Original Research

DIGITAL HEALTH

Digital Health
Volume 5: 1–21
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2055207619886463
journals.sagepub.com/home/dhj
SAGE

Development of the mHealth App Trustworthiness checklist

Afua van Haasteren, Felix Gille, Marta Fadda and Effy Vayena

Abstract

Background: Mobile health applications (mHealth apps) currently lack a consensus on substantial quality and safety standards. As such, the number of individuals engaging with untrustworthy mHealth apps continues to grow at a steady pace.

Objective: The purpose of this study was to investigate end-users' opinions on the features or actions necessary for trustworthy mHealth apps; and to convey this information to app developers via a succinct but informative checklist: the mHealth app trustworthiness checklist.

Methods: The checklist was formulated in three stages: (a) a literature review of studies identified the desirable features of the most prolific mHealth apps (health and fitness apps); (b) four focus group sessions with past or current users of these apps ($n = 20$); and (c) expert feedback on whether the checklist items are conceivable in a real-life setting ($n = 6$).

Results: Five major themes emerged from the focus group discussions: informational content, organizational attributes, societal influence, technology-related features, and user control factors. The mHealth app trustworthiness checklist was developed to incorporate these five themes and subsequently modified following expert consultation. In addition to the trustworthiness themes, we identified features that lie between trust and mistrust (limited digital literacy and indifference) as well as 10 features and actions that cause end-users to mistrust mHealth apps.

Conclusion: This study contributes to the evidence base on the attributes of trustworthy mHealth apps. The mHealth app trustworthiness checklist is a useful tool in advancing continued efforts to ensure that health technologies are trustworthy.

Keywords

Trustworthiness, trust, application developers, focus group, qualitative study, checklist

Submission date: 21 April 2019; Acceptance date: 3 August 2019

Introduction

Healthcare is increasingly engaging with mobile health (mHealth) and their accompanying software applications colloquially referred to as 'apps'. The push towards mHealth is largely driven by encouraging statistics showing that up to 58% of US mobile phone owners had downloaded some form of a health app as of 2015; and also, that there were around 325,000 health apps on the market in 2017.[1,2] Not only are these apps projected to remedy negative health outcomes and medical errors; but also they may improve access to care while mitigating rising healthcare costs.[3,4] A quick glance through either the Apple iTunes app store or the Google Play store, for instance, will reveal various mHealth apps (or simply health apps) with disparate claims and objectives.[5,6]

Department of Health Sciences and Technology (D-HEST), ETH Zürich, Switzerland

Corresponding author:
Effy Vayena, Department of Health Sciences and Technology (D-HEST), ETH Zürich, Health Ethics and Policy Lab, Auf der Mauer 17, 8092, Zürich, Switzerland.
Email: Effy.vayena@hest.ethz.ch
Twitter: @EffyVayena

There is ongoing debate about the relationship between current regulations governing mHealth apps and their overall clinical utility.[7] Due to concerns that stringent regulations can stifle innovation, however, regulatory bodies have embraced a laissez-faire approach to overseeing mHealth apps.[3,4,8] This relaxed regulatory climate has led to questionable – and at times dangerous – claims by some of these apps.[9] As such, it is crucial to carve out better measures to uphold the quality, safety, effectiveness and data security of mHealth apps.[1,4,10–15]

Several governing bodies are taking action to mitigate some of the safety and security issues facing mHealth apps. In the UK for instance, the National Health Service (NHS) has launched the NHS Apps Library.[16] This library mandates app developers to satisfy a strict set of criteria laid out in the Digital Assessment Questionnaire to have their products approved for the library.[17] This additional layer of oversight is intended to render an app safe enough for health providers to feel comfortable to recommend it to their patients. More recently, the National Institute for Health and Care Excellence (NICE) has released the Evidence Standards Framework for Digital Health Technologies with the aim of enlightening the developers of digital health technologies such as mHealth apps about standards in effectiveness and economic impact.[18]

A recent review of the literature on health apps revealed that researchers appear to analyse three key areas when assessing app quality.[1] First, the evidence base of the app; examining issues such as adherence to existing guidelines, as well as clinical and scientific relevance. Second, they examine the app from the perspective of the end-user analysing issues relating to usability, aesthetics and ease of use. The third approach involves scrutinising the trustworthiness of an app by examining factors such as transparency, privacy, data management, data protection and data reuse.[1] Although this current study accommodates all of these techniques, the issue of trustworthiness is of most importance. We focus on trustworthiness for a variety of reasons that we shall describe below.

To decipher what constitutes trustworthy mHealth apps, it is beneficial to define the concepts of trust and trustworthiness. Trust, is a relational phenomenon that signifies the willingness of one party to become vulnerable to another presumably competent, reliable and honest party in the hope of an optimistic outcome as a result of this relationship.[19,20] In the context of this research, we are interested in the trust relationship that occurs between an end-user and their mHealth app. In this instance, the end-user may trust a particular mHealth app to benefit their health on the basis that it can provide relevant feedback from analysing his/her data.

Trustworthiness on the other hand, refers to attributes that compel an individual to consider another individual or entity worthy of their trust. To be perceived as trustworthy, the entity must strategically signal their honesty, competence and reliability to make others comfortable to place their trust in them.[20] In the context of this study, an end-user is likely to consider mHealth apps which are capable of performing their intended duties – such as recording data – accurately as trustworthy. Trustworthiness judgements are often subjective and influenced by past experiences, upbringing, nationality or even race.[21] In a nutshell, trust is a relational concept between at least two parties, whereas trustworthiness is a characteristic of the trusted party.

Results from a 2015 survey evaluating the reasons why end-users adopt, abandon or continue to use mHealth apps, revealed mistrust of developers as a culprit.[22] This mistrust likely results from cynicism – or heightened suspicion – about the competence, honesty and reliability of app developers.[20,23] A viable option to improve trust in mHealth apps, therefore, may be to provide end-users with the necessary tools to flag up untrustworthy apps.[24] This approach of focusing on end-users, however, is likely to be affected by the same confounders that plague public understanding of information.[25] Thus, it may be prudent to prioritise app developers when tackling the issue of why mHealth apps are mistrusted.[9]

A plethora of tools in the form of checklists and scales have been developed to assist mHealth app users to assess certain attributes such as quality and effectiveness. For example, the World Health Organization developed the mHealth evidence reporting and assessment (mERA) checklist to standardise the reporting of mHealth interventions.[26] Similarly, a checklist approved by the United States Agency for International Development (USAID) seeks to allow mHealth program managers to assess and plan mHealth programs holistically by addressing security, privacy and confidentiality concerns.[27] In terms of mHealth scales, the Mobile App Rating Scale (MARS), for instance, has been developed to appraise app quality.[28]

The App Synopsis is one of the few checklists that focuses on evaluating the trustworthiness of mHealth apps. Its aim is to standardise the manner in which end-users, developers and distributors report on the functionality, information quality, rationale, validity, and reliability of health apps.[29,30] While the App Synopsis helps to examine the trustworthiness of health apps, it does not necessarily relay the concerns of end-users to app developers. Meanwhile, communicating end-users' concerns directly to app developers will place the developers in a much better position to create apps that end-users can trust.

Hence, in this current study, we seek to uncover the features or actions that encourage end-users to judge mHealth apps as trustworthy. Ultimately, we intend to convey this information to app developers by summarising the opinions of end-users into a succinct but informative mHealth App Trustworthiness (mHAT) checklist.

Methods

To commence this study, we selected the most prevalent type of health apps: physical activity apps.[22] The reasoning for doing so, was to establish a baseline of mHealth app characteristics that a large section of end-users were likely to be familiar with. The mHAT checklist was thus formulated in three stages: (a) a literature review of studies that shed light on the desirable features of health and fitness apps, (b) four focus group sessions with past or current users of physical activity or fitness apps, and (c) expert feedback on whether the checklist items are conceivable in a real-life setting. In the subsequent sections, we elaborate on each of these stages.

Literature review

Materials. To derive appropriate themes to steer the focus group discussions, we proceeded with a literature review to identify end-users' opinions about health and fitness apps. Specifically, we aimed to identify and summarise the features or characteristics that improved end-users' experience with such apps. Lessons from a previous review, conducted by some of the authors of this article, were useful in structuring the review process.[31]

Procedure. To derive search terms related to trust and trustworthiness, we developed and conducted test runs of relevant keywords, synonyms, medical subject heading (MeSH) terms in several electronic databases. These preliminary actions were meant to identify the terms and databases that would yield appropriate articles. After these test runs were completed, we searched the PubMed, Medline and Scopus electronic databases using the terms: (competen* OR accuracy OR positive intentions OR ethics) AND (trust* OR mistrust OR distrust OR credib* OR confiden* OR relia*) AND (physical activity app* or fitness app*). All of these searches took place in July 2018.

In total, we recovered 2280 article: 1925 articles from Medline, 73 from PubMed and 282 from Scopus. Two coders (AvH and MF) independently screened the titles and abstracts of the uncovered articles culminating with the full-texts of articles to be included in this study. Cohen's kappa – which provides the proportion of agreement between two coders factoring in chance during the title and abstract screening – was calculated to assess the level of agreement among the coders.

Analysis. The three main questions in the Critical Appraisal Skills Program (CASP) qualitative checklist were used to assess the quality of each study.[32] The primary questions that the CASP qualitative checklist attempts to answer are: (a) are the results valid? (b) what do the results actually reveal? and (c) is the study useful in this context? MF and AvH bore these questions in mind while assessing the eligible full-text articles. Articles that described the features and characteristics of health and fitness apps that enhance end-users' experience were included.

After agreeing on the full-text articles to be included in the review, the coders identified the themes and subthemes emanating from each of the articles through thematic analysis.[33] To do so, both coders independently reviewed the full-texts thoroughly to record and categorise all of the emergent themes. Additionally, the coders created a condensed version of the themes for the focus group discussions. Throughout this process, the coders conferred with each other about disagreements to reach a consensus. A third author (FG) was consulted to resolve any differences in opinion between the two coders.

Focus groups

Materials. Focus groups are a practical way of exploring the experiences, viewpoints and beliefs of participants on a particular topic.[34,35] This study employed the focus group methodology because we sought to identify end-users' perspectives on the features and characteristics of physical activity apps that make them trustworthy. To begin, we created open-ended questions from our knowledge of the literature and compiled a list of the condensed themes generated from the preliminary literature review. We also formulated a topic guide to ensure uniformity among each of the groups. Consent forms and a descriptive characteristics form were also created and distributed to the participants.

Procedure (recruitment and sampling). After obtaining ethical approval from the Ethics Commission of ETH Zurich, Switzerland, we recruited a purposive sample of 20 participants by distributing posters on the ETH Zurich campus. To qualify as a participant, individuals had to be past or present users of physical activity apps. The participants were allocated into four focus groups comprising of seven, six, four and three people respectively. Each focus group session lasted for about one hour and was conducted by two authors: AvH moderated the session while FG took detailed notes in an observing role.

At the beginning of each focus group, participants were informed that the session would be audiotaped and the consent and descriptive characteristics forms were distributed. After going through all of the questions on the topic guide, participants were asked to rank the relevance of the condensed themes derived from the literature review. The purpose of these rankings was not to generate quantitative data but, rather, to stimulate further discussions relevant for understanding participants' perspectives on the topic. After each focus group session, AvH and FG discussed their preliminary thoughts.

Analysis. The focus group discussions were analysed via content analysis. To do so, we followed the step-by-step process laid out by Taylor-Powell and Renner in 2003.[36] First, AvH transcribed all of the audio recordings. Following this, MF and AvH independently coded the data by reading the transcripts multiple times to familiarise themselves with their contents bearing in mind the aspects of the transcripts that answered our main question: what features and characteristics of health apps encourage end-users to perceive them as trustworthy?

The coding process was iterative as the coders identified patterns and themes both within and across the focus groups. The coders employed ethnographic analysis to draw out the main themes and subthemes by interpreting the focus group dialogues and presenting direct quotes to support each category.[34] To avoid simply identifying themes without pursuing their meanings within specific contexts, the coders ensured that they captured all of the differing opinions expressed by the participants. Whenever the two coders disagreed on an issue, a third author (FG) was consulted.

Checklist development and expert feedback

Materials. The themes generated from the focus groups were compiled into the mHAT checklist. Since the checklist aims to signpost app developers into creating trustworthy health apps, it was essential to obtain expert opinion on whether its items are feasible in a real-life setting. By 'expert', we mean individuals with in-depth knowledge and experience on a particular subject-matter.[37] The experts chosen for this study possessed either of these abilities: (a) highly skilled in the processes entailed in developing apps; or (b) knowledgeable in some form of a programming language such as Python or JavaScript which are commonly used as app development software.

Procedure. The checklist is made up of five sections to reflect the themes and subthemes derived from the focus group transcripts. Altogether, a purposive sample of six software engineers, computer scientists, and programmers provided written feedback on the checklist via email. Our initial email to the experts was sent out in November 2018 and comprised of the mHAT checklist, its intended purpose as well as a detailed layout of its methodology. To assist the experts in their critical analysis, we asked each one to assess the checklist from an app developer's perspective considering the following questions:

- Are the checklist contents feasible for an app developer?
- Can these questions help incorporate trust in an app? If not, what is missing?
- Are the questions straight to the point and self-explanatory?
- Do you have any suggestions on what to include to improve the checklist?

Analysis. Upon receiving all of the expert responses in January 2019, we proceeded to analyse them both individually and collectively. On the whole, there appeared to be a consensus among the experts on which checklist items had to be modified. Thus, we responded to the expert comments in a subsequent email. After two rounds of email correspondence with the experts, we derived a modified and final version of the mHAT checklist that includes all of the suggested changes.

Results

Literature review

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart in Figure 1 summarises the selection process used to obtain the included studies.[38] Overall, 15 peer-reviewed English-language studies published between 2008–2018 were included in this review.[39–53] Based on the above-described inclusion criteria, each of the 15 studies provided details about the features of physical activity apps that enhanced end-users' experience. In Table 1, we show the condensed themes designed for the focus group discussions. The complete list of the themes derived from the full-text articles can be found in Appendix 1.

The Cohen's kappa value obtained for the title and abstract screening was 0.34, which widely used classifications denote as only fair agreement.

These classifications however, ignore that the interpretation of the kappa value depends heavily on the true proportion of positives among the examined units. In other words, the low kappa value does not indicate that the two screeners disagreed on the articles included but rather very few of the recovered studies were included in the study: less than 1% (15 out of 2,099). Under such circumstances, Bland demonstrated that a kappa of 0.3 actually corresponds to very good agreement.[54]

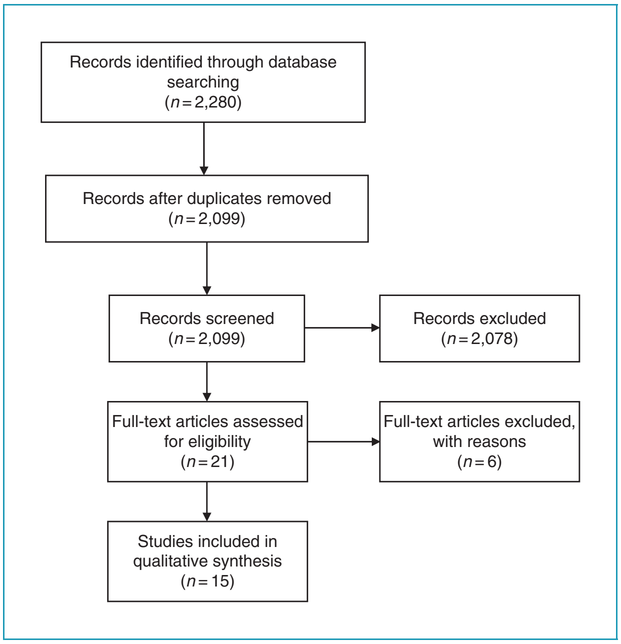
Records identified through database searching
(n = 2,280)
Records after duplicates removed
(n = 2,099)
Records screened
(n = 2,099)
Records excluded
(n = 2,078)
Full-text articles assessed for eligibility
(n = 21)
Full-text articles excluded, with reasons
(n = 6)
Studies included in qualitative synthesis
(n = 15)

Figure 1. Flowchart detailing article selection process.

Focus groups

Table 2 provides a summary of the demographic characteristics of participants in the focus groups. Although the topic guide for the focus group discussions was designed with physical activity apps in mind, participants kept referring to their experiences with a broad range of health apps. Along with the factors that positively influenced the trustworthiness of mHealth apps, participants elaborated on mHealth app features that may lead to mistrust. In some cases, participants expressed views that questioned the role and relevance of trustworthiness in mHealth apps altogether. Below, we elaborate on these divergent findings by identifying the trustworthiness factors, the factors that lie between trust and mistrust as well as the features that encourage mistrust.

Trustworthiness factors

The focus group findings indicate that five major themes affect end-user trust in mHealth apps:

1. Informational content
2. Organizational attributes
3. Societal influence
4. Technology-related features
5. User control

Informational content

This section highlights the considerations participants tend to make when judging the trustworthiness of the information provided by a mHealth app.

Table 1. Condensed list of themes derived from the full-text articles.

Theme	Meaning
Autonomy	Users have the freedom to manage or restrict data access.
Costs	Users either download apps for free or pay for the apps.
Feedback	The app communicates data back to users to help them improve.
Easy functional characteristics	Users can learn to use the app quickly.
Engagement	The features of the app capture users' attention.
Source and content	The app displays accurate content that is easy to understand.
Reputation	The institution curating the app is viewed in a favourable light and recognised as both reliable and competent.
Technical properties	The features engineered into the app that dictate how it operates.
Tracking	Users have ready access to the global positioning system (GPS) data generated by the app.

App: application.

Table 2. Descriptive characteristics of participants.

Variable	Sample size ($n = 20$)
Age (years)	
20–25	2
26–30	14
31–35	1
36–40	2
40–45	1
Gender	
Male	5
Female	15
Highest educational level	
Bachelor's degree	2
Master's degree	16
Doctoral degree	1
Unknown	1
Place of origin	
Europe	15
North America	3
Australia	2

Information accuracy. Participants expected the contents of health apps to be informed by robust research. There was a general awareness that these apps are unlikely to be void of errors and that these errors could be exacerbated by improper use of the app. Nevertheless, participants stressed the need for accurate health apps with negligible margin of errors. Regular content updates and an independent third-party review were two actions that appeared to boost the trustworthiness of the information provided by a health app.

> So, I haven't seen an app which would have an error or a margin of error. You know they all claim to be very accurate and that's obviously not the case. (Participant 15, Female)

> The measurements were not accurate because I didn't wear it always the way I was supposed to. And therefore the measurements were not accurate and therefore I didn't trust it anymore. But I knew it was my fault. (Participant 7, Female)

Understandability. Participants noted that it was difficult to verify and comprehend the information sources used to generate an app's contents. Consequently, they requested that app developers make the reference lists of the research used to develop the app available. There was an expectation that these reference lists would be both easy to locate and lucid enough for lay people to understand. With many apps perpetuating false claims that could be harmful to end-users, participants highlighted that they are more likely to perceive an app as trustworthy if it is accompanied by safety guidelines.

> Because I think that showing the video of the movement and you could see it properly so it was less dangerous to repeat a movement that you can see it multiple times. (Participant 8, Female)

> I think [. . .] citing the data source makes that app more trustworthy. So if someone is interested they can actually go to the data source and they can look themselves to verify the information that they provide and if it's in line with the data source. (Participant 13, Male)

Transparency. Most participants highlighted the need for the purveyors of health apps to be forthcoming about the outcomes of using a particular app. Participants interpreted lengthy terms of service and privacy policies as an attempt by developers to obscure possible risks that could result from using their app. The handling of app users' personal information was described as opaque; thus, participants perceived apps that require 'too much' personal information to either download or use as untrustworthy. Interpretations of how much personal information was too much varied from one participant to the other.

> I would like the app to tell me exactly if I ask [. . .] what the app is doing, what it's recording me about my data and what it's going to do with them. You know but in a very transparent and in a friendly manner. The way I can get it in like two minutes read and decide whether or not to download this app. I think transparency is lacking nowadays. (Participant 9, Female)

> I think for me there's one thing, two things, that's having clear privacy policies so it's clear and transparent what they do with the data and you don't have to

> go through 250 lines to find out in the fine print. (Participant 5, Female)

Organizational attributes

This theme underlines the institutional level factors that stand to influence the trustworthiness of mHealth apps.

Brand familiarity.

Health apps from well-known brands were generally perceived as more trustworthy than those of unknown brands. This heightened favourability stems from an assumption that a company with several products on the market will have the capacity and resources to manufacture better quality products. One other factor that improved the trustworthiness of apps from well-known brands was the perception that they will employ skilled personnel to handle all issues related to the app.

> Yea, I want something that's been on the market, that has a good product in terms of maybe fitness wear, or fitness equipment that I've used in the past, and that is basically it's been on the market, it's a trustworthy name and therefore you'd associate the fact that because of their good reputation so far they've invested in the app properly and they've really invested in the knowledge that went into that app in terms of suggestions and their training programs. (Participant 18, Female)

> Like I'm more comfortable downloading an app from a brand that I know of even if they may share my data or not. But like I know that I'd be more comfortable than a very obscure company that was putting out this app. (Participant 16, Female)

Reputation.

Not-for-profit entities – such as universities, research institutions, and government agencies – were perceived as more trustworthy than their profit-making counterparts. Consequently, participants preferred to contribute their data to not-for-profit entities based on the presumption of benevolent outcomes. Nevertheless, both for-profit and not-for-profit organizations with positive track-records of adhering to strict data protection regulations such as the General Data Protection Regulation (GDPR) were perceived as trustworthy.

> For me, maybe having some logo on the website and this shows there is some collaboration from research academic researchers or with government bodies. (Participant 13, Male)

> So it's not like I don't wanna give you my data, I just wanna know [...] which way are you using the data right? If it is I don't know for research of something like cardiac arrest or I don't know like something related to health, then I can help and I don't really mind. If it is because they are making money out of my data selling those to other companies you know, then I might not be trusting this company. (Participant 11, Female)

Societal influence

This category refers to the societal-level factors that influence trustworthiness judgements towards mHealth apps.

Recommendations.

Fully aware that positive reviews and high download figures could be misleading, participants still chose to interpret mHealth apps with such statistics as trustworthy. Apps that appeared in the top results of search engines as well as those recommended by family, friends or acquaintances were also viewed as highly trustworthy.

> Well, I rely on a friend's opinion. A specific friend who I know has two good qualities. So, one he's good at sports and very knowledgeable about it and second he's very sensitive about data protection issues. So, he's not using anything that there is the slightest chance of being leaky. Right so, I will get things that he would use. (Participant 15, Female)

> They have reviews from other users that you can use and you can see what are most common things used on the app and what are less commonly used and that gives you a little more trustworthiness in some of the features. (Participant 16, Female)

External pressures: 'everybody uses it'

Participants who purchased and were satisfied with devices such as Fitbit or the Apple Watch felt inclined to extend their trust of these devices onto their accompanying apps. In an effort to take advantage of the features of these devices, participants felt compelled to download their accompanying apps. Participants acknowledged however, that downloading the apps improved their overall user experience.

> Well, I only use what some [sic] like how much it goes with my device like the watch that I use. So I was, I didn't really intend to use the app but it turns out it was rather useful. (Participant 15, Female)

> Well, I'm using a [Device name] and you can only connect it with the [App name] app so obviously I downloaded that one. (Participant 17, Female)

Cost. Participants expressed competing ideas about the advantages and disadvantages of paying for health apps. On the one hand, some participants perceived paid apps as more trustworthy than unpaid apps by equating the levy for access with attributes such as quality and data protection. On the other hand, advocates of unpaid apps argued that free apps were no different from paid ones given that the personal data originating from both paid and free apps were subject to malicious activities.

> If I don't pay I don't trust it. Because otherwise what's the business? [. . .] If it's free, it's obvious there's something that they are doing business for. (Participant 3, Male)

> I assume when I pay for something I'm paying a little bit to protect my privacy or my data [. . .] I mean I'm always downloading free apps but I have the assumption that if I would pay perhaps I would be paying for something which has a bit of quality. (Participant 19, Female)

Technology-related features

This theme emphasises the technical features of mHealth apps that reinforce their trustworthiness.

Usability. Participants preferred apps that were easy to use and well-suited for the purposes for which they were created. Customisable apps that factored in aesthetics generally improved the trustworthiness of an app. Well-designed apps convinced participants that a lot of time and effort had been invested into producing the app thereby increasing its trustworthiness. Pop-up advertisements were one feature of health apps that decreased their trustworthiness according to most participants.

> Yea the first impression but not from the app, the user-friendliness and the fact that It's nice actually to use and to look at gave me a good impression and then increase my trust. (Participant 9, Female)

> I think just easy usability as well. Just in terms of like how the functions are laid out, what it looks like, if it's not too complicated, if it allows you pretty much to be mobile with it and use it wherever you go. I'd rather download an app like that rather than an app that I can only just use in the gym or I have to have equipment to do it. But it's an app that gives you the options either way. (Participant 18, Female)

Privacy. Participants emphasised that they are more likely to perceive health apps as trustworthy if they uphold high privacy and security standards. The commitment to these standards was assessed by safeguards to protect users' data and to keep out unauthorised individuals. The participants with technical backgrounds advocated for apps to employ measures such as end-to-end encryption, while 'lay' participants stressed the need for apps to ensure that individual users are not identifiable among the pool of data generated from the app.

> When they have the data on their side, um some of them sometimes say that they will encrypt the data but as soon as it's not encrypted they can do whatever they want [. . .] for me personally more the technology blocks because uh yes if I trust a fitness app, most of the time it's also about the data storage. (Participant 12, Male)

> Well, one thing that I'm really hating about apps these days is so many of them require, request a Facebook log in into the app. I will be much more comfortable giving my data out knowing that it's not linked like it's more anonymous in that way. (Participant 16, Female)

User control

This category highlights the intrinsic and extrinsic motivations that influence the trustworthiness of the data generated from mHealth apps.

Autonomy. Participants often highlighted the desire to have control over the personal data derived from health apps. Hence, they expressed concerns about apps that deprived them of the freedom and authority to determine who had access to their personal data.

> So I will say yea I trust these apps if I can control it, if I can disable the location, if I can falsify the details and if I can actually see if my data is safe or not. (Participant 14, Female)

> So, I like that I can download uh somethings from an app that gives me a bit more let's say the idea that I have some control on it. Like the GPS files. (Participant 10, Female)

Empowerment. Participants rallied against health apps that enrolled them in all data sharing schemes by default: opt-out systems. Instead, they preferred apps that employed opt-in systems to allow them to actively choose the types of data that an app can store. The act of opting-in was interpreted as an empowering and trustworthy strategy. Apps that did not delete end-users' data after discontinued use as well as those that compelled them to share posts on social networking sites were perceived as untrustworthy.

> So if there is an ability to share the data if I'm able to see what the other users are doing with the same app, it makes [it] for me less trustworthy. Because when I can see what other people are doing with the app, I get the suspicion that other people can see what I'm doing... So if there's no option not to share and if I always see what the other people are doing then it's not very trustworthy. (Participant 15, Female)

> Yea, I'd like to have complete control um and I also think I mean I've downloaded apps before and then deleted them for storage issues and then reloaded them and then it says okay we already know that you have an account, we already know it's registered to your Facebook, and then it's like okay, so why have you kept that data and what have you done with it in the mean time? (Participant 18, Female)

Factors that downplay the value of trustworthiness

This category underlines the factors that undercut the value of trustworthy mHealth apps.

Limited digital literacy. During the focus group discussions, it appeared that some participants discovered for the first time the mishaps that could result from the misuse of personal data. These participants were generally unaware of the adverse effects that could result from malicious access to sensitive personal information such as geolocation data or home address. Consequently, they did not factor in trustworthiness when downloading health apps.

> But then again it depends what you like. If I put in I'm 35, I'm female, and right now I'm fertile, I mean who can really use this information? (Participant 7, Female)

> An employer for example? (Participant 4, Female)

> No I'm not aware of the risks so because I'm not aware I am not aware I don't care [laughs] (Participant 8, Female)

Indifference. Some participants chose to overlook the potential adverse effects that could result from the misuse of their personal data on the grounds that it was unlikely that they would be singled out among a large pool of health app users. To such individuals, it was their responsibility to streamline the personal data that an app could amass on them. As such, they were willing to take the responsibility if malevolent activities were to occur with their data.

> About my data, I don't care. I'm aware that they are going to use it. So I do the other way around, the worst case is that they're going to use it for very bad purposes and I don't care. The moment I install it, I'm aware that that can happen and I don't care [...] But I don't care. That's the problem because I know that they're going to use my data for their business. That can be bad or not I don't care but they're business not mine. (Participant 3, Male)

> I think there's always a trade-off between the amount of data you give and the amount of service they provide. If you don't give any data, or delete your own data and everyone does that then they cannot provide the good service that they can provide now. (Participant 13, Male)

> To your point I believe there is also a group effect you call it like this in English? Like you know WhatsApp for instance you think okay everyone uses it so why should I be worried you know? If they do something wrong to me they do something to everyone and then someone may go and complain (Participant 9, Female)

Mistrust factors

Although we set out to identify the features of health apps that make them trustworthy, participants highlighted 10 app characteristics that make them less trustworthy. These 'mistrust factors' have been summarised in the Box 1.

Expert feedback: the mHAT checklist

The final version of the mHAT checklist is laid out in Table 3. There was broad agreement among the experts on the need to strengthen the language within the information accuracy and privacy sections of the checklist. Overall, the mHAT checklist is useful throughout the life cycle of health apps especially during the design and update phases. The checklist can also be beneficial when restructuring organizational practices such as

Box 1. Mistrust factors

1. *Incessant tracking*: apps that constantly monitor users' geolocation.
2. *Lengthy privacy policies*: apps with extensive legal documents that obscure how the data derived from the app is governed.
3. *Low download figures*: apps with low download figures.
4. *Poor reviews*: apps with substandard reviews about their functionality.
5. *Mandatory social networking*: apps that can only be used in concert with social networking sites e.g. Facebook or Twitter.
6. *Excessive personal details*: apps that grant access only after users have entered multiple personal details e.g. billing information and address.
7. *Persistent data after deleting app*: apps that store the personal details of former users
8. *Excessive advertisements and pop-up content*: apps that inundate users with advertisements and pop-up contents.
9. *Unnecessary phone functions*: apps that demand seemingly unnecessary phone functions to operate e.g. microphone, camera etc.
10. *Perceptions of regulatory enforcement*: the level of trust in the ability of a particular government to provide oversight of the products marketed to the general public.

Table 3. The mHealth App Trustworthiness (mHAT) checklist.

Question		Yes	No	Not applicable	In progress	Comments
Informational content						
Information accuracy	Does the app provide accurate measurements?	□	□	□	□	
	Does the app inform end-users about errors in measurements?	□	□	□	□	
	Does the app ensure that personalised data tailored to end-users are precise?	□	□	□	□	
	Is the information on the app certified by an: a. in-house team? b. external third-party team?	□	□	□	□	
	Is the information provided by the app backed by robust research?	□	□	□	□	
	Does the app recommend regular updates to: a. fix bugs inherent within the app? b. amend app contents based on improved research?	□	□	□	□	
Understandability	Is the app accompanied by clear end-user safety guidelines?	□	□	□	□	
	Is the research-backed evidence used to create the app easy to locate and understand?	□	□	□	□	
Transparency	Does the app highlight potential risks or side-effects resulting from its use?	□	□	□	□	
	Are the 'terms of service' concise and easy to read?	□	□	□	□	
	Does the app require only minimal personal data of end-users?*	□	□	□	□	
	Are the privacy policies concise, clear and easy to understand?	□	□	□	□	

(continued)

Table 3. Continued.

Question		Yes	No	Not applicable	In progress	Comments
Organizational attributes						
Brand familiarity	Does the company have other reputable products or services to associate the app with?	□	□	□	□	
Reputation	Does the company curating the app have clear policies on how to handle end-user data?	□	□	□	□	
	Does the company make their data handling history and data breaches available to end-users?	□	□	□	□	
	Is the app affiliated with a non-governmental organization or a reputable government agency?*	□	□	□	□	
	Does the company value data protection regulations?	□	□	□	□	
	Does the company utilise skilled personnel within the app development domain?	□	□	□	□	
	Has the company developed similar apps in the past?					
Societal influences						
Recommendations	Can end-users readily suggest the app to others?	□	□	□	□	
	Does the app have good reviews?	□	□	□	□	
	How easily can end-users locate the app? Does it appear: a. in the top results of search engines? b. as a featured app in the app store?	□	□	□	□	
	Does the app store display how often the app has been downloaded?	□	□	□	□	
External factor	Does the app accompany a wearable device?	□	□	□	□	
Technology-related features						
Usability	Is the app easy to use and have a friendly end-user interface?	□	□	□	□	
	Is the app visually appealing (aesthetics)?	□	□	□	□	
	Does the app send out a reasonable number of notifications?*	□	□	□	□	
	Are the features of the app customisable?	□	□	□	□	
	Is the app accessible by its target audience?*	□	□	□	□	
Privacy	Is the data generated from the app secured by end-to-end-encryption?	□	□	□	□	

(continued)

Table 3. Continued.

Question		Yes	No	Not applicable	In progress	Comments
	How is the data generated from the app stored: a. locally on the device? b. encrypted?	□	□	□	□	
	Is privacy a core consideration throughout the app design phase, i.e. a privacy by design approach?	□	□	□	□	
	Is the data generated from the app anonymised so that individuals are non-identifiable?	□	□	□	□	
	Can users easily access all of their data e.g. address, billing information?	□	□	□	□	
User control						
Autonomy	Do the functions of the app give end-users the overall impression of freedom to control the use of their data?	□	□	□	□	
Empowerment	Does the app allow end-users to restrict data sharing to third-parties such as social networking sites?	□	□	□	□	
	Do end-users act as the proprietors of the data generated from the app?	□	□	□	□	
	Does the app seek explicit end-user permission before sharing data with third-parties?	□	□	□	□	
	Does the app allow end-users to opt-in and decide which data can be stored or processed?*	□	□	□	□	
	Does the app allow end-users to easily delete their data?	□	□	□	□	

App: application.
Certain items in the checklist have been marked with asterisks to signify that they are context-specific.

marketing campaigns relating to a health app. The contents of the checklist have been reviewed and deemed feasible by the six experts.

There are six main columns in the mHAT checklist comprising of the following titles: questions, yes, no, not applicable, in progress, as well as a comments section. Certain items in the checklist have been marked with asterisks to signify that they are context-specific. Therefore, for these checklist items, it is advised to interpret their meaning on a case-by-case basis. For instance, if a paid app requires the billing information of an end-user before it can be downloaded, entering one's card details may not be considered as requiring excessive information. Overall, users of the mHAT checklist should aim for more 'yes' than 'no' answers. In Appendix 2, we provide more details on the checklist.

Discussion

In this study, we have investigated end-users' opinions on the features and actions that make mHealth apps trustworthy and charted these findings in the mHAT checklist. One major advantage of the mHAT checklist is that its format (i.e. a checklist) makes it convenient to use. Aside from being lucid enough for lay people to understand and employ, checklists are vital in reducing errors caused by omissions.[55]

The checklist is a suitable layout to present the findings obtained in this study. Referring back to some mHealth oversight examples in the UK context, the current version of the NHS Apps Library's Digital Assessment Questionnaire V2.1 as well as the NICE Evidence Standards Framework targeting app developers have similar formats.[17,18] Since the purpose of the mHAT checklist is to signpost app developers to the features and actions that make mHealth apps trustworthy, it is vital that it is presented in a format that gets its message across concisely.

The mHAT checklist is not only useful in the initial stages of creating an app but throughout its life cycle. The life cycles of health technologies typically consist of four stages: (a) a development stage that involves needs assessment; (b) an implementation stage whereby crucial managerial decisions are taken to introduce a technology; (c) an integration stage that entails embedding the new program into an already existing one, and (d) a sustained operation stage.[56] The mHAT checklist meets these variable needs as it attempts to bring the expectations of end-users to light. For example, knowing how end-users will perceive a lengthy terms of service document is not only important for the development stage of a mHealth app but also its implementation stage.

Implications of this study

The participants in this study highlighted conflicting views about how the cost of a mHealth app affects its trustworthiness. While some participants interpreted paid apps to mean that they were of better quality, others did not subscribe to this view at all. The basis for the latter argument was that the cost of an app is meaningless in relation to its trustworthiness. This difference in opinion resembles the findings of other studies. In one study aimed at cataloguing mHealth apps, free health and fitness apps were consistently rated better than paid ones.[57] The results of a different study show pricier apps to be perceived as more credible and trustworthy.[53] Future studies should attempt to uncover the underlying relationship between cost and trustworthiness.

Participants generally agreed that credible organizations must be charged with cross-checking the claims made by mHealth apps. In line with these oversight tasks, there were suggestions to hand out accreditation that is commensurate with compliance levels. When asked to identify which groups they consider credible and which aspects of the app to assess, participants stressed the need for unbiased, non-profit making entities such as non-governmental organizations or government agencies. Indeed, these expectations are not far-fetched and do in fact support the idea that individuals seek reassurances against potential risks and unpredictability brought forth by increasing complexity in society.[57] The increased credibility of non-profit making entities may lie in a heightened perception of benevolence in line with the concept of trust and trustworthiness.

Future accreditation bodies that are established to assess health apps will need novel and more innovative funding streams to succeed. Traditional certification companies within the health technology industry thrive in part because hospitals demand objective reviews about their costly software systems. As uncovered in this study, a substantial number of mHealth app end-users prefer free apps and are thus unlikely to shell out resources to ascertain the trustworthiness of mHealth apps. An alternative option may be for an accreditation body to work directly with app developers. In such cases however, conflicts of interests are bound to emerge.[9] Thus, there is the need to identify viable means of guaranteeing funding streams for such an organization.

Echoes of privacy and data security concerns were evident in participants' descriptions of their interactions with mHealth apps. Particularly, participants expressed discomfort with apps that provided lengthy privacy policies, permitted use only after entering many personal details and refused to delete end-users' data after discontinued use. These sentiments appear to be justified, however, as there is evidence to substantiate these concerns. For instance, some mHealth apps lack privacy policies altogether; in cases where policies do exist, they may be composed of boilerplate text and may require a college-level education to understand.[58,59]

Two main reasons were responsible for why participants justified their engagement with mHealth apps that they found untrustworthy: limited digital literacy and indifference. Theories purporting a relationship between ignorance and trust may be useful in explaining this phenomenon. According to Gambetta, cognitive dissonance results when an individual wrestles with their beliefs about the trustworthiness of another entity – in this case mHealth apps. To counteract this discomfort, therefore, individuals may be willing to act irrationally by prioritising need over trust.[60] This theory, along with our findings, demonstrates that it is likely that needing something may confound efforts to uncover what really makes it trustworthy. Further studies should be conducted to determine the interdependencies of these two concepts.

To address the relevance of trustworthiness of health apps, it is vital to assess whether participants make distinctions between medical and health apps. During the focus group discussions, participants kept referring to their broad experiences with both types of apps. Although this aided in making the findings of our study generalisable to mHealth apps, it highlighted that these types of apps may not be perceived differently at all. Considering that such evidence will be vital in shaping education campaigns on increasing public awareness of trustworthy apps, this issue should be investigated further.

Strengths and weaknesses of the approach

A major strength of this study is the use of the focus group methodology. The group dynamics involved in focus groups allowed participants to counteract each other and to reveal the nuances informing their decisions to perceive an app as trustworthy. The ranking of themes exercise used during the focus group discussions was best suited for this kind of setting as the focus groups were largely homogenous. In the midst of peers, participants tend to feel comfortable to express their views, however unpopular.[34]

We went to great lengths to ensure that the focus groups were made up of the suggested 6–8 people. However, due to drop out rates and difficulty in scheduling a suitable time for participants, we were unable to meet these standards. Nevertheless, two out of four focus groups were comprised of 6–8 people. Another limitation of this study is selection bias as all of the participants were young and highly educated. Although this is unlikely to skew our findings, it may have affected the issues highlighted by the participants. For example, most of the participants appeared more knowledgeable about privacy and security issues than the average person.[61] Lastly, only six experts with technical backgrounds validated the mHAT checklist. This decision was influenced by time and resource limitations. The mHAT checklist will be better served if validated among a wider and more diverse group of experts. Nonetheless, the checklist can be considered robust as the experts provided feedback to refine its contents.

Conclusion

This study contributes to our understanding of the attributes of trustworthy mHealth apps. Our findings reveal that app developers must be conversant with variable issues ranging from informational content, organizational attributes, societal influence, technology-related features and user control factors when attempting to create trustworthy health apps. The mHAT checklist is a useful tool in advancing continued efforts to ensure that health technologies are trustworthy.

Acknowledgements: The authors would like to thank the focus group participants and experts for allocating time to contribute to this study. The authors would also like to thank Walter Karlen and John Powell for their valuable comments.

Contributorship: AvH designed the project under the supervision of EV. AvH and MF conducted the literature review; AvH and FG designed the study under the supervision of EV; AvH and FG conducted the focus group discussions. All of the authors reviewed and edited the manuscript and approved the final version of the manuscript.

Conflict of interest: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: The ethics committee of ETH Zurich approved this study (REC number: EK 2018-N-60).

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by grants from the Swiss National Science Foundation (grant number: PP000P3-157556; 4075210-167223).

Guarantor: EV.

Peer review: This manuscript was reviewed by reviewers, who has chosen to remain anonymous.

References

1. Paglialonga A, Lugo A and Santoro E. An overview on the emerging area of identification, characterization, and assessment of health apps. *J Biomed Inform* 2018; 83: 97–102.
2. Krebs P and Duncan D. Health app use among US mobile phone owners: A national survey. *JMIR* 2015; 3: e101.
3. Cortez N, Cohen IG and Kesselheim A. FDA regulation of mobile health technologies. *N Eng J Med* 2014; 371: 372–379.
4. Bates DW, Landman A and Levine DM. Health apps and health policy: What is needed? *JAMA* 2018; 320: 1975–1976.
5. Ventola CL. Mobile devices and apps for health care professionals: Uses and benefits. *P T* 2014; 39: 356–364.
6. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M, et al. Mobile apps in cardiology: Review. *JMIR Mhealth Uhealth* 2013; 1: e15.
7. Shuren J, Patel B and Gottlieb S. FDA regulation of mobile medical apps *JAMA* 2018; 320: 337–338.
8. Plante T. Behind the paper: The rise of 'snake oil' mobile health apps in an era absent of regulation, https://npjdigitalmedcommunity.nature.com/users/172239-timothy-plante/posts/37590-this-might-be-why-some-snake-oil-mobile-health-apps-sell-like-hotcakes (2018, accessed 14 January 2019).
9. Powell AC, Landman AB and Bates DW. In search of a few good apps. *JAMA* 2014; 311: 1851–1852.
10. Buijink A, Visser B and Marshall L. Medical apps for smartphones: Lack of evidence undermines quality and safety. *Evid Based Med* 2013; 18: 90–92.
11. Bindhim NF and Trevena L. Health-related smartphone apps: Regulations, safety, privacy and quality. *BMJ Innov* 2015; 1: 43–45.
12. Torous J and Roberts LW. Needed innovation in digital health and smartphone applications for mental health: Transparency and trust. *JAMA Psychiatry* 2017; 74: 437–438.

13. van Velthoven M and Powell J. Do health apps need endorsement? Challenges for giving advice about which health apps are safe and effective to use. *Digit Health* 2017; 3: 2055207617701342.
14. Krieger WH. When are medical apps medical? Off-label use and the Food and Drug Administration. *Digit Health* 2016; 2: 2055207616662782.
15. Goyal S and Cafazzo JA. Mobile phone health apps for diabetes management: Current evidence and future developments. *QJM* 2013; 106: 1067–1069.
16. NHS. NHS Apps Library, https://www.nhs.uk/apps-library/ (2019, accessed 7 March 2019).
17. NHS Digital. Digital assessment questions V2.1. Report, https://developer.nhs.uk/wp-content/uploads/2018/09/Digital-Assessment-Questions-V2.1-Beta-PDF.pdf, 16 August 2018.
18. National Institute for Health and Care Excellence. *Evidence standards framework for digital health technologies*. UK, March 2019.
19. Levi M and Stoker L. Political trust and trustworthiness. *Annu Rev Polit Sci (Palo Alto)* 2000; 3: 475–507.
20. O'Neill O. Trust, trustworthiness and transparency, https://www.efc.be/human-rights-citizenship-democracy/trust-trustworthiness-transparency/ (2015, accessed 10 March 2019).
21. Galeser E, Laibson D, Scheinkman J, et al. What is social capital? The determinants of trust and trustworthiness. *NBER Work Pap Ser* 1999; 7216: 1–62.
22. Murnane EL, Huffaker D and Kossinets G. Mobile health apps: Adoption, adherence, and abandonment. *UBICOMP/ISWC* 2015: 261–264.
23. Abelson J, Miller FA and Giacomini M. What does it mean to trust a health system? A qualitative study of Canadian health care values. *Health Policy* 2009; 91: 63–70.
24. Albrecht UV. Transparency of health-apps for trust and decision making. *J Med Internet Res* 2013; 15: e277.
25. Sinatra GM and Hofer BK. Public understanding of science. *Policy Insights Behav Brain Sci* 2016; 3: 245–253.
26. Agarwal S, LeFevre AE, Lee J, et al. Guidelines for reporting of health interventions using mobile phones: Mobile health (mHealth) evidence reporting and assessment (mERA) checklist. *BMJ* 2016; 352: i1174.
27. Spigel L, Wambugu S and Villella C. mHealth data security, privacy and confidentiality guidelines: Companion checklist, https://www.measureevaluation.org/resources/publications/ms-17-125b (2018, accessed 7 March 2019).
28. Stoyanov SR, Hides L, Kavanagh DJ, et al. Mobile app rating scale: A new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015; 3: e27.
29. Albrecht U-V, Noll C and Von Jan U. App-synopsis: Self-assessment on trust or distrust of health-apps. *Stud Health Technol Inform* 2014; 202: 233–236.
30. Albrecht U-V. Transparency of health-apps for trust and decision making. *J Med Internet Res* 2013; 15: e277.
31. Adjekum A, Blasimme A and Vayena E. Elements of trust in digital health systems: Scoping review. *J Med Internet Res* 2018; 20: e11254.
32. Critical Appraisal Skills programme. CASP qualitative checklist, https://casp-uk.net/wp-content/uploads/2018/03/CASP-Qualitative-Checklist-2018_fillable_form.pdf (2018, accessed 7 March 2019).
33. Maguire M and Delahunt B. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *AISHE-J* 2017; 8: 3351–33514.
34. Wilkinson S. Focus group methodology: A review. *Int J Soc Res Methodol* 1998; 1: 181–203.
35. Bertrand J, Brown J and Ward V. Techniques for analyzing focus group data. *Eval Rev* 1992; 16: 198–209.
36. Taylor-Powell E and Renner M. Analyzing qualitative data. Report, University of Wisconsin-Extension Cooperative Extension Madison, Wisconsin Program Development and Evaluation, USA, 2003.
37. Clayton MJ. Delphi: A technique to harness expert opinion for critical decision-making tasks in education. *Educ Psychol (Lond)* 1997; 17: 373–386.
38. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 2009; 6: e1000097.
39. Ancker JS, Witteman HO, Hafeez B, et al. 'You get reminded you're a sick person': Personal data tracking and patients with multiple chronic conditions. *J Med Internet Res* 2015; 17: e202.
40. Giunti G, Guisado Fernandez E, Dorronzoro Zubiete E, et al. Supply and demand in mHealth apps for persons with multiple sclerosis: Systematic search in app stores and scoping literature review. *JMIR Mhealth Uhealth* 2018; 6: e10512.
41. Guo Y, Bian J, Leavitt T, et al. Assessing the quality of mobile exercise apps based on the American College of Sports Medicine guidelines: A reliable and valid scoring instrument. *J Med Internet Res* 2017; 19: e67.
42. Hebden L, Cook A, van der Ploeg HP, et al. Development of smartphone applications for nutrition and physical activity behavior change. *JMIR Res Protoc* 2012; 1: e9.
43. Howe KB, Suharlim C, Ueda P, et al. Gotta catch'em all! Pokémon GO and physical activity among young adults: Difference in differences study. *BMJ* 2016; 355: i6270.
44. Lau Y, Cheng LJ, Chi C, et al. Development of a healthy lifestyle mobile app for overweight pregnant women: Qualitative study. *JMIR Mhealth Uhealth* 2018; 6: e91.
45. Lee HE and Cho J. What motivates users to continue using diet and fitness apps? Application of the uses and gratifications approach. *Health Commun* 2017; 32: 1445–1453.
46. Lobelo F, Kelli HM, Tejedor SC, et al. The wild wild West: A framework to integrate mHealth software applications and wearables to support physical activity assessment, counseling and interventions for cardiovascular disease risk reduction. *Prog Cardiovasc Dis* 2016; 58: 584–594.
47. Mohadis HM, Mohamad Ali N and Smeaton AF. Designing a persuasive physical activity application for older workers: Understanding end-user perceptions. *Behav Inf Technol* 2016; 35: 1102–1114.

48. Parpinel M, Scherling L, Lazzer S, et al. Reliability of heart rate mobile apps in young healthy adults: Exploratory study and research directions. *J Innov Health Inform* 2017; 24: 224–227.
49. Puszkiewicz P, Roberts AL, Smith L, et al. Assessment of cancer survivors' experiences of using a publicly available physical activity mobile application. *JMIR Cancer* 2016; 2: e7.
50. Robertson MC, Tsai E, Lyons EJ, et al. Mobile health physical activity intervention preferences in cancer survivors: A qualitative study. *JMIR Mhealth Uhealth* 2017; 5: e3.
51. Short CE, Finlay A, Sanders I, et al. Development and pilot evaluation of a clinic-based mHealth app referral service to support adult cancer survivors increase their participation in physical activity using publicly available mobile apps. *BMC Health Serv Res* 2018; 18: 1–11.
52. Wen D, Zhang X, Liu X, et al. Evaluating the consistency of current mainstream wearable devices in health monitoring: A comparison under free-living conditions. *J Med Internet Res* 2017; 19: e68.
53. West JH, Hall PC, Hanson CL, et al. There's an app for that: Content analysis of paid health and fitness apps. *J Med Internet Res* 2012; 14: e72.
54. Bland M. *An introduction to medical statistics*. 4th ed. Oxford: Oxford University Press, 2015.
55. Scriven M. *The logic and methodology of checklists. Report*, Western Michigan University, USA, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.588.7093&rep=rep1&type=pdf, 2000.
56. Khoja S, Durrani H, Scott RE, et al. Conceptual framework for development of comprehensive e-health evaluation tool. *Telemed J E Health* 2013; 19: 48–53.
57. Lewis JD and Weigert A. Trust as a social reality. *Soc Forces* 1985; 63: 967–985.
58. Sunyaev A, Dehling T, Taylor PL, et al. Availability and quality of mobile health app privacy policies. *J Am Med Inform Assoc* 2015; 22: e28–e33.
59. Rosenfeld L, Torous J and Vahia IV. Data security and privacy in apps for dementia: An analysis of existing privacy policies. *Am J Geriatr Psychiatry* 2017; 25: 873–877.
60. Gambetta D. 'Can We Trust Trust?' In Gambetta D. (ed) *Trust: Making and breaking cooperative relations*. Oxford: Blackwell, 1988, pp.213–237.
61. Obar JA and Oeldorf-Hirsch A. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Inf Commun Soc* 2018; 1–20.
62. Cavoukian A. *Privacy by design in law, policy and practice: A White Paper for regulators, decision-makers and policy-makers*. Ontario, Canada: Information and Privacy Commissioner, August 2011.

Appendix 1. List of the themes derived from the full-text articles.

Themes	Ancker, 2015[39]	Giuinti, 2018[40]	Guo, 2017[41]	Hebden, 2012[42]	Howe, 2016[43]	Lau, 2018[44]	Lee, 2017[45]	Lobelo, 2016[46]	Mohadis, 2016[47]	Parpinel, 2017[48]	Puszkiewicz, 2016[49]	Robertson, 2017[50]	Short, 2018[51]	Wen, 2017[52]	West, 2012[53]
Data sharing															
Manage access to data: authorised personnel only						X		X				X			
App is password protected				X		X									
Control of location services (i.e. for services weather forecasts)									X			X			
Users control access to information		X					X	X				X			
Engagement															
Access to social networking (e.g. social media, forums)		X				X	X		X		X	X	X		X
Augmented reality					X										
Detailed instructions on app functions													X		
Note-taking function	X			X			X								
Use of gamification elements (e.g. tasks and rewards)		X		X	X		X		X			X	X		
Source and content															
Accurate and evidence-based content		X	X	X		X	X	X	X		X	X			
Cost of access		X						X							X
Credible and verifiable information sources		X					X	X	X	X	X	X			X

(continued)

Appendix 1. Continued.

Themes	Ancker, 2015[39]	Giuinti, 2018[40]	Guo, 2017[41]	Hebden, 2012[42]	Howe, 2016[43]	Lau, 2018[44]	Lee, 2017[45]	Lobelo, 2016[46]	Mohadis, 2016[47]	Parpinel, 2017[48]	Puszkiewicz, 2016[49]	Robertson, 2017[50]	Short, 2018[51]	Wen, 2017[52]	West, 2012[53]
Endorsement of standard guidelines (e.g. ACSM)			X			X		X	X	X	X				
Explicit safety warnings and risks			X			X			X		X	X			
Reputation of developers (e.g. commercial vs government)		X							X			X			
Inclusion of broad range of activities, information or app versions		X				X			X		X	X	X		X
Motivating content						X	X		X		X	X	X		
Plain, casual, concise and lay terms to describe complicated concepts	X						X	X				X	X		
Recommendations from healthcare professionals													X		X
Testimonials of previous users									X		X	X			
Theory-driven behaviour change strategies				X				X					X		
Visual representation of instructions and data		X	X	X		X		X	X		X	X	X		
Technical and functional characteristics															
Minimal measurement error										X				X	
Accessible from multiple operating systems,				X				X				X	X		

(continued)

Appendix 1. Continued.

Themes	Ancker, 2015[39]	Giuinti, 2018[40]	Guo, 2017[41]	Hebden, 2012[42]	Howe, 2016[43]	Lau, 2018[44]	Lee, 2017[45]	Lobelo, 2016[46]	Mohadis, 2016[47]	Parpinel, 2017[48]	Puszkiewicz, 2016[49]	Robertson, 2017[50]	Short, 2018[51]	Wen, 2017[52]	West, 2012[53]
platforms and devices: (e.g. Web, app)															
Easy searches for app content and functions	X					X						X	X		
Easy to find and install app in store				X							X				
Inclusion of smartphone multimedia functions		X				X		X							
Manageable number of steps to carry out tasks						X						X	X		
Aesthetic design			X	X		X	X						x		
Reminders		X							X		X	X			
Speed of content upload				X						X					
User accounts with login requirements		X		X											
Tracking and feedback															
Tailor app to individual goal or target preferences	X	X	X	X		X		X	X		X	X	X		
Remote monitoring by healthcare professionals outside of clinical settings		X				X		X			X				
Users can track and self-evaluate their performance	X	X				X	X		X		X	X			X

(continued)

Appendix 1. Continued.

Themes	Ancker, 2015[39]	Giuinti, 2018[40]	Guo, 2017[41]	Hebden, 2012[42]	Howe, 2016[43]	Lau, 2018[44]	Lee, 2017[45]	Lobelo, 2016[46]	Mohadis, 2016[47]	Parpinel, 2017[48]	Puszkiewicz, 2016[49]	Robertson, 2017[50]	Short, 2018[51]	Wen, 2017[52]	West, 2012[53]
Positive feedback mechanisms									X		X	X			
Real-time feedback on progress						X			X			X	X		X
Facilitate communication between healthcare professionals						X		X			X		X		

Appendix 2: The mHAT checklist

What is the goal of this checklist?

The mHAT checklist is a tool that intends to inform the custodians of physical activity apps such as developers and management about end-user criteria for trustworthy apps. Although the primary focus of the mHAT checklist is physical activity, its contents are potentially resourceful for the broad community of app developers.

How was the checklist developed?

The premise for this checklist was derived from a comprehensive review of the literature on trust and digital health.[31] The empirical evidence used to create this checklist was obtained from four focus groups comprising of 20 end-users of physical activity apps. Within each focus group, participants described the features, resources, circumstances and characteristics of these apps that portray them as either trustworthy or untrustworthy. The mHAT checklist builds on the conceptualization of trustworthy apps as discussed by the focus group participants.

What are the contents of the checklist?

This checklist contains actions and considerations that must be present for end-users to trust mHealth apps. The checklist is clustered into five main categories:

- Informational content: this category outlines end-user considerations about the information provided by the app.
- Organizational attributes: this category highlights institutional level factors that influence trustworthiness of the app.
- Societal influences: this category emphasises the societal-level factors that influence trustworthiness judgements of the app.
- Technology-related features: this category describes the technical features of the app that affect its trustworthiness.
- User control: this category explains the intrinsic and extrinsic motivations that influence the trustworthiness of the data generated from the app.

How should the checklist be used?

The mHAT checklist should be used throughout the life cycle of health apps. It is especially relevant for the design and update phases of app development. It can also be used to structure organizational practices and marketing campaigns. Checklist users are encouraged to follow-up on the items marked with asterisks (*), by reviewing studies that shed more light on end-user preferences in specific contexts. This extra effort is necessary because focus group participants expressed differing opinions on these items making them context specific.

To respond to the questions in the checklist, check the box next to one of these options:

- Yes
- No
- In progress
- Not applicable

How can the results be interpreted?

Checklist users must aim for more ‘yes’ than ‘no’ answers when using the checklist. Each yes answer amounts to one point and the maximum number of points that can be obtained is 44. It is highly recommended to use your discretion when interpreting the questions that are flagged as either ‘not applicable’ or ‘in progress’. The ‘comments’ column is provided to allow room for jotting down notes next to each question.

Glossary of checklist terms

- *Anonymised data*: individuals cannot be identified from the data generated by the app.
- *Data protection regulations*: legislation (e.g. GDPR) that regulates the actions of the processors and controllers of personal data.
- *Encrypted data*: security mechanisms that inhibit unauthorised persons from accessing the data generated from the app.
- *End-users*: individuals that use and interact with the app.
- *End-user interface*: the manner in which an individual interacts with an app.
- *Margin of error*: the degree to which app measurements deviate from real values.
- *Personal data*: any information in the data generated by the app that links directly to an individual.
- *Privacy policies*: disclosures about the methods used to gather, manage and utilise end-user data.
- *Safety guidelines*: recommendations that provide assistance on safe use of the app.
- *Terms of service*: rules that dictate the conditions that end-users must agree to in order to use the app.
- *Third-parties*: individuals or entities that are not directly involved with the app.
- *Privacy by design*: an approach that implements privacy at the outset of app development taking into consideration the design phase, network infrastructure and a responsible business culture.[62]