



This is a repository copy of *Involving users in the evaluation of apps for specific health conditions*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/125283/>

Version: Accepted Version

Article:

Powell, L. orcid.org/0000-0003-0230-8722, Joddrell, P. and Parker, J. (2017) Involving users in the evaluation of apps for specific health conditions. *Studies in Health Technology and Informatics*, 242. pp. 646-653. ISSN 0926-9630

<https://doi.org/10.3233/978-1-61499-798-6-646>

The final publication is available at IOS Press through
<http://dx.doi.org/10.3233/978-1-61499-798-6-646>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Involving users in the evaluation of apps for specific health conditions

Keywords. Dementia, ADHD, apps, evaluation, methodology

1. Plain Language Overview

Many mobile applications (apps) target specific long-term conditions. There is little evidence of how to ascertain their suitability for the complex needs of these populations. This paper presents two examples of how dementia and Attention Deficit Hyperactivity Disorder app suitability can be ascertained.

2. Introduction

An estimated 46% of the worldwide population has an Internet connection, which when compared with the 1% proportion in 1995 provides evidence of the rapid growth of our global connectivity [1]. In the United Kingdom (UK), 87.9% of adults are Internet users [2]. This rise has affected how people engage with healthcare services and make decisions about their own health, with potentially positive and negative consequences [3].

An increasingly popular form of engagement with the Internet is through mobile applications (apps), which have become synonymous with modern smartphones, tablets and other portable devices. Since the release of the 'App Store' for Apple 'iOS' devices in 2008, 140 billion apps have been downloaded [4] and there are currently more than two million apps available for download in this store and Google's 'Play Store' [5]. Apps can be used for a multitude of different purposes, including gaming, productivity, creativity and socialising.

Both Apple and Google feature categories in their online app stores for 'Medical' or 'Health & Fitness' apps, respectively. These apps cover such functions as health promotion, fitness guides and public health [6]. There are also apps targeting specific conditions; to provide self-management health information or tailored content to meet the needs of the person living with condition [7]. For example, in a review of apps to support people with diabetes, the most common app functions included self-monitoring of blood glucose, a tool to track insulin or oral diabetic medications, and a dose calculator for prandial insulin [8].

Assessing the quality of health-related apps is an important issue, given the risks that are associated with providing misleading information or unsubstantiated claims [7] or when storing confidential patient information [6]. High-profile examples of where these risks became reality included the data protection compliance issues with

accredited apps in the NHS Health Apps Library [9] and the challenge and subsequent Federal Trade Commission charges against the creators and marketers of the ‘brain-training’ app ‘Lumosity’ for deceiving customers with unfounded claims [10, 11].

As in both cases; the involvement of researchers in the assessment of publicly available apps is clearly warranted to offer independent, unbiased and, most importantly, evidence-based guidance. A generic tool for assessing the quality of health apps exists [12], but this does not allow for a comprehensive evaluation of the functions and features of condition-specific apps with input from a range of stakeholders, most importantly including people living with the condition themselves. There have also been examples of evaluations of apps that are condition-specific, including asthma [13] diabetes [8], bi-polar disorder [14] and chronic pain [15]. Whilst these studies have certainly addressed the content of apps with regards to their quality and the underlying evidence, they have not involved service users to address the suitability of apps for their own needs. This was therefore the approach adopted by the researchers to look at the suitability of apps for specific health conditions.

3. Method

This paper presents two examples that have been adopted by University of Sheffield researchers to explore what makes an app suitable for a population. Two differing examples are presented: firstly, to find apps where the content has been identified for a specific population; and secondly, to find apps with generic content but featuring design characteristics that make them suitable for a specific population.

3.1. Attention Deficit Hyperactivity Disorder

Example one involves the exploration of what makes an app suitable for children and young people diagnosed with Attention Deficit Hyperactivity Disorder (ADHD). Therefore, the research question was “Are the top ten listed apps specifically designed for children and young people suitable and what are the key components for apps to be suitable for this population?”

In this context, the strategy involves identifying suitable mobile applications and conducting qualitative semi-structured interviews exploring what the target population believe makes an app suitable for them.

3.1.1. Search and identification of apps

In June 2016, a search of mobile apps in the Apple iTunes Store and the Android Google Play Store was conducted. These databases were selected due to displaying systematically organised app rankings defined by unique algorithms unique to each app store, commonly known as App Store optimisation (ASO). For Apple the primary factor is number of downloads however there are also many other secondary factors such as keywords and visuals [16]. Similarly the Android database is filtered according to multiple criteria including the volume of ratings, value of ratings and download growth [17]. Although this gives rise to potential bias as apps are selected according to the databases own ASO, it is to some extent unavoidable unless all of the search results are downloaded for testing [18].

The term “ADHD” was searched in both Google Play and iTunes app stores.

Preliminary screening was conducted based on app titles, full marketing description, and screenshots of the apps potentially relevant for inclusion. The first five apps were included from each app store, leaving a total of ten apps for inclusion. See Table 1 for inclusion criteria of apps.

Table 1. Inclusion and exclusion criteria of apps designed for people with ADHD

Inclusion Criteria	Exclusion Criteria
<ul style="list-style-type: none"> States aimed at ADHD The user is the child/young person with ADHD/ADD Mobile application App is available in the English language 	<ul style="list-style-type: none"> Does not state app is aimed at ADHD or ADD Not targeted at the child/young person with ADHD/ADD (e.g targeted at parents, clinicians or adults) Not a mobile application Not available in the English language Duplicate app

Duplicate apps were then removed (See Figure 1). This was applicable if there was more than one version of an app. It was decided that the app version to be included was the app that appeared first on the app store list. Duplicates across both app stores did not apply.

The remaining ten apps were downloaded onto tablet devices and their content was summarised into a tabular format to help assist participants (clinicians only) during the semi-structured interviews.

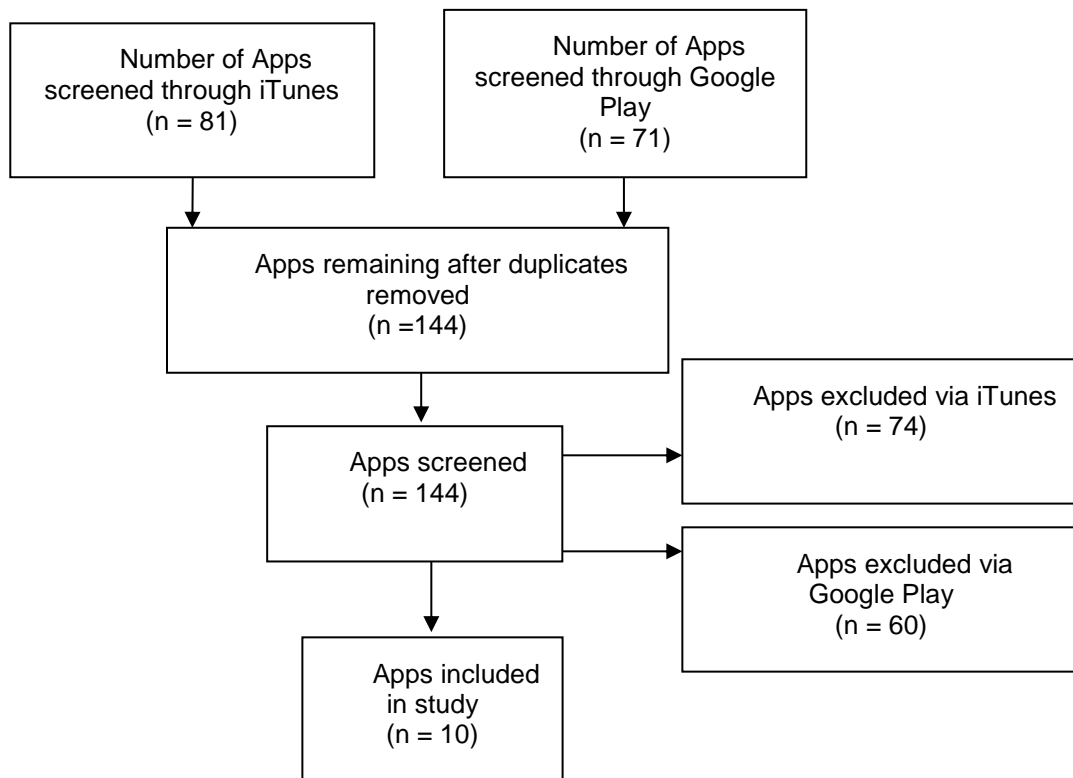


Figure 1. App selection process [19]

3.1.2. Participants and procedure

A convenience sample of five clinicians and five children and young people (YP) were invited to take part in semi-structured interviews. Interviews lasted up to 60 minutes with the clinicians, and up to 45 minutes with the child/YP.

During the interviews, clinicians and children/young people were presented the identified apps. Interview discussions were guided by an interview schedule covering four key areas: what they believe makes a successful app, what doesn't make a successful app, what an app's function should be for this population and how they believe apps could help manage ADHD/address difficulties in YP with ADHD. Participants were given the opportunity to use the apps themselves during the interview, provide their opinions and also were asked what they believe makes a suitable app for YP with ADHD. The two groups provided two unique perspectives: user perspective and a clinical perspective. Participants also completed a short questionnaire on their demographic characteristics. The interviews focused on suitability of the apps for the population.

3.1.3. Data Analysis

All interviews were audio recorded and transcribed. Thematic analysis [20] was used to search for data patterns within and across the participant groups. Authors LP and JP independently identified codes and themes from the transcripts. Discrepancies were resolved through group discussion in an iterative fashion between authors. Themes identified aimed to capture the essence of the participant's views. The themes were used to develop criteria that make an app suitable for this population. The apps were then scored against these criteria to demonstrate the suitability of the apps included in the study.

3.2. Dementia

In the second example for identifying suitable apps for a specific health condition, the research focused on gaming and activity apps for people living with dementia. In contrast to the first example, this project was concerned with generic apps (i.e. that had not been designed for a specific health population) and how certain design features could make them more accessible for people experiencing cognitive impairment because of their dementia diagnosis.

3.2.1. Search and identification of apps

Two game types were identified based on previous research by colleagues in the Netherlands [21] and discussions with members of a local patient and public involvement group. These games were 'Solitaire' (also known as 'Patience') and a 'bubble matching' puzzle game. Terms to represent the two games were selected ('Solitaire' or 'Patience' and 'bubble') and these were entered into the search bar of the Apple iTunes Store. Ten apps of each of the two identified game types were downloaded based on the inclusion and exclusion criteria presented in Table 2. As with the first example, it is acknowledged that by downloading the top ten listed apps as organised by the app store, the unavoidable potential for bias is present.

A literature review was conducted to find out existing knowledge of the use of touchscreen apps by people living with dementia [22], and this information was used to inform the selection of the most accessible version of each of these games from the ten representations downloaded. Once the most accessible version of each game type was identified, these were tested directly with people living with dementia in a quasi-experimental study.

Table 2. Inclusion and exclusion criteria of ‘Solitaire’ or ‘bubble matching’ apps

Inclusion Criteria	Exclusion Criteria
<ul style="list-style-type: none"> • Accurate representation of target game • App is available in the English language 	<ul style="list-style-type: none"> • Rules different to basic version of game • Duplicate app

3.2.2. Participants and procedure

Thirty people living with dementia were recruited from care services and asked to play one of the two apps independently on three separate occasions. All gameplay was recorded using two video cameras mounted on tripods to capture a view of the participants’ faces and their interactions with the app. After each gameplay session, the researcher administered a brief questionnaire to assess their experience of the app (a more detailed version of the procedure can be accessed in [23]). This approach was used to ensure that participants who can experience difficulty with communication because of their dementia diagnosis, were able to provide feedback on their experience of testing the app through direct and indirect means.

3.2.3. Data analysis

Using the video recordings, each of the 84 gameplay sessions were analysed using the Observer video analysis software to code every participant interaction with the app. This allowed the researchers to see what features of the apps were facilitating successful interactions and what features were contributing to unsuccessful interactions. The questionnaire responses were collated and this information was used to assess whether participants’ own experiences of testing the apps were positive or negative.

4. Results

4.1. Attention Deficit Hyperactivity Disorder

Exploring the suitability for apps for children and young people with ADHD revealed that in order to engage the population, it is important for the apps to be visually appealing, the user should be able to interact with the app rather than watch visual stimuli on the screen alone, it is important to be able to relate to the app, to personalise the app (e.g. change what a character looks like to represent the user), provide instant reward and audio feedback. These were all themes identified from clinicians and children/young people data. In addition, clinicians stated that this population would also benefit from being able to monitor their symptoms via an app, and for an app to focus on strengthening relationships with others [19].

4.2. Dementia

Engaging directly with people living with dementia to evaluate the suitability of existing apps revealed many key design features that facilitated successful interactions and also several features that proved to be barriers. The findings have been applied in the design of a framework for identifying further apps that might be suitable for people living with dementia [18]. The results have also been used to improve the two tested apps through collaboration with the developers to implement accessibility features into their existing app. The updated versions of these apps are now available to download in both the Apple and Google app stores, demonstrating the potential for collaborative research involving people living with a specific health condition and app developers.

5. Discussion

Two potential frameworks for evaluating mobile apps have been constructed. Although different, both frameworks have in common that they take into account the views and opinions of the target audience of certain mobile applications to ascertain their suitability. They are novel methods as they have been applied to two very different conditions (ADHD and Dementia) which shows these frameworks have potential to be applied across many more conditions and age groups.

The combination of identifying suitable apps and involving people living with the target condition to evaluate them is the core of this article. It could be argued that there are existing examples of app evaluation methods that are more rigorous [8] and/or larger in scale [14] than the current approach, however these do not directly incorporate the thoughts and opinions of the target user group. To achieve this it was decided that a balance needed to be struck between the work undertaken by the researchers in identifying suitable apps and the amount of work required by participants representing the user group to evaluate the identified apps. For example, it would be unrealistic to expect participants to directly test and give feedback on a high volume of apps (as in [14]), therefore the present approach advocates the narrowing of the pool to a more feasible number of apps by researchers prior to evaluation by users.

In the present day, technology usage and connectivity is increasing. This has meant that mobile apps are becoming more popular. Apps can be used for a number of purposes including the self-management of specific conditions. This can be achieved in many ways such as using apps for monitoring symptoms, medication reminders and games that engage people with complex needs. The results of this paper demonstrate that both YP diagnosed with ADHD and people living with dementia do enjoy using apps for a variety of functions, however, they have complex needs compared to those with a neurotypical profile and this was reflected in the participants' verbal and observed responses in both research projects.

Where the authors acknowledge the value of accounting for the opinions of app target audiences when developing apps, they also recognise that the method presented in this paper is not without its limitations. For example, unlike in one of the referenced examples evaluating diabetes self-management apps [8], the authors of the present method have not assigned rating scores for the quality of apps.. However, the authors of the ADHD example have provided a score for each included app, to reflect the app suitability for the population, based on the criteria suggested by their participants, and the author of the dementia example rates apps on their accessibility for the user group.

Future research could consider developing broader criteria that could be applied to apps aimed at those with complex long term conditions. In order to build on this, a toolkit could then be produced with generic criteria for apps and some condition specific advice to meet the more complex needs of these individual populations such as those living with dementia and ADHD.

Acknowledgement

To be added if accepted. This section has been temporarily removed to ensure anonymity, where possible, of the submission.

References

1. Internet Live Stats. (2016). Number of Internet Users. Retrieved January 16, 2017, from <http://www.internetlivestats.com/internet-users/>
2. Office for National Statistics. (2016). Internet users in the UK. Retrieved January 16, 2017, from <https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2016>
3. Powell, J. A., Darvell, M., & Gray, J. A. M. (2003). The doctor, the patient and the world-wide web: how the internet is changing healthcare. *Journal of the Royal Society of Medicine*, 96(2), 74–6.
4. Statista. (2016). Cumulative number of apps downloaded from the Apple App Store from July 2008 to September 2016 (in billions). Retrieved January 16, 2017, from <https://www.statista.com/statistics/263794/number-of-downloads-from-the-apple-app-store/>
5. Statista. (2016). Number of apps available in leading app stores 2016. Retrieved January 16, 2017, from <https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/>
6. Kratzke, C., & Cox, C. (2012). Smartphone Technology and Apps : Rapidly Changing Health Promotion Smartphones and Health Apps for Health Promotion Interventions Healthcare. *International Electronic Journal of Health Education*, 15(1), 72–82. doi:ISSN-1529-1944
7. Kamel Boulos, M. N., Brewer, A. C., Karimkhani, C., Buller, D. B., & Dellavalle, R. P. (2014). Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online Journal of Public Health Informatics*, 5(3). doi:10.5210/ojphi.v5i3.4814
8. Demidowich, A. P., Lu, K., Tamler, R., & Bloomgarden, Z. (2012). An evaluation of diabetes self-management applications for Android smartphones. *Journal of Telemedicine and Telecare*, 18(4), 235–238. doi:10.1258/jtt.2012.111002
9. Huckvale, K., Prieto, J. T., Tilney, M., Benghozi, P.-J., & Car, J. (2015). Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Medicine*, 13(1), 214.

doi:10.1186/s12916-015-0444-y

10. Max Planck Institute for Human Development and Stanford Center on Longevity. (2014). A Consensus on the Brain Training Industry from the Scientific Community. doi:10.1007/s00426-014-0564-6
11. Federal Trade Commission. (2016). Lumosity to Pay \$2 Million to Settle FTC Deceptive Advertising Charges for Its “Brain Training” Program. Retrieved January 17, 2017, from <https://www.ftc.gov/news-events/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges>
12. Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth*, 3(1), e27. doi:10.2196/mhealth.3422
13. Huckvale, K., Morrison, C., Ouyang, J., Ghaghda, A., & Car, J. (2015). The evolution of mobile apps for asthma: an updated systematic assessment of content and tools. *BMC Medicine*, 13(1), 58. doi:10.1186/s12916-015-0303-x
14. Nicholas, J., Larsen, M. E., Proudfoot, J., & Christensen, H. (2015). Mobile apps for bipolar disorder: A systematic review of features and content quality. *Journal of Medical Internet Research*, 17(8), e198. doi:10.2196/jmir.4581
15. Wallace, L. S., & Dhingra, L. K. (2014). A systematic review of smartphone applications for chronic pain available for download in the United States. *Journal of opioid management*, 10(1), 63–8. doi:10.5055/jom.2014.0193
16. Cailean, I. (2015). Extensive Guide to App Store Optimization (ASO) in 2015 - Part 1: Apple App Store. Retrieved January 25, 2017, from <http://www.trademob.com/app-store-optimization-guide-apple/>
17. Cailean, I. (2015). Extensive Guide to App Store Optimization (ASO) in 2015 - Part 2: Google Play Store. Retrieved January 25, 2017, from <http://www.trademob.com/app-store-optimization-guide-google/>
18. Joddrell, P., Hernandez, A., & Astell, A. J. (2016). Identifying Existing, Accessible Touchscreen Games for People Living with Dementia. In K. Miesenberger, C. Buhler, & P. Penaz (Eds.), *Lecture Notes in Computer Science* (pp. 509–514). Berlin: Springer International Publishing. doi:10.1007/978-3-319-41264-1_69
19. Powell, L., Parker, J., Robertson, N., & Harpin, V. (n.d.). ADHD: Is there an app for that? A suitability assessment of apps for children and young people with ADHD. *Journal of mHealth and uHealth*. (Under review)
20. Pope, C., Ziebland, S., & Mays, N. (2000). Analysing qualitative data. *BMJ*, 320(7227).
21. Groenewoud, H., Schikhof, Y., Astell, A. J., Goumans, M., & de Lange, J. (2014). iPad Happy Games for People With Dementia As Pleasant and Meaningful Activity. In *Alzheimer's & Dementia* (Vol. 10, p. 181). Elsevier. doi:10.1016/j.jalz.2014.04.201
22. Joddrell, P., & Astell, A. J. (2016). Studies Involving People With Dementia and Touchscreen Technology: A Literature Review. *JMIR Rehabilitation and Assistive Technologies*, 3(2), e10. doi:10.2196/rehab.5788
23. Astell, A. J., Joddrell, P., Groenewoud, H., de Lange, J., Goumans, M., Cordia, A., & Schikhof, Y. (2016). Does familiarity affect the enjoyment of touchscreen games for people with dementia? *International Journal of Medical Informatics*, 91, e1–e8. doi:10.1016/j.ijmedinf.2016.02.001

