A Comparison of Two Methods to Assess Mobile Hand-Held Communication Device Use

S. Berolo[1], I. Steenstra[2], B.C. Amick III[2,3,4], R.P. Wells[1,2,3]

[1]*Department of Kinesiology, University of Waterloo, Waterloo, Canada.* [1]*sberolo@uwaterloo.ca*

[2]*Institute for Work and Health, Toronto, Canada.*

[3]*Centre for Research Expertise for the Prevention of Musculoskeletal Disorders, University of Waterloo, Waterloo, Canada*

[4]*School of Public Health, University of Texas Health Science Center at Houston, Houston, USA*

Abstract

The purpose of this study was to: 1) examine the agreement between self-reported measures of mobile hand-held device use and direct measures, and 2) understand how respondents think about their device use when completing the questionnaire. 35 participants were recruited from a Canadian university to participate. The questionnaire was previously developed to estimate six types of usage potentially related to musculoskeletal disorders in cell phone users. Direct measures were collected on a custom software usage logging application resident on the participants' phone. Bland-Altman plots demonstrated that, generally speaking, participants' self-reports overestimated their logged usage; overestimates tended to be low at low mean usage times, and became more variable as mean usage time increased. Six themes were identified from the interviews: 1) types of mobile hand-held devices, 2) tasks attributed to different categories of device use, 3) their thought process used to arrive at usage, 4) ease of reporting usage, 5) physical interaction with device, and 6) completeness of list of categories of device use. The preliminary findings indicate that it may be challenging for respondents to provide good self-reports of usage. The root of this challenge does not appear lie with the design of the questionnaire; rather, it may be attributed to the respondents' difficulty in estimating usage, partly due to the variability of device use both within a day and a week.

Introduction

Despite concern that musculoskeletal symptoms are associated with mobile hand-held device use [2], there is a lack of knowledge concerning the use and exposures associated with these devices. To understand the use of mobile technology and determine if use is associated with any musculoskeletal or other health effects, accurate measures of how these devices are used is required. Previous research has examined agreement between self-reported computer, keyboard, and mouse use and direct measures of use collected using computer logging software or observation [3, 4], however no such literature exists comparing self-reported hand-held device use with directly measured use. Before progressing with further studies of hand-held device use, we thought it important to better understand the accuracy of the exposure measures collected by a questionnaire. The study described herein was designed to: 1) examine agreement between self-reports collected by a previously developed questionnaire [2] and direct measures of use collected using a custom software logging application, and 2) examine the usability of the questionnaire by conducting interviews to better understand how respondents think about their device use.

Methods

Data collection. 35 students/recent graduates (17 female and 18 male) under the age of 35 were recruited from a Canadian university to participate. Inclusion criterion included being an Android device user. There was no exclusion criterion. The study was approved by the Office of Research Ethics at the University of Waterloo, Ontario, Canada.

Mobile hand-held device use on a "typical" day last week was assessed using: 1) a paper and pencil version of the previously used questionnaire [2], and 2) the custom Android device logging software. The usage logging program recorded participants' hand-held device use on second-by-second basis. It recorded generic information on the applications used but did not record any personal or identifying information. For example, it would show that an email or messaging application was used, but would not display contents of emails or messages, and would not record voice or keys pressed. It also recorded the number and duration of phone calls made and received, and the number of text messages sent and received. It was developed for the Android platform due to the relative ease of accessing usage information on this operating system.

The questionnaire was previously developed to estimate six types of hand and upper limb usage, potentially related to musculoskeletal disorders among cell phone users. These six categories of use were chosen to represent different intensities of thumb and upper limb usage [2]. The questionnaire assessed how much time on a typical day last week participants spent using their mobile hand-held device(s) for: 1) emailing, texting, and instant messaging; 2) scheduling (calendar, appointments); 3) internet browsing; 4) making phone calls and talking on the phone; 5) listening to music, watching videos, and taking pictures; and 6) gaming using a mobile device. A "typical" day referred to time at work, as well as time away from work, at home, and with friends. Participants met the researchers on two occasions; on their first visit they were shown how to download the logger software onto their device, and on their second visit, one week later, they uploaded the data log from their device and filled out the questionnaire. Participants also provided demographic data and information regarding the make and model of their device(s). A subgroup of respondents, 17 out of 35, then participated in an interview.

Interviews were carried out with an aim to better understand of how respondents think about their device use. An *intensive interview* framework was used, and followed these general steps: 1) the respondent was reminded of the question and the response he/she gave, 2) the respondent was asked what he/she thought the question meant, 3) the respondent was asked to describe the thought process used to arrive at his/her answer, and 4) the respondent listened to the question, again, and was asked whether he/she would change his/her answer, and if yes, what he/she would change it to [1, 5]. Each interview took approximately 20 minutes.

Analysis. A program was developed using Visual Basic to summarize logged device usage. Each participant's data record was broken down by calendar day so average daily use could be compared to self-reported use on a "typical" day. There are currently over 400,000 Android applications available [6], and the applications used by each participant were quite different. Since the logged data was to be compared with self-reports of six kinds of usage, after the questionnaire was complete, participants were asked to categorize the applications recorded by the logging program into one of the six categories of device use specified in the questionnaire. A seventh category called 'other' was added to this mapping task for those applications that did not fit into one of the six categories. Using the results from the mapping task just described, the time spent performing each of the six measures of device use found in the questionnaire on each of the 7 days prior to the participants' second visit was calculated. Bland-Altman plots were used to examine agreement between logged and self-reported usage. The 17 interviews were transcribed. A preliminary coding scheme of major themes was developed by reading through all interviews once. "Themes" were defined as concept trends or patterns. The coding scheme was confirmed by two members of the research team upon detailed examination of each interview, and was used to block sections of each transcript that corresponded to a theme.

Results

Comparison of self-reported and logged device usage

Bland-Altman plots for categories of device use emailing, texting, and instant messaging (Figure 1) and gaming (Figure 2) are shown. These two categories of device use demonstrated strong relationship to self-reported pain in the upper extremity [2]. If there were perfect agreement between self-reported and

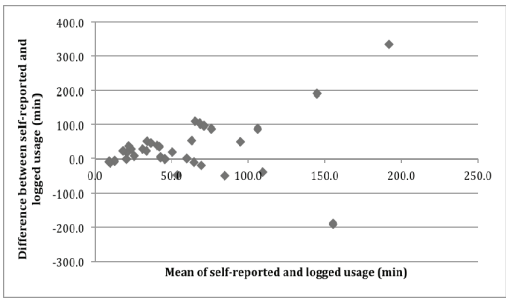

Figure 1. Bland-Altman plot for emailing, texting, and instant messaging on a "typical" day (N=35).

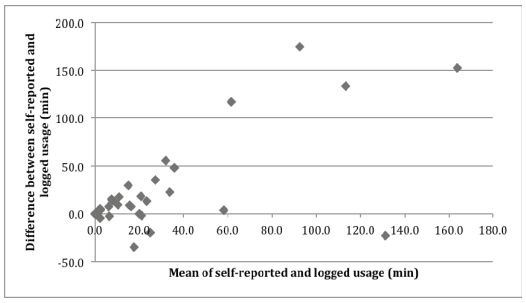

Figure 2. Bland-Altman plot for gaming on a "typical" day (N=35).

logged use, all data points in these plots would lie at 0 on the y-axis. Self-reports of zero usage time will lead to perfect agreement or an underestimation of logged use. In this population, generally speaking, for the six categories of device use, participants' self-reports overestimate their logged use; overestimates tend to be low at low mean usage times, and become more variable as mean usage time increases.

Themes from interviews

A total of six themes were identified from interviews: 1) types of mobile hand-held devices, 2) tasks attributed to different categories of device use, 3) their thought process used to arrive at usage, 4) ease of reporting usage, 5) physical interaction with device, and 6) completeness of list of categories of device use.

While all participants agreed those devices listed in the questionnaire are mobile hand-held devices (cell phones, personal digital assistants, and gaming devices), many had suggestions to be added to the list, tablets for example. A second theme was tasks participants attributed to each of the six categories of device use. For example, some participants considered use of social networking sites, like Facebook and Twitter, part of email, texting, and instant messaging, while others considered them part of Internet browsing. A third theme was the thought process to arrive at self-reported usage. A variety of processes were described by participants, including averaging, estimating based on numbers of events, thinking about time spent with or without a computer/laptop, and thinking about one's daily schedule. A fourth theme that surfaced was the difficulty some participants had in reporting device use on a typical day. One participant described this challenge by comparing phone use to computer use, and another explained it was difficult because each day was very different. A number of participants also spoke about the physical interaction they have with their device. For example, some participants indicated that they used both thumbs to type most of time, and others noted they do not use two thumbs for reasons related to device size and screen function.

Discussion

Participants' self-reports tended to overestimate their logged use; overestimates tended to be low at low mean usage times, and became more variable as mean usage time increased. This may suggest it is easier for respondents to provide self-reports that agree with logged usage when time spent performing a task is low. If respondents know, for example, that they do not perform a task, or that they spend little time performing a task, it may be more straightforward for them to provide a self-report than it would be if they performed that task more heavily. Some of the greater differences might be explained by inappropriate application matching. For example, the highest overestimate of logged use for gaming (Figure 2) likely results from the game Tetris being categorized by that participant as 'other' as opposed to 'gaming'. Such misclassification is a possible consequence of participants categorizing their applications; however, considering the number of applications available for download, the large variation in applications used by participants, and the fact that it is the agreement between self-reports

and logged usage under study, application matching was the best way to set up the logged data for comparison with self-reports.

Themes from interviews were useful in identifying possible reasons for differences between logged and self-reported usage. Interviews highlighted that there can be blurring of boundaries between categories of device use. Should a respondent use both a mail application and an Internet browser application to check email, for example, the boundary between emailing, texting, and instant messaging and Internet browsing would blur. Although logged usage from such logging software has not been demonstrated as a gold standard, *per se*, the application likely provides a more refined representation of use compared to self-reports. That said, like similar exposure models in computing work, it might not simply be actual usage that is critical in predicting health outcomes. For example, holding a hand over a mouse may produce significant muscle loading that does not get captured in direct measurement.

The comparisons made in this study between self-reported and logged device usage are preliminary and made among a small, restricted university sample. A questionnaire may be the most convenient way to assess use in epidemiological studies. These findings, however, suggest it may be challenging for respondents of a questionnaire to provide self-reports of device use that do not overestimate logged use. The root of this challenge does not appear lie within the design of the questionnaire; instead, it may be attributed to the variability of time spent using a device. It remains unknown which method of estimating device usage is better predictive of musculoskeletal symptoms.

Conclusions

Going forward, research investigating the relationship between mobile hand-held device use and musculoskeletal symptoms and other health outcomes may want to consider using a logging application to examine exposure simultaneously with self-reported exposure and observational approaches to better understand the sources of any hazardous exposures.

References

1. Belson, W.A. (1981). *The Design and Understanding of Survey Questions*. Hants, England: Gower Publishing Co. Ltd.

2. Berolo, S., Wells, R.P., Amick, B.C.III (2011). Musculoskeletal symptoms among mobile hand-held device users and their relationship to device use: A preliminary study in a Canadian university population. *Applied Ergonomics* **42**, 371-378.

3. Douwes M., de Kraker H., & Blatter B.M. (2007) Validity of two methods to assess computer use: Self-report by questionnaire and computer use software. *International Journal of Industrial Ergonomics* **37**, 425-431.

4. Mikkelsen S., Vilstrup I., Lassen C.F., Kryger, A.I., Thomsen, J.F., & Andersen, J.H. (2007). Validity of questionnaire self-reports on computer, mouse and keyboard usage during a four-week period. *Occupational and Environmental Medicine* **64**, 541-547.

5. Presser, S., Couper, M.P., Lessler, J.T., Martin, E., Martin, J., Rothgeb, J.M. Singer, E. (2004). Methods for Testing and Evaluating Survey Questions. *Public Opinion Quarterly* **68**, 109-130.

6. Wikipedia. Google Play. <http://en.wikipedia.org/wiki/Android_Market>. Published 6 March 2012. Accessed 28 March 2012.