Title:

Not all Sensors are Created Equal: A Framework for Evaluating Human Performance Measurement Technologies

Authors:

***Brian Caulfield** (Corresponding Author), The Insight Centre for Data Analytics, O'Brien Centre for Science, University College Dublin, Dublin 4, Ireland and Applied Research for Connected Health, NexusUCD, Belfield Office Park Block 9/10, Clonskeagh, Dublin 4, Ireland. Email: b.caulfield@ucd.ie Telephone: 01 7166515.

***Brenda Reginatto,** Applied Research for Connected Health, NexusUCD, Belfield Office Park Block 9/10, Clonskeagh, Dublin 4, Ireland

***Patrick Slevin**, The Insight Centre for Data Analytics, O'Brien Centre for Science, University College Dublin, Dublin 4, Ireland.

*These authors contributed equally to this work.

Abstract

Recent years have witnessed an explosion in the number of wearable sensing devices and associated apps that target a wide range of biomedical metrics, from actigraphy to glucose monitoring to lung function. This offers big opportunities for achieving scale in the use of such devices in application contexts such as telehealth, human performance and behaviour research and digitally enabled clinical trials. However, this increased availability and choice of sensors also brings with it a great challenge in optimising the match between the sensor and a specific application context. There is a need for a structured approach to first refining the requirements for a specific application, and then evaluating the available devices against those requirements. In this paper we will outline the main features of such an evaluation framework that has been developed with input from stakeholders in academic, clinical and industry settings.

1. Introduction

The market availability of digital devices that measure different aspects of human performance and behaviour has significantly increased in recent years. Human performance and behaviour measurement technology refers to consumer and medical grade health and wellbeing devices across a number of fields such as wearable, digital health and remote monitoring technologies. It is estimated that the number of connected wearable devices worldwide will increase from 325 million in 2016 to 929 million by 2021 [1]. Similarly, the digital health consumer base is growing in tandem, it is forecasted that by 2021, the number of people availing of remote monitoring programmes will grow to 52 million globally [2]. Although the increased availability of such devices is leading to greater research and commercial opportunity, it can also create significant confusion, especially for professionals who are attempting to select appropriate technologies that meet the requirements of their specific application, whether it is clinical trial, a research study, or a digital health service. To the authors' knowledge, there are currently no standardised methods to help professionals identify, evaluate and compare the numerous human performance devices available with respect to their

specific application requirements. In the absence of such a method, several issues exist for professionals who are undertaking device evaluations.

The first of these issues is the need for a tool that helps professionals identify devices that satisfy their application requirements. In many cases, when technologies are chosen and later evaluated, it is often not the device that emerges as the problem per se. It is that the device was, at the time of selection, not appropriate given the specific needs and requirements of the service provider and/or the user. Therefore, to address such an outcome, the authors would argue, that the application requirements should be the driver of the process (i.e. device identification, evaluation and comparison). This creates a fresh emphasis for the professional to understand the nuances of their specific application.

Though there are fuzzy boundaries between them, it is useful to consider three primary application contexts for human performance devices: Wellness/Fitness; Healthcare; and Clinical Trials/Research each with different use cases, depending on the primary motivation for use (Figure 1). Each application will have their own particular set of requirements to consider when deciding upon the type of device to deploy there within.

To help demonstrate the diverging requirements that can exist between applications in relation to a specific device, Figure 2 compares the potential high-level requirements for deploying a wearable activity tracker as part of an employee wellness programme versus those of a clinical trial endpoint.

It is evident from this example that the requirements of an application can be more nuanced and complex than one might imagine. The requirements appear similar in both use cases, yet even at this high level there are some critical differences and even more would be likely to emerge on a detailed analysis of the discrete requirements for each use case. This raises an issue for professionals while attempting to choose an appropriate device due to the difficulty of accounting for the plethora of requirements within an application context. For example, professionals may not be familiar with establishing a set of requirements? Or more importantly, they may have questions about the various

criteria that are important to define when choosing a device? To help address questions like these, the authors would argue that a process to thoroughly guide professionals through the definition of their application requirements could decrease the risk of selecting a device that does not fully account for the needs of the service or user, and are, therefore, not fit-for-purpose.

Another issue facing professionals in this space is the lack of a holistic tool for evaluating human performance devices. Several tools are available to help professionals evaluate digital health technologies but these tools are heavily biased towards measuring human factors criteria. A reason for this is that many of these tools have been developed within the discipline of human-computer interaction where the evaluation of user-interfaces associated with web applications and mobile technologies are a core focus [3–9]. Elsewhere, tools have also been developed to evaluate the acceptance, user-experience and usability of both the hardware and the software aspects of digital systems, products and services [10–15]. However, while these tools can be useful, there is a need to also evaluate aspects such as regulatory compliance, technical specifications and capabilities and scientific evidence supporting the use of a given device. These evaluation domains can be particularly relevant in highly regulated applications such as clinical trials. Once more, the availability of a holistic evaluation tool, which takes all such aspects in consideration, could support professionals to determine more effectively whether a device is indeed fit-for-purpose, according to their specific application requirements.

A final issue that should be highlighted is the lack of a tool to evaluate human performance devices prior to their implementation. As mentioned above, the available evaluation tools are primarily focused on measuring human factors criteria. Because of their nature, as tools that are focused on the outcome of a person's interaction with a product, they are frequently administered post-implementation. Yet, it is not until a post-implementation evaluation is conducted, that the devices' appropriateness to the service provider and the end-user is discovered, at which time the device could emerge as not fit-for-purpose. Long before this point however, a decision to invest in a device,

100 or several devices, was made. Such scenarios illustrate that an opportunity exists for a tool that can
101 help mitigate the risk of spending resources on devices that are not appropriate, by extending the
102 evaluation process to the pre-purchase phase where discrete devices are identified and evaluated
103 against the application requirements so that the most appropriate device can be selected in a
104 systematic and informed manner.

105 The aim of this paper is to address the gaps highlighted above by describing a framework for
106 evaluating human performance technologies. The framework guides professionals through the
107 processes of defining application requirements, searching for and selecting candidate devices, and
108 finally performing a structured evaluation of these devices against application requirements – all
109 with a view to helping them determine if a device is fit-for-purpose and worthy of field testing based
110 on their specific requirements. Whether these requirements are in the context of a clinical trial, a
111 pilot study, or a digital health service, the outcome should reflect a systematic and rigorous
112 evaluation.

113 2. Results

114 The evaluation framework follows a three-step process, (1) Requirements Definition, (2) Device
115 Search and (3) Device Evaluation (Figure 3). Each step of this framework is supported by relevant
116 templates, which guide the user through the process and to allow for the clear documentation of the
117 rationale for their choice. In this regard the user is defined as the person/group responsible for
118 selecting the device for deployment in the specified application. Though it is recommended that the
119 framework be employed in a systematic manner, the steps could be applied in sequence or users
120 could elect to apply isolated elements of the framework if constrained by resources and time. For
121 example, there may be situations where one or more devices of interest have already been
122 identified as part of an ad-hoc process. In this case, the user could complete step one (Requirements
123 Definition), skip step two (Device Search) and proceed to step three (Device Evaluation) to
124 determine which of the pre-selected devices is most fit-for-purpose and / or worthy of field testing,

according to their specific requirements. On the other hand, some users may not have the resources or time to enable completion of a formal field-testing phase, and therefore, this could limit the application of the framework to a desk-based evaluation of identified devices.

Step 1 – Requirements Definition

Defining the application requirements at the beginning of the process enables the user to conduct a more systematic and efficient device search and evaluation. The template provided by the framework guides the user through this process, prompting the consideration of different aspects including: application description and goals, device requirements (e.g. what data needs to be collected through the device) and user profile (i.e. who are the people expected to use the device and any specific design requirements they may have). Other aspects to reflect upon include: budget, setting (e.g. home, hospital), geographical location (e.g. urban or rural area), technical requirements (e.g. operational system preferences, compatibility with other equipment and connectivity requirements) and any ethical dilemmas associated with the use of the device (e.g. users are part of a vulnerable population or device is likely to place undue burden or stress on users). Figure 4 offers an excerpt from the requirements definition template.

Finally, the user is encouraged to categorize requirements as essential or secondary, according to how critical they are to the achievement of the application goals. This is important because it helps users remain grounded in those aspects that are most important, which can be often challenging when evaluating and comparing devices that offer multiple features and functionalities. Additionally, the essential requirements form the basis for the device search strategy, as described below.

It is important to note, that the extent and intricacy of the requirements list is at the discretion of the user. A more intricate requirements list will, in general, reduce the pool of devices unearthed in the search, while a high-level requirements list will, in general, broaden the scope of the devices identified.

Step 2 – Device Search

The second step of the framework aims to help the user identify available devices that match their essential application requirements in an efficient and yet comprehensive manner. Firstly, the user is guided through the process of generating keywords based upon the essential application requirements and using such keywords to conduct a systematic web search. Several recommendations are also provided on how to optimise the search, for example, by using particular words or symbols to widen or restrict results, and reviewing the search engine settings to avoid biased results (e.g. based on the user's location or previous search history).

The user is then prompted to use a comparison matrix template to shortlist devices worthy of a comprehensive desk-based evaluation. It is recommended that only those devices which satisfy all essential requirements are taken to the third step of the framework (Device Evaluation). Figure 5 presents an example of the comparison matrix. In this case, only Devices 2 and 3 satisfy all essential requirements and are deemed worthy of a comprehensive desk-based evaluation.

Step 3 – Device Evaluation

The third step of the framework allows the user to conduct a comprehensive desk-based device evaluation and determine whether one or more devices are worthy of field testing. The template provided prompts the user to answer a number of questions and scrutinize each device according to six domains: 1) Background Information; 2) Cost and Supply Information; 3) Regulatory Compliance; 4) Scientific Evidence; 5) Technical Evaluation and 6) Human Factors. Figure 6 offers an excerpt of the device evaluation template for illustrative purposes. A description of each domain is discussed below.

Background Information

The user is prompted to gather background information on the company supplying the device. This may include, for example, information on the size of the organisation, number of years they have been operating, where the company is based and whether they have experienced any product

recalls in the past. The goal of this section is to give the user a sense of trust in the company behind the device and clarify whether they possess the required infrastructure to support the use of the device for the purpose specified by the user. Such knowledge can be of critical importance if the application requires a steady supply of a large number of devices.

Cost and Supply Information

This section allows the user to determine whether the device is affordable and available. It covers the costs of the device (including the need for additional or recurrent purchases, shipping fees, or technical support subscription charges), as well as relevant supply information, such as availability in the target country, minimum order requirements and the possibility of obtaining a free sample.

Regulatory Compliance

This domain requires the user to consider whether the device evaluated complies with relevant regulatory standards, with due regard to the territory or territories in which the device will be deployed. This includes not only safety and performance standards, but also data protection regulation applicable to the target location where the device will be used.

Scientific Evidence

The user is encouraged to examine the scientific evidence supporting the intended use of the device. This includes, but is not limited to, evidence demonstrating the validity and accuracy of the device's target measurement in comparison to the gold standard, data quality under field conditions (as opposed to highly controlled, in-lab environments), clinical safety and performance, technical feasibility and usability. Where the intended use stated by the manufacturer differs from the user application, it is important to investigate whether there is evidence supporting the latter. For instance, if the user wishes to use an activity tracker originally designed for athletes with a cohort of geriatric patients, it would be important to determine whether there is any data published on the use of the device by older people.

Technical Evaluation

This domain scrutinizes the device's technical specifications and capabilities. The intention is to give the user a deep understanding of how the device operates and what technical infrastructure may be required. Examples of sub-sections within this domain include device dimensions, battery life and charging methods, calibration requirements, operational system compatibility, connectivity requirements (e.g. wired, Wi-Fi, BT), data access and storage (e.g. is it possible to access raw data from the device? Where is the data stored?), data security (e.g. how is user data protected?) and data visualisation (e.g. does the device provide feedback? In this case, where is it displayed?).

Human Factors

The final domain of the evaluation template relates to the device usability and other human factors. The questions in this section help the user examine the level of end-user interaction required, as well as any obvious design issues, which may hinder usability and user experience. Other aspects to consider include the device material (e.g. is it washable? Is it durable? Could it cause allergy or skin irritation?) and the quality of educational materials provided.

Data Gathering and Interpretation

The user may refer to a variety of sources to obtain the information required to complete the device evaluation process. These might include the supplier's website, news outlets, blogs, scientific journals, discussion forums and communication with the supplier. Where information regarding a query cannot be located, it is recommended that this is clearly stated (e.g. 'information not found'), instead of making assumptions around the device features and capabilities. Documenting the access date and source of information is also highly encouraged as there may be discrepancies depending on when and where information is garnered (e.g. different resellers may offer different prices). Documenting the information source is also particularly beneficial when revisiting the decision-making process in the future.

Once all queries have been answered, the columns on the right-hand side of the template prompt the user to compare the devices. This can be seen on the 'Requirements Fulfilled?' column presented in Figure 6. For each query the user should try to determine whether the relevant application requirement is fulfilled for each device. It is recommended to clearly document if the query is not relevant to the user requirements or if further information is required to finalise a decision. Once more, this is beneficial when revisiting the decision-making process in the future.

By comparing how well the devices satisfy the application requirements under each domain, the user should be in a much more informed position to determine which device or devices are worthy of field testing. It is recommended to clearly document the rationale for the decision made, as well as any specific areas that require further investigation through field testing. In the case where a conflict between two devices emerges, and the user is satisfied that they have obtained all the information they can to help inform their desk-based evaluation, it is recommended that the user field tests both devices to determine the most fit-for-purpose device.

3. Discussion

The evaluation framework presented in this paper was developed as a collaboration between academic, industry and clinical stakeholders to address the lack of an existing structured approach to help professionals evaluate human performance technologies. The framework provides a comprehensive tool which enables the user to define their specific requirements, conduct a systematic web search and complete a holistic desk-based evaluation, to determine whether one or more devices are fit-for-purpose and / or worthy of field testing.

The first two steps of the framework, Requirements Definition and Device Search, are unique in comparison to existing resources. To the authors' knowledge, this is the first tool to prompt users to thoroughly reflect upon and prioritise their requirements prior to selecting a device. It is believed that this will enable users to conduct a more efficient search and grounded evaluation, decreasing

247 the risk of selecting devices that fail to fully account for the specific needs of their application.
248 Similarly, no other resources have been found to support professionals in conducting a systematic
249 web search to identify devices that match such requirements.

250 While existing tools may help users evaluate specific aspects of a digital health device, these
251 resources are not conducive of a holistic evaluation. The six domains presented on the third step of
252 the framework, address this gap by allowing users to conduct a comprehensive desk-based
253 evaluation regardless of their own area of expertise. It is also expected that this exercise may help
254 users identify areas where they require specialist input to help them decide whether a device
255 matches their application requirements.

256 Finally, it is important to note that the desk-based evaluation process described in this paper is not
257 expected to replace the need for field testing of selected devices. It is, however, believed that it will
258 greatly help users identify critical issues in a timely manner, i.e. before significant time or resources
259 are spent on implementing devices that are not fit-for-purpose. This offers a significant advantage
260 over existing resources, which mainly focus on evaluating devices post-implementation.

261 ## 4. Methods

262 The three-step evaluation process outlined above was developed using an iterative participatory
263 design approach, as described by Simonsen and Hertzum [16]. This is a hybrid design approach that
264 emphasises the involvement of potential future end-user's expertise and experiences primarily for
265 the design of technologies, businesses and social innovative products and services [16,17]. Moreover, as
266 well as being an inclusive design process it is also iterative, where researchers and potential future
267 end-users work collaboratively to discover, explain, reflect and integrate knowledge at various time-
268 points in the process to aid in the productive development of the design [18].

269 The approach was felt to be most appropriate considering the cross-collaborative nature of the
270 research which required the input of various types of expertise in the health technology field, from

271 both research and industry, throughout the design process of the evaluation framework. Two
272 industry collaborators were involved at different points throughout the process. In both cases,
273 selection of human performance measurement technologies was a critical issue for their business,
274 one company being involved in tele-health service provision, and the other being a clinical research
275 organisation. Figure 7 illustrates the main project stages including the point at which each
276 stakeholder group were involved.

277 Defining the Device Search and Requirements Definition Processes

278 The initial phase of the project aimed to define the Device Search Process (Step 2 in Results). Early
279 iterations of this process were trialled by the researchers (BR and PS) using two discrete smart blood
280 pressure monitors as the focus for the desk-based web search. These searches were unstructured
281 but did entail the formulation of keyword searches. When all keyword sequences were saturated,
282 the researchers reconvened to critically evaluate the process used.

283 The core aspect to emerge from this early work, was that though the process made sense in terms of
284 formulating keyword sequences to identify potential devices, without the requirements of a specific
285 use case, the web search findings were too expansive. For instance, a plethora of smart blood
286 pressure monitors were identified but without contextual information such as, a set budget per
287 device, there was no early mechanism to filter down the large number of devices garnered from the
288 web search.

289 Through further consensus, the researchers decided that hypothetical applications with specific
290 requirements should be developed first, one academic in nature and one from a health technology
291 industry perspective. Doing so would allow the researchers to assess the flexibility of the tool in
292 relation to the diverse needs of potential end users. Crucially, the specific application requirements
293 would help focus the identification process while providing the device evaluation with a more
294 purposeful direction. A hypothetical academic application, concerning a diabetes self-monitoring

study requiring the use of a smart glucometer, was then developed, forming the basis for the Requirements Definition Process (Step 1 in Results).

Industry Workshop 1

A one-hour workshop with the two industry collaborators was then arranged with the aim of (a) gaining feedback regarding the overall project design and trajectory and (b) defining an industry-led hypothetical application. During the workshop, the collaborators were presented with the hypothetical academic application use case. Based on this example, they were then given the task of developing an industry-led application use case that reflected the requirements of a provider style study. The primary outcome of the workshop was the development of industry-led hypothetical use case namely a medication adherence programme, requiring the use of a smart pill adherence tracking device

Expert Consultation 1: Specification of Device Evaluation Template

The aim of this phase was to develop the specifications for the device evaluation template. To align with the iterative participatory approach, the research centre's existing in-house healthcare technology expertise was leveraged to identify the specifications. To provide the basis for the feedback sessions with the expert group (n=7), the authors developed an alpha version of the device evaluation template, using a list of device evaluation criteria that had originally been used by one of them (BC) as a teaching tool.

Each expert was invited to participate in a 30-minute brainstorming session with researchers (BR and PS). The experts came from a variety of digital health backgrounds including biomedical, software and systems engineering, human factors, regulatory, clinical and digital health expertise. Each expert was provided with a copy of the alpha version at the beginning of the session. They were asked for their feedback regarding the domain content in the context of using it to evaluate a digital health and wellbeing device. Notes were taken by the researchers regarding the relevant points made by each expert. Upon completing the feedback sessions, the notes gathered were collated and

examined for patterns. As patterns were identified, they were cross checked for consistency and compiled into iteration additions. The researchers then refined the domains and the alpha template was subjected to its first iteration.

Expert Consultation 2: Domain Question Development

The aim of this phase was to develop the questions within each domain. To ensure these iterations complied with initial feedback and comments, the same experts (n=7) were invited to partake in a follow-up 30-minute feedback session with two researchers (BR and PS). The experts received a copy of the iterated template at the beginning of the session and asked for their feedback regarding the domains and questions there within. As the expert critiqued the iterated template, notes were taken, and afterwards collated and examined for patterns. If patterns were identified, they were cross checked for consistency and compiled into iteration additions. The insights gathered informed the final iteration. The researchers now had a beta version of the template ready for testing.

Pre-Evaluation Template Testing

In this phase, the researchers (BR and PS) aimed to define the devices, per hypothetical use case, that would be allocated to the external researchers for testing the device evaluation template. To explain how the devices were chosen for the testing phase, it will be instructive to provide an example using the smart pill box hypothetical use case. The researchers followed the first two steps of the process: Requirements Definition and Device Search. Leveraging the requirements defined by the industry collaborators (Figure 8), the researchers conducted the device web search based on these criteria:

smart or connected device; ability to track medication (pill / tablet) adherence; portable device; offline use enabled (i.e. store and forward); potential ability for other services to access; monitoring data (near) real time; currently available for purchase; distributed in Ireland; and compliance with EU regulations (CE marking, EU Data Protection Directive).

Using the Google search engine, a search was conducted to identify potentially suitable devices. The researchers went 4 pages deep into the Google search engine (10 results displayed per page) for each keyword permutation. In each case, the first 40 results were examined to search for relevant devices. Once new search combinations were not yielding any new device search information within the 40 results, the researchers concluded that saturation had been reached and ceased the search. The following keywords were used: smart / connected / wireless / Bluetooth; monitor / track; medication / pill / tablet; and adherence / compliance.

A number of keyword permutation combinations were tested. One researcher (BR) tested combinations using the keyword 'monitor' (e.g. smart + monitor + medication + adherence + device) while another researcher (PS) tested combinations containing the word 'track' (e.g. smart + track + medication + adherence + device).

The following combinations retrieved the largest number of new results: (smart + monitor + medication + adherence + device); (smart + monitor + pill + adherence + device; wireless + monitor + medication + adherence + device); (Bluetooth + monitor + medication + adherence + device); and (Bluetooth + monitor + medication + compliance + device). A record was not kept of the number of unique devices that were found using each search combinations.

From this search, 21 devices were initially identified. All smartphone apps, connected blister packs and smart ingestible pills, totalling 5, were excluded since they were outside the scope of this medication adherence programme. A further 2 devices were excluded because they were not focused on pill or tablet adherence monitoring. Another 2 devices were excluded because they were bound to a service that did not allow integration with external services. The remaining 12 devices were subjected to the requirements comparison matrix. Based on this comparison, 9 devices were excluded because they didn't meet one or more criteria or not enough information was available despite contacting manufacturing company and three devices were shortlisted as suitable to allocate to the evaluation template testing phase as seen in Figure 9.

The same process was followed by researchers to determine the devices to be allocated to the those testing the evaluation template based on the smart glucometer hypothetical use case. In total, 10 devices were identified of which 2 were found to be suitable for the testing phase.

Evaluation Template Testing

For this phase, the aim was to finalise the device evaluation template in terms of its usability and general experience. To ensure objectivity, external researchers (n=5) from a range of digital health and wellbeing backgrounds, other than those used in the specification of the device evaluation template, were recruited to test the beta version. Both hypothetical application use cases were tested; two participants were allocated the smart glucometer academic use case, and three were allocated the smart pill-box industry use case. Each participant was emailed a copy of the beta version template, a copy of the application requirements plus an outline of the devices to be evaluated. No parameters were defined for the testing other than to test the device evaluation template using the devices allocated to them. The testing was completed at their convenience.

Upon completion, comments were received via email from the participants in relation to the usability, user-experience and perceived usefulness of the evaluation template. Their feedback was collated and examined for patterns. When patterns were identified, they were cross checked for consistency and compiled into iteration additions. These concluding insights informed the final iteration of the beta version device evaluation template.

Industry Workshop 2

A final one-hour workshop was conducted with the two industry collaborators. The aim was to present the framework and garner final feedback. Particularly, the authors wanted to explore if they felt (a) that the three-step framework process was a useful and relevant guide and (b) that the device evaluation template was flexible enough to meet their specific needs. Notes were taken and feedback was incorporated to the final version of the evaluation framework.

5. Acknowledgements

The authors would like to thank the centre for Applied Research for Connected Health (ARCH) for funding this work. They also thank ICON plc, S3 Group, Kinesis Health Technologies and colleagues at the Insight Centre for Data Analytics for their feedback and input into this project.

6. Competing Interests

The authors have no competing interests to declare.

7. Contributions

BR and PS contributed to study design, data collection and analysis and writing/reviewing the manuscript. BC supervised the project, contributed to study design, provided the list of device evaluation criteria which formed the basis for the alpha version of the device evaluation template and critically reviewing the manuscript. All authors have approved the final version of this manuscript and take accountability for all aspects of the manuscript.

8. Funding

This project received funding from Enterprise Ireland and the Irish Development Authority's joint initiative, the Technologies Centres programme. Grant Number: TC20140018

9. References

1. Statista. Connected wearable devices worldwide 2016-2021. (2018). Available

at: https://www.statista.com/statistics/487291/global-connected-wearable-devices/. (Accessed: 5th March 2018)

2. Berg Insights. *mHealth and Home Monitoring - 8th Edition*. (2017).

3. Chin, J. P., Diehl, V. a & Norman, L. K. Development of an instrument measuring user satisfaction of the human-computer interface. *CHI '88- Proc. SIGCHI Conf. Hum. factors Comput. Syst.* 213–218 (1988). doi:10.1145/57167.57203

4. Davis, F. D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q.* **13,** 319 (1989).

5. Nielsen, J. *Usability Engineering*. *Usability Engineering* **44,** (1993).

6. Lewis, J. R. Computer System Usability Questionnaire. *Int. J. Hum. Comput. Interact.* **7,** 57–78 (1995).

7. Brooke, J. SUS - A quick and dirty usability scale. *Usability Eval. Ind.* **189,** 4–7 (1996).

8. Lin, H. X., Choong, Y. Y. & Salvendy, G. A proposed index of usability: A method for comparing the relative usability of different software systems. *Behav. Inf. Technol.* **16,** 267–277 (1997).

9. Madsen, M. & Gregor, S. Measuring Human-Computer Trust. *Proc. Elev. Australas. Conf. Inf. Syst.* 6–8 (2000).

10. Demers, L., Weiss-lambrou, R. & Ska, B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2 . 0): An overview and recent progress. *Technol. Disabil.* **14,** 101–105 (2002).

11. Mugge, R. & Schoormans, J. P. L. A Longitudinal Study of Product Attachment and its Determinants. *Adv. Consum. Res.* **7,** 641–647 (2006).

12. Tedesco, D. P. & Tullis, T. S. A Comparison of Methods for Eliciting Post-Task Subjective Ratings in Usability Testing. *Upa 2006* 1–9 (2006).

13. Finstad, K. Interacting with Computers The Usability Metric for User Experience. *Interact. Comput.* **22,** 323–327 (2010).

14. Stoyanov, S. R. *et al.* Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. *JMIR mHealth uHealth* **3,** e27 (2015).

15. Hirani, S. P. *et al.* Quantifying beliefs regarding telehealth: Development of the Whole Systems Demonstrator Service User Technology Acceptability Questionnaire. *J. Telemed. Telecare* **23,** 460–469 (2017).

16. Simonsen, J. & Hertzum, M. in *Design Research: Synergies from Interdisciplinary Perspectives* 16–32 (2010). doi:10.4324/9780203855836

17. Manzini, E. & Rizzo, F. Small projects/large changes: Participatory design as an open participated process. *CoDesign* **7,** 199–215 (2011).

18. Kristensen, M., Kyng, M. & Palen, L. Participatory Design in Emergency Medical Service : Designing for Future Practice. *Chi 2006* 161–170 (2006). doi:10.1145/1124772.1124798

10. Figure Legends

Figure 1: Three Primary Application Contexts for Human Performance Devices

Figure 2: Example of High Level Requirements for Deploying a Wearable Tracker in Different Application Contexts

Figure 3: Evaluation Framework's Three Step Process

Figure 4: Excerpt from the Requirements Definition Template

Figure 5: Example of Comparison Matrix in Practice

Figure 6: Excerpt from the Device Evaluation Template

Figure 7: Main Project Stages and Stakeholder Input

Figure 8: Application Requirements

Figure 9: Device Shortlisting