itself is quite new. Studies were only excluded when all participating researchers confirmed the exclusion. The screening process removed 1266 studies from the initial set, with 23 remaining.

(3) *Eligibility Test:* During the review and data extraction process, the eligibility based on the full text was tested in parallel, as it was not expected that many studies would be excluded after the intensive screening process of the previous step. Two studies were excluded during the eligibility/review process, as they were research-in-progress papers and seemed to provide no relevant information. One study was excluded because the language and image quality of the study were too low to extract any results. Finally, the paper of [32] was excluded, as it focuses on the concept and challenges of digitized products and not on quality. Although it was discussed whether the challenges themselves might represent quality factors, the team finally agreed that it would require too much interpretation to extract possible quality factors and excluded the paper. Therefore, after the exclusion of the four studies, 19 studies remained as the final set.

(4) *Review and Data Extraction:* Two researchers reviewed a single study and extracted its contents independently into a spreadsheet program. Afterwards, the two compared their results; if consensus could not be reached, a third researcher helped resolve the disagreement. The data extraction spreadsheets contain general information on the studies (e.g., aims of the study, research questions, sample description, domain) and information related to our research question (e.g., quality definition, type of wearable, stakeholders in focus).

(5) *Synthesis of Results:* Four researchers participated in discourse in the aggregative synthesis to integrate the results of the studies. The results are presented and discussed in the following section.

3. Results and discussion

This section describes the results obtained from the SLR and discusses their implications for the research question. It first presents general information about the selected studies, and then details the findings related to the research question and the identified quality-influencing factors. The section closes by comparing the findings with the ISO 25000 standards.

3.1. Publication information of the studies

The final set of review papers consisted of 19 studies (1.47% of the initial set), which are listed in Table 1. The table includes IDs, which will be used later to identify the studies in Table 2. The low acceptance rate is the result of our comprehensive selection process, which ensured to only keep promising studies in the process.

Table 1. Nineteen resulting studies of the SLR with IDs

ID	Study	ID	Study
A	Alnsour et al. 2016 [2]	LU	Lundell and Bates 2016 [27]
B	Bruns and Jacob 2014 [5]	MA	Martinez-Perez et al. 2013 [28]
CL	Calvaresi et al. 2017 [6]	ME	Meulendijk et al. 2014 [29]
CR	Carroll and Richardson 2016 [7]	MO	Moilanen et al. 2014 [30]
CV	Carvalho et al. 2016 [8]	NE	Neuhuettler et al. 2017 [31]
D	Dunn et al. 2016 [11]	PE	Peischl et al. 2015 [33]
G	Gao et al. 2015 [15]	SI	Simons et al. 2013 [42]
HA	Hazarika et al. 2015 [17]	SU	Suryadi and Kim 2017 [43]
HS	Hsiao et al. 2013 [18]	ZA	Zapata et al. 2015 [46]
ID	Idri et al. 2016 [19]		

We did not exclude publications based on the publication date; however, the earliest identified publication dates are as recent as 2013. Self-tracking itself is not a new phenomenon, but the new technological possibilities are, and as such, quality in the context of self-tracking solutions is still quite a new topic.

As the review occurred in 2017, only three studies from that year could be considered. Thus, the studies have been published between 2013 and 2017 with a peak of five studies in 2016 indicating a small, but steady, increase in the interest in the topic.

Of the 19 studies, 10 were published in journals and nine were published as part of conference proceedings. A plurality of the journal articles, three, were published in the Journal of Medical Systems; others include the Software Quality Journal and the Business & Information Systems Engineering Journal. The conference papers were evenly distributed between different conferences such as the Americas Conference on Information Systems (AMCIS), the European Conference on Information Systems (ECIS), and conferences with a focus on Human Computer Interaction (HCI). The distribution of the studies in different journals and conferences shows the interdisciplinary interest in the topic, which has to be considered in future research.

3.2. Quality-influencing factors overview and stakeholder-group categorization

The extraction process revealed 217 quality-influencing factors, including duplicates, synonyms, and homonyms. After removing the duplicates, the synonyms and homonyms were identified by comparing the definitions either provided by the study (or a referenced quality model, theory, or standard) or by searching for a well-known definition provided, for example, by a standard.

However, few studies provided adequate definitions or even used a quality-oriented model, theory, or standard as a foundation for their research: Nine studies did not use any kind of quality model/theory or quality-related model. The remaining 10 studies referenced different theories, standards, and models. Multiple citations included the ISO 25000, ISO 9000, and the Technology Acceptance Model, while some other models, such as SERVQUAL, were mentioned only once. This indicates a missing foundation for quality of self-tracking solutions and the plurality of approaches in use that have to be considered in future research. Further, there seems to be no generally valid standard used in this research field that is sufficient to address all aspects of quality.

After removing synonyms and renaming the homonyms to differentiate them, 114 factors remained. Of these 114 factors, 15 were excluded because the team agreed that they were not quality factors. Some of the excluded factors did not affect the quality perceptions of stakeholders and could not be measured and influenced by changes on the manufacturer side, while others only supplied context and had no influence on the product or quality. Examples of these types of factors include Deceptiveness, Health Care Need, Product Type, and Voluntariness.

While it appeared to be a quality factor at first glance, user experience was also excluded. After comparing different definitions (the corresponding paper did not provide a complete definition), the team concluded that it is not a real factor in itself, but the result of nearly all other factors in combination. Thus, it was excluded for being too high level and being more a result than a factor.

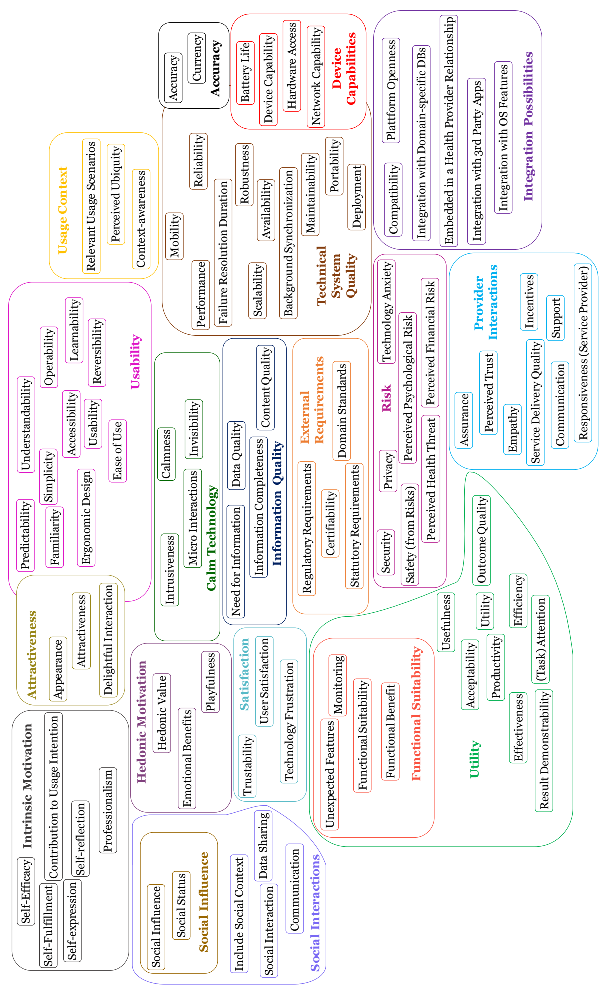

Figure 1. Quality factor map with clusters

Simultaneously with the exclusion of synonyms/duplicates and the renaming of homonyms, all factors were clustered onto a map, as described in the following process, to create a visual overview of the general topics covered by the factors.

In a first step, the factors were placed on the map by comparing their conceptual definitions: The greater the consistency of the definitions, the closer the factors were arranged on the map.

Afterwards in a second step, thematic clusters were identified based on the arrangement, since similar factors were places next to each other. The clusters were labeled to provide categories for the factors. A category therefore contains similar factors and introduces a greater level of abstraction.

The resulting map is shown in Figure 1. The biggest clusters on the map are the *Usability*, *Technical System Quality*, and *Utility* categories, containing 10 to 11 factors each. The smallest clusters are the *Accuracy* and *Social Influence* categories, containing two factors each. The map shows the wide range of quality subtopics covered by the literature.

Regarding the stakeholder groups addressed in the studies, only four groups could be identified: consumers, medical professionals, manufacturers, and developers.

Of the 98 quality-influencing factors, 89 factors were identified as consumer-oriented, 28 as medical-professional-oriented, 21 as developer-oriented, and nine as manufacturer-oriented, with 28 factors relating to more than one stakeholder group (e.g., appearance, privacy, or usability). This could indicate that research on consumer-related quality factors is more profound than that on other stakeholder groups such as manufacturers. However, research opportunities exist not only for manufacturer-related quality factors, but also for those of stakeholder groups that are not present in the studies, such as legislative institutions, which may also have specific quality requirements (e.g., regarding data protection).

The factors are also related to different entities of the wearable ecosystem (wearables, apps, and platform). Forty-eight factors relate to more than one entity, 56 relate to apps, 33 to wearables, and 20 to smartphones.

Table 2. Quality factors in relation to stakeholder groups and entities of the wearable ecosystem, with source studies (A=App, S=Smartphone, W=Wearable, O=Overall)

Quality Factor	Consu-mer	Med. Prof.	Manu-fact.	Deve-loper	Source Studies
Category: Accuracy					
Accuracy	A, W	A	-	A	MA, ME, MO
Currency	W	-	-	-	MO
Category: Attractiveness					
Appearance	A, W, O	A	-	A	LU, MA, NE
Attractiveness	A, O	-	-	-	A, SI, ZA
Delightful Interaction	W	-	-	-	LU
Category: Calm Technology					
Calmness	W, O	-	-	-	CV, MO
Intrusiveness	W	-	-	-	MO
Invisibility	O	-	-	-	MO
Micro Interactions	W	-	-	-	LU
Category: Device Capabilities					
Battery Life	A, W	-	-	-	LU, SU
Device Capability	O	A	-	-	CV, PE
Hardware Access	-	S	-	-	PE
Network Capability	S, O	-	-	-	CV
Category: External Requirements					
Certifiability	A	A	-	A	ME
Domain Standards	-	O	-	-	SI
Regulatory Requirements	-	-	O	O	CR
Statutory Requirements	-	-	O	O	CR
Category: Functional Suitability					
Functional Benefit	A, W, O	O	O	O	A, CL, LU, SI
Functional Suitability	A, W	-	-	-	G, ID, MO
Monitoring	A, O	O	O	O	CL, SU
Unexpected Features	W	-	-	-	LU
Category: Hedonic Motivation					
Emotional Benefits	A	-	-	-	SI
Hedonic Value	A, W, S, O	-	-	-	B, G, NE
Playfulness	O	-	-	-	NE
Category: Information Quality					
Content Quality	A, O	A	-	A	MA, NE
Data Quality	A	-	-	-	SI
Information Completeness	W	-	-	-	MO
Need for Information	A, S	-	-	-	B
Category: Integration Possibilities					
Compatibility	A, O	A	-	-	CV, ID, PE
Embedded in a HealthProvider Relationship	A, O	-	-	-	SI
Integration with 3rd Party Apps	W	-	-	-	LU
Integration with Domain-specific Databases	A	-	-	-	SI
Integration with OS Features	-	S	-	-	PE
Platform Openness	-	A, S	-	-	PE
Category: Intrinsic motivation					
Contribution to Usage Intention	A, W	-	-	-	MO, SI
Professionalism	A, S	-	-	-	B
Self-Efficacy	A, W, O	-	-	-	G, MO, NE, SI
Self-Expression	A, S	-	-	-	B
Self-Fulfillment	A, S	-	-	-	B
Self-Reflection	W	-	-	-	MO
Category: Provider Interactions					
Assurance	A, O	-	-	-	NE, SI
Communication	O	-	-	-	NE
Empathy	A, O	-	-	-	NE, SI
Incentives	O	-	-	-	NE
Perceived Trust	A	-	-	-	SI
Responsiveness (Service Provider)	A, O	-	-	-	NE, SI
Service Delivery Quality	O	-	-	-	NE
Support	O	-	-	-	NE
Category: Risk					
Perceived Financial Risk	A	-	-	-	SI
Perceived Health Threat	W	-	-	-	G
Perceived Psychological Risk	A	-	-	-	SI
Privacy	A, O	A, S	-	A	CV, ME, PE
Safety (from Risks)	O	-	-	-	CV
Security	A, O	A, S	-	A	CV, ID, MA, ME, NE, PE

Quality Factor	Consu-mer	Med. Prof.	Manu-fact.	Deve-loper	Source Studies
Technology Anxiety	O	-	-	-	NE
Category: Satisfaction					
Technology Frustration	A	-	-	-	HA
Trustability	A, O	A	-	A	CV, ME
User Satisfaction	A, O	-	-	-	CV, SI, ZA
Category: Social Influence					
Social Influence	W, O	-	-	-	G, HS, MO, NE
Social Status	W, O	-	-	-	LU, NE
Category: Social Interactions					
Communication	W, O	O	O	O	CL, LU
Data Sharing	W	-	-	-	MO
Social Context	A	-	-	-	SI
Social Interaction	A, S	-	-	-	B
Category: Technical System Quality					
Availability	A, O	A, S	-	A	CV, MA, PE
Background Synchronization	-	S	-	-	PE
Deployment	A	A, S	-	-	PE, SI
Failure Resolution Duration	A, S, O	-	-	A, S, O	D
Maintainability	A	-	-	-	ID
Mobility	A, O	-	-	-	CV, SI
Performance	A, W	A	-	A	ID, MA, MO
Portability	A	-	-	-	ID
Reliability	A, W, O	A, S	-	A	CV, ID, ME, MO, NE, PE, SI
Robustness	O	-	-	-	CV
Scalability	-	-	O	O	CV
Category: Usability					
Accessibility	A	A	-	A	ME
Ease of Use	A, W, S, O	A	-	A	B, CV, G, HS, MA, NE, SI
Ergonomic Design	W	-	-	-	G
Familiarity	A, O	-	-	-	CV, SI
Learnability	A	A	-	A	MA, ZA
Operability	A	-	-	-	ID, ZA
Predictability	W, O	-	-	-	CV, MO
Reversibility	O	-	-	-	CV
Simplicity	O	-	-	-	CV
Understandability	A, O	-	-	-	NE, ZA
Usability	A, W, O	A, O	O	A, O	CL, CR, CV, LU, ME, NE, PE
Category: Usage Context					
Context-Awareness	O	-	-	-	CV
Perceived Ubiquity	A, S, W	-	-	-	B, HS
Relevant Usage Scenarios	W	-	-	-	LU
Category: Utility					
Acceptability	O	-	-	-	CV
Effectiveness	A, W, O	O	O	O	A, CL, CV, G, ZA
Efficiency	A, W, O	O	O	O	CL, CV, MO, NE, SO, ZA
Outcome Quality	O	-	-	-	NE
Productivity	A, S	-	-	-	B
Result Demonstrability	O	-	-	-	NE
(Task) Attention	O	-	-	-	CV
Usefulness	A, W	A	-	-	G, HS, PE, SI
Utility	O	-	-	-	CV, NE

No platform-related factors were identified, but 55 factors were identified as being generally relevant. Thus, most of the identified factors relate either to apps or to the ecosystem as a whole.

Similarly, most factors (73 in total) stem from research on mHealth, while 22 stem from general smartphone and mobile app research. Thirty-one factors stem from research on HCI, 27 from ubiquitous systems research, and only six from research on ambient assistant living (AAL). However, unlike HCI or AAL, mHealth was one of the keywords of this SLR, though it was an optional keyword, as it was only connected via OR in the search string. Thus, there might be biases in these results due to the chosen keywords.

The five most frequently mentioned factors are ease of use (7), reliability (7), usability (7), efficiency (6) and security (6). Even if this is not clear evidence of the importance of the factors, at least it is an indication and for example shows the importance of the *Usability* cluster regarding the quality perception.

The most factors from any study, 27 in total, were provided by [8], which is not surprising, as the authors conducted a systematic mapping study to collect quality characteristics and measures. Therefore, their study is similar to ours, but it focuses on the quality of interactions and ubiquitous systems that are “transparent and calm and keep the user’s attention on his/her main activities” [8]. Thus, they excluded studies that did not match their narrow definition of ubiquitous systems but that would otherwise be relevant in our context.

An overview of the factors, related stakeholder groups, as well as related entities of the wearable ecosystem and source studies is provided in Table 2. Due to the page limit of the conference, we could not include the definitions of the quality factors within this paper.

3.3. Comparison with the ISO 25000 standard

We compared our results to the ISO 25000 series, which represents the prevailing quality standard in the field of systems engineering, to identify uncovered factors in the standard, to check for quality factors not present in the analyzed literature, and to gain further insights for our research question. As explained in the introduction, the standard is domain-independent and thus might not provide all the necessary quality factors for either self-tracking solutions or for SCPs in general. Nevertheless, its popularity, scope, and focus on software products and services makes the ISO standard a proper evaluative tool to provide more insight into our findings.

For the comparison, we used the ISO 25010 System/Software Product Quality Model and Quality in Use Model [21], ISO 25011 IT Service Quality Model and Quality in Use Model [23], and the ISO 25012 Data Quality Model [20].

Each quality model consists of multiple so-called characteristics (e.g., satisfaction) on a first level, and most characteristics consist of multiple sub-characteristics (e.g., usefulness, trust, pleasure, and comfort in satisfaction) on a second level. Therefore, the ISO models are similar to our concept of categories (first level) and quality factors (second level) (cf. Figure 1).

Although the Quality in Use model from the ISO 25011 is based on the model of the ISO 25010, it is more focused on service-level agreements. Thus, we considered both of the Quality in Use models in the comparison. However, during analysis, it became obvious that the service-level agreement specialization did not make any difference. Thus, both models are merged in our resulting figure (cf. Figure 2). It should also be noted that the page limit of the conference did not allow us to display a detailed figure with all sub-characteristics and quality factors.

First Level Comparison: In the first step, we compared the four identically named categories and ISO characteristics: *Accuracy*, (*Functional*) *Suitability*, *Satisfaction*, *and Usability*. Regarding their definitions, the categories and ISO characteristics match with each other. However, they differ in their quality factors/sub-characteristics. For example, *Satisfaction* contains the three quality factors Technology Frustration, Trustability, and User Satisfaction, while the ISO 25010 contains the four sub-characteristics Usefulness, Trust, Pleasure, and Comfort [21].

In the second step, we identified synonyms on the categories / ISO characteristics level based on their corresponding definitions. We identified two synonyms in total. The category *Accuracy*, for example, has the same meaning as the ISO characteristic *Precision*, but these terms again differ in their respective sub-characteristics / quality factors.

First / Second Level Comparison: In the third step, we compared the quality factors with the namely identical ISO characteristics, as well as comparing the namely identical ISO sub-characteristics with our categories. Regarding their corresponding definitions, we found ten matches that are identical in name as well as having the same meaning. For example, User Interface Aesthetics as part of Usability in the ISO 25010 matches with our category *Attractiveness*. We also identified ten matches of quality factors / ISO characteristics and categories / ISO sub-characteristics that are synonymous.

Second Level Comparison: Finally, we identified two synonym matches of quality factors / ISO sub-characteristics and one namely identical match of quality factors / ISO sub-characteristics.

It is notable that approximately 50 percent of the matched categories have roughly the same scope of factors. The biggest difference between the categories can be found in one of our broadest categories: *Usability*. The ISO standard provides only five quality categories, whereas we identified 11 quality characteristics in our research (cf. Table 2). In addition, seven categories and their corresponding quality factors could not be matched with the ISO-provided categories or characteristics; this occurred especially in the areas of *User Motivation* (hedonic, intrinsic), *Social Interactions*, and *Social Influence*, but also in more technical, domain-oriented fields like *Device Capabilities*. The reason for this could be the universality of the ISO standard and the domain orientation of our research.

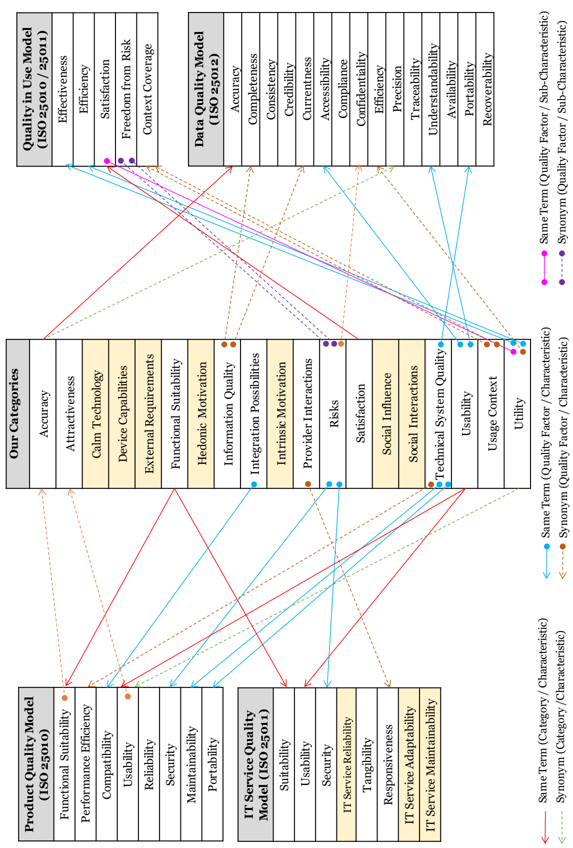

Figure 2. ISO standard category matching

As a result, our research on the one hand shows the importance of customer-oriented quality factors in the domain of self-tracking solutions and on the other hand makes clear that most of these factors are not covered by commonly used quality standards like the ISO standards.

Nevertheless, the ISO standards cover important quality factors (e.g., Satisfaction and Suitability), and therefore are helpful in the context of quality perception of self-tracking solutions. We therefore recommend, that manufacturers should at least use such a generally accepted standard, even if it is lacking some customer-oriented quality factors, instead of avoiding using a quality standard at all.

4. Limitations

To answer our research question, we used a comprehensive SLR, which comes with possible threats to validity. The analyzed biases are derived from the biases mentioned by [25] and [34].

General biases like publication bias (publication of positive results only) cannot truly be avoided. We tried to mitigate this threat in two ways. First, we allowed not just journal articles, but also conference papers. Second, we did not exclude studies based on their publication date in order to consider a wide range of studies.

Research bias, or descriptive validity, concerns the objectivity of the research process and the factual accuracy of the account (e.g., selection driven by researcher expectations, too many interpretation steps, etc.). We tried to mitigate this threat by using a pre-defined protocol, as recommended by [25], basing our selection procedure on other published SLRs and conducting the study with four researchers who collaboratively worked in discourse during the screening, extraction, and synthesis processes. This approach also allowed us to mitigate threats to the theoretical/internal validity (capturing what we intended) and interpretative validity (biases in drawing conclusions).

A special limitation of this study that equally concerns its interpretative and descriptive validity is the lack of standardized terminology in both the self-tracking and quality literature, as well as absent definitions in the analyzed studies. The broader field of SCPs especially suffers from a shifting terminology, a common problem in information systems literature in general [40]. A pre-test of the search terms and search string, as well as knowledge of previous research, was used to minimize the amount of missing results in the search. Discursive collaboration in the extraction and synthesis processes was used to map terms and mitigate this threat.

Other threats to validity lie in only using digital databases and possible biases in the primary studies. Both aspects must be addressed in future research; for example, by using forward and backward snowballing, as proposed by [26]. Due to missing resources, this step could not be included in the current research.

5. Conclusion and outlook

The goal of this paper was to identify factors that affect the quality perceptions of stakeholders within a wearable ecosystem. To achieve this purpose, we used an SLR based on a pre-defined protocol. The final set used for data extraction comprised 19 studies. Based on a comprehensive process, we extracted 98 synonym- and duplicate-free quality-influencing factors and clustered them based on similarity on a map.

We also identified four stakeholder groups (consumers, medical professionals, manufacturers, and developers) in the literature and showed which of the 98 quality factors are relevant for them. Most of our identified factors are consumer-oriented, which shows that research seems to be more profound in that area. Regarding the coverage of the elements of the ecosystem, most factors relate to either the app or the ecosystem as a whole. The comparison with the ISO 25000 standard showed that, despite multiple matches, user-oriented factors and more technical and domain-oriented fields are not present in the standard.

Limitations of this study, as explained in the previous section, should be addressed by future research. Additionally, the overview of the factors also does not include the relationships between them or their possible measurements. Thus, future research should aim to provide more insight into this topic, either through additional empirical research or through comparisons with existing models and standards other than the ISO 25000.

Future research should also consider the emerging possibilities of ongoing quality management that SCPs offer by enabling “continuous monitoring of real-world performance data, allowing companies to identify and address design problems that testing failed to expose” [36].

6. References

[1] Alam, K.A., R. Ahmad, A. Akhunzada, M.H.N.M. Nasir, and S.U. Khan, "Impact analysis and change propagation in service-oriented enterprises: A systematic review", Information Systems, 54, 2015, pp. 43–73.

[2] Alnsour, Y., B. Hazarika, and J. Khuntia, "The Role of Effectiveness, Appeal and Functionality on Evaluation of Health Apps", in AMCIS 2016 Proceedings, AMCIS 2016 - 22nd Americas Conference on Information Systems. 2016.

[3] Becker, M., A. Kolbeck, C. Matt, and T. Hess, "Understanding the Continuous Use of Fitness Trackers: A Thematic Analysis", in Proceedings of the 21th Pacific Asia Conference on Information Systems (PACIS 2017), 21st Pacific-Asia Conference On Information Systems (PACIS 2017), Langkawi Island, Malaysia, 16-20 July 2017. 2017.

[4] Boell, S.K. and D. Cecez-Kecmanovic, "On being ‘systematic’ in literature reviews in IS", Journal of Information Technology, 30(2), 2015, pp. 161–173.

[5] Bruns, K. and F. Jacob, "Value-in-Use and Mobile Technologies", Business & Information Systems Engineering, 6(6), 2014, pp. 349–359.

[6] Calvaresi, D., D. Cesarini, P. Sernani, M. Marinoni, A.F. Dragoni, and A. Sturm, "Exploring the ambient assisted living domain: A systematic review", Journal of Ambient Intelligence and Humanized Computing, 8(2), 2017, pp. 239–257.

[7] Carroll, N. and I. Richardson, "Software-as-a-Medical Device: Demystifying Connected Health regulations", Journal of Systems and Information Technology, 18(2), 2016, pp. 186–215.

[8] Carvalho, R.M., R.M. de Castro Andrade, K.M. de Oliveira, I. de Sousa Santos, and C.I.M. Bezerra, "Quality characteristics and measures for human–computer interaction evaluation in ubiquitous systems", Software Quality Journal, 2(2), 2016, pp. 743–795.

[9] De Moya, J.-F. and J. Pallud, "Quantified Self: A Literature Review Based on the Funnel Paradigm", in Proceedings of the 25th European Conference on Information Systems (ECIS), European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017. 2017.

[10] DIN, "Qualitätsmanagementsysteme – Grundlagen und Begriffe (ISO 9000:2015); Deutsche und Englische Fassung EN ISO 9000:2015", 9000th edn., Deutsches Institut für Normung e. V., Beuth, Berlin, 01.040.03(9000), 2015.

[11] Dunn, B.K., M.L. Jensen, and R. Ralston, "Eye of the Blamestorm: An Exploration of User Blame Assessment within Compound Digital Platforms", in SIGHCI 2016 Proceedings, SIGHCI 2016 - 15th Annual Pre-ICIS Workshop on HCI Research in MIS. 2016.

[12] Dybå, T. and T. Dingsøyr, "Empirical studies of agile software development: A systematic review", Information and Software Technology, 50(9-10), 2008, pp. 833–859.

[13] Fawcett, T., "Mining the Quantified Self: Personal Knowledge Discovery as a Challenge for Data Science", Big data, 3(4), 2015, pp. 249–266.

[14] Fleisch, E., M. Weinberger, and F. Wortmann, "Geschäftsmodelle im Internet der Dinge", Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung (zfbf), 67(4), 2015, pp. 444–465.

[15] Gao, Y., H. Li, and Y. Luo, "An empirical study of wearable technology acceptance in healthcare", Industrial Management & Data Systems, 115(9), 2015, pp. 1704–1723.

[16] Gimpel, H., M. Nißen, and R.A. Görlitz, "Quantifying the Quantified Self: A Study on the Motivation of Patients to Track Their Own Health", in Proceedings of the 34th International Conference on Information Systems (ICIS 2013), International Conference on Information Systems, Milan, Italy, 15.12. - 18.12.2013. 2013.

[17] Hazarika, B., J. Karimi, J. Khuntia, and M. Parthasarathy, "Technology Frustration and Consumer Valuation Shift for Mobile Apps: An Exploratory Study", in AMCIS 2015 Proceedings, AMCIS 2015 - 21st Americas Conference on Information Systems. 2015.

[18] Hsiao, C.-H., M.-C. Chen, and K.-Y. Tang, "Investigating the Success Factors for the Acceptance of Mobile Healthcare Technology", in PACIS 2013 Proceedings. 2013.

[19] Idri, A., M. Bachiri, and J.L. Fernandez-Aleman, "A Framework for Evaluating the Software Product Quality of Pregnancy Monitoring Mobile Personal Health Records", Journal of Medical Systems, 40(3), 2016, p. 50.

[20] ISO/IEC, "Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model", 25012nd edn.(25012), 2008-12-15.

[21] ISO/IEC, "Systems and software engineering - Systens and software Quality Requriements and Evaluation (SQuaRE) - System and software quality models (25010:2011)", 25010th edn., Switzerland(25010), 2011-03-01.

[22] ISO/IEC, "Systems and software engineering - Systems and Software Quality Requriements and Evaluation (SQuaRE) - Guide to SQuaRE (25000:2014)", 25000th edn., Switzerland(25000), 2014-03-15.

[23] ISO/TS, "Information technology - Systems and software quality requirements and evaluation (SQuaRE) - Service quality models", 25011st edn., Switzerland(25011), 2017-06.

[24] Jain, R. and L. Jalali, "Objective Self", IEEE Multimedia, 21(4), 2014, pp. 100–110.

[25] Kitchenham, B. and S. Charters, Guidelines for performing Systematic Literature Reviews in Software Engineering: Version 2.3, 2007.

[26] Levy, Y. and T.J. Ellis, "A systems approach to conduct an effective literature review in support of information systems research", Informing Science: International Journal of an Emerging Transdiscipline, 9(1), 2006, pp. 181–212.

[27] Lundell, J. and C. Bates, "Understanding User Experience Journeys for a Smart Watch Device", in HCI in Business, Government, and Organizations: Information Systems: Third International Conference, HCIBGO 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17-22, 2016, Proceedings, Part II, F.F.-H. Nah and C.-H. Tan, Editors. 2016. Springer International Publishing: Cham.

[28] Martinez-Perez, B., I. de La Torre-Diez, S. Candelas-Plasencia, and M. Lopez-Coronado, "Development and evaluation of tools for measuring the quality of experience (QoE) in mHealth applications", Journal of Medical Systems, 37(5), 2013, p. 9976.

[29] Meulendijk, M., E. Meulendijks, P. Jansen, M. Numans, and M. Spruit, "What concerns users of medical apps? Exploring non-funcitonal requirements of medical mobile applications", in ECIS 2014 Proceedings, ECIS 2014 - 22nd European Conference on Information Systems. 2014.

[30] Moilanen, P., M. Salo, and L. Frank, "Inhibitors, enablers and social side winds Explaining the use of exercise tracking systems", in BLED 2014 Proceedings, BLED 2014 - 27th Bled eConference: eEcosystems. 2014.

[31] Neuhuettler, J., W. Ganz, and J. Liu, "An Integrated Approach for Measuring and Managing Quality of Smart Senior Care Services", in Advances in The Human Side of Service Engineering: Proceedings of the AHFE 2016 International Conference on The Human Side of Service Engineering, July 27-31, 2016, Walt Disney World®, Florida, USA, T.Z. Ahram and W. Karwowski, Editors. 2017. Springer International Publishing: Cham.

[32] Novales, A., M. Mocker, and D. Simonovich, "IT-enriched "Digitized" Products: Building Blocks and Challenges Diego, CA, USA, August 11-14, 2016", in Proceedings of the 22nd Americas Conference on Information Systems (AMCIS 2016), AMCIS 2016, San Diego, CA, USA,, August 11-14, 2016. 2016. Association for Information Systems.

[33] Peischl, B., M. Ferk, and A. Holzinger, "The fine art of user-centered software development", Software Quality Journal, 23(3), 2015, pp. 509–536.

[34] Petersen, K. and C. Gencel, "Worldviews, Research Methods, and their Relationship to Validity in Empirical Software Engineering Research", in 2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement, 2013 Joint Conference of the 23nd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement (IWSM-MENSURA), Ankara, Turkey, 23.10.2013 - 26.10.2013. 2013. IEEE.

[35] Porter, M.E. and J.E. Heppelmann, "How Smart, Connected Products are Transforming Competition", Harvard Business Review, 92(11), 2014, pp. 64–88.

[36] Porter, M.E. and J.E. Heppelmann, "How Smart, Connected Products Are Transforming Companies", Harvard Business Review, 93(10), 2015, pp. 96–114.

[37] Poslad, S., Ubiquitous Computing: Smart Devices, Environments and Interactions, 2nd edn., John Wiley & Sons Ltd, Hoboken, 2009.

[38] PwC, The Wearable Life 2.0: Connected living in a wearable world, Consumer Intelligence Series, PwC, 2016.

[39] Rooksby, J., M. Rost, A. Morrison, M. Chalmers, and M.C. Chalmers, "Personal Tracking as Lived Informatics", in CHI '14 - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 32nd Annual ACM Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada, April 26 - May 01, 2014. 2014. Assoc. for Computing Machinery: New York, USA.

[40] Schryen, G., A. Benlian, F. Rowe, S. Gregor, K. Larsen, S. Petter, G. Paré, G. Wagner, S. Haag, and E. Yasasin, "Literature Reviews in IS Research: What Can Be Learnt from the Past and Other Fields?", Communications of the Association for Information Systems, 41(1), 2017, pp. 759–774.

[41] Shih, P.C., K. Han, E.S. Poole, M.B. Rosson, and J.M. Carroll, "Use and Adoption Challenges of Wearable Activity Trackers", in iConference 2015 Proceedings, iConference, Newport Beach, CA, USA. 2015.

[42] Simons, L.P.A., J.F. Hampe, and N.A. Guldemond, "Designing healthy living support: Mobile applications added to hybrid (e)Coach solution", Health and Technology, 3(1), 2013, pp. 85–95.

[43] Suryadi, D. and H.M. Kim, "Identifying Sentiment-Dependent Product Features from Online Reviews", in Design Computing and Cognition ’16, J.S. Gero, Editor. 2017. Springer International Publishing: Cham.

[44] Swan, M., "Sensor Mania!: The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0", Journal of Sensor and Actuator Networks, 1(3), 2012, pp. 217–253.

[45] Vom Brocke, J., A. Simons, K. Riemer, B. Niehaves, and R. Plattfaut, "Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research", Communications of the Association for Information Systems(37), 2015, pp. 206–224.

[46] Zapata, B.C., J.L. Fernandez-Aleman, A. Idri, and A. Toval, "Empirical studies on usability of mHealth apps: A systematic literature review", Journal of Medical Systems, 39(2), 2015, p. 1.