

RESEARCH ARTICLE

Smartphone-based evaluations of clinical placements—a useful complement to web-based evaluation tools

Jesper Hessius^{1*}, Jakob Johansson²¹Faculty of Medicine, Uppsala University, Uppsala, Sweden; ²Department of Surgical Sciences, Uppsala University, Uppsala, Sweden**Abstract**

Purpose: Web-based questionnaires are currently the standard method for course evaluations. The high rate of smartphone adoption in Sweden makes possible a range of new uses, including course evaluation. This study examines the potential advantages and disadvantages of using a smartphone app as a complement to web-based course evaluation systems. **Methods:** An iPhone app for course evaluations was developed and interfaced to an existing web-based tool. Evaluations submitted using the app were compared with those submitted using the web between August 2012 and June 2013, at the Faculty of Medicine at Uppsala University, Sweden. **Results:** At the time of the study, 49% of the students were judged to own iPhones. Over the course of the study, 3,340 evaluations were submitted, of which 22.8% were submitted using the app. The median of mean scores in the submitted evaluations was 4.50 for the app (with an interquartile range of 3.70-5.20) and 4.60 (3.70-5.20) for the web ($P = 0.24$). The proportion of evaluations that included a free-text comment was 50.5% for the app and 49.9% for the web ($P = 0.80$). **Conclusion:** An app introduced as a complement to a web-based course evaluation system met with rapid adoption. We found no difference in the frequency of free-text comments or in the evaluation scores. Apps appear to be promising tools for course evaluations.

Key Words: *Clinical education; Course evaluation; Education evaluation; Smartphone based tools*

INTRODUCTION

There is a long tradition in education of using course evaluations to generate feedback on how students perceive their courses and to improve the quality of their education [1]. Response rates are important for reliable evaluations, but achieving a high response rate can be difficult. One factor that has been shown to have considerable impact on the response rate is the ease of use of the evaluation system [2]. Traditionally, evaluations have been paper-based, distributed by the teaching institution to be filled in by the students taking the courses concerned. Previously, the impact on the quality of the information gathered during the transition from paper-based to Internet-based course evaluations have been studied [3]. A

number of studies investigated such factors as response rates and average evaluation scores [3-5], and, generally speaking, it was found that Internet-based tools tended to have lower response rates but little or no effect on the average scores attained [3-6]. There were fears that the transition from paper to web-based evaluations would lead to the submission of fewer free-text comments, but studies have shown the opposite to be the case [6,7].

Since the introduction of web-based evaluations, several other evaluation formats have been tested, including the use of short messaging services [8] and personal digital assistants, so called 'personal digital assistants (PDAs)', a type of handheld personal computer popular before the advent of the smartphone [9]. Then, in only a couple of years, the information technology landscape went through a seismic shift as the first smartphones appeared on the market and rapidly became popular. In 2012, it was estimated that 86% of Swedish adults aged between 15 and 64 had a smartphone. Sweden was then the country with the third-highest smartphone penetration in

*Corresponding email: jesper.hessius@ltv.se

Received: November 5, 2015; Accepted: November 29, 2015;

Published online: November 30, 2015

This article is available from: <http://jeehp.org/>

the world [10]. The permanent access to the Internet offered by smartphones has led to the use of mobile applications, or 'apps,' in a growing range of situations. Among medical students, apps are extensively used, both for learning and to help in the student-patient encounter [11]. However, studies of technical solutions developed primarily for mobile devices are notable by their absence; indeed, to our knowledge, there are no reports in the literature on the use of apps for course evaluations. A tool developed specifically as a mobile complement to a proven web-based evaluation system should make existing course evaluation questionnaires more accessible, and would thus help improve the courses' quality assurance. We have developed just such an app in order to investigate whether the quality of the evaluations is affected by the new format.

METHODS

Materials

The study looked at evaluations of courses at the Faculty of Medicine at Uppsala University, Sweden between August 2012 and June 2013, and took the form of a 'data-mining' statistical analysis of quantitative evaluations drawn from the database of the university's 'KlinikKurt' course evaluation system. The medical training at Uppsala University consists of 11 terms, of which 5.5 terms are spent in clinical placements, where students receive specialised training in approximately 30 different clinical fields. In the autumn term of 2012, 499 students were enrolled in the clinical programme, while in the spring term of 2013 there were 509.

Criterion for inclusion

- (1) Questionnaires submitted to KlinikKurt during the autumn term of 2012 or the spring term of 2013.

Criteria for exclusion

- (1) Questionnaires that failed to note the location or nature of the placement.
- (2) Questionnaires in which every question with a Likert scale was left unanswered.
- (3) Questionnaires that were identical to a previously submitted questionnaire in terms of location, placement, and anonymous user ID.

Technical information

Uppsala University's Faculty of Medicine has its own web-based tool, KlinikKurt, which students use to evaluate all their clinical placements during training [12,13], whether at Uppsala University Hospital or other hospitals in central Sweden. It takes the form of a questionnaire with ten questions, each of which reflects a different aspect of clinical supervision, answer-

ed using a six-point Likert scale. Each question has a field for optional, free-text comments. Students are expected to submit a course evaluation for each of their clinical placements. Evaluations may be submitted at any point from the start of term until approximately a week after term ends. The evaluations are submitted anonymously.

A mobile app for submitting evaluations to the KlinikKurt system was developed. The app was designed to communicate with the same database as the web-based tool, and differences in the way the two tools worked were minimised. The development framework Apache Cordova/Adobe Phonegap was chosen. This allowed the functionality of the web-based tool to be duplicated as closely as possible, and will simplify later conversion to other mobile platforms. It also meant the app could be written largely using the same programming language as the web-based tool. To assist interface design and asynchronous communication with the script that enters evaluations into the KlinikKurt database, the app was based on jQuery Mobile. However, significant modifications were made to the built-in user interface so that the app resembled the KlinikKurt homepage, to ensure that users familiar with the web-based tool would feel at home using the app (Fig. 1).

The user interface for the app was designed to resemble the radio buttons of the web-based tool but with a number of adaptations to allow for the smaller screen size—for example, the text and buttons were made larger (Fig. 2). Unlike the web-based tool, the app's user interface was designed to be active even when Internet access was not available, enabling the user to fill in evaluations offline, although not to submit them. To enable us to distinguish between evaluations sent in from the app and those from the web-based tool, a variable, hidden from the user, was incorporated into both tools to indicate which had been used to submit the evaluation.

All the data used in the statistical analysis were drawn from the database of responses submitted to KlinikKurt. Throughout the two terms in question, students were encouraged to evaluate each completed clinical placement using either www.klinikurt.se or the KlinikKurt app, which was available for iOS free of charge from the App Store digital distribution platform. To advertise the launch of the app, posters were put up in the student canteen and in the hospital research library in the autumn term of 2012. To gauge how many students had an iPhone, and were thus able to use the app, a visit was paid to each course at the end of the autumn term or the beginning of the spring term. Wherever possible, a lecture or similar event with obligatory attendance was chosen. The question 'Do you have an iPhone?' was put to the students, making it clear that the question excluded other smartphones and other mobile devices such as iPads.



Fig. 1. The homepage of the web based tool (A) and the app's opening screen (B).

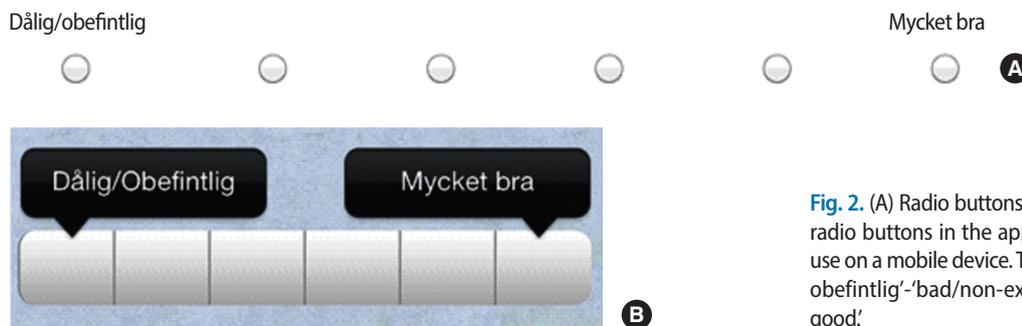


Fig. 2. (A) Radio buttons in the web-based tool and (B) radio buttons in the app, adapted for greater ease of use on a mobile device. Translation of end points: 'dålig/obefintlig'-'bad/non-existent' and 'mycket bra'-'very good'.

Statistics

All the evaluations satisfying the above criteria from the autumn term of 2012 and spring term of 2013 were combined into a single data set, and were then divided into two populations according to whether they were submitted using the app or the web-based tool. The number of unique user IDs was recorded for the two sub-populations and for the data set as a whole. For each evaluation that met the inclusion criteria, an 'evaluation score' was calculated by taking the mean of its Likert scale values. Questions left unanswered did not contribute to the score. Because the evaluation scores were not normally distributed (Fig. 3), the median was calculated for the scores from each sub-population using Microsoft Excel 2010, and this is presented together with the interquartile range. Signifi-

cance in differences between evaluation scores in the respective populations was tested with a Mann-Whitney U-test, from which the P-value is presented, calculated using STATISTICA 64 (StatSoft Inc., Tulsa, OK, USA).

The frequency of free-text comments for each sub-population was calculated by subtracting the number of blank comment fields from the total number of available fields in the population's evaluations. The frequency of complete evaluations—those in which every question with a Likert scale had been answered—was calculated by subtracting the number of responses containing at least one blank answer from the total number of evaluations in the population using Microsoft Excel 2010. The frequency of evaluations submitted during the final week of the evaluation period was calculated by separat-

Table 1. Collated statistics for course evaluations submitted using the app and the web

Variable	App	Web	P-value
Proportion of total evaluations	22.8%	77.2%	
Median evaluation score (and interquartile range)	4.50 (3.70-5.20)	4.60 (3.70-5.20)	0.24
Proportion of evaluations with free-text comments	50.5%	49.9%	0.80
Proportion of complete evaluations	98.9%	96.1%	<0.001
Proportion of evaluations submitted in the final week	16.2%	26.3%	<0.001

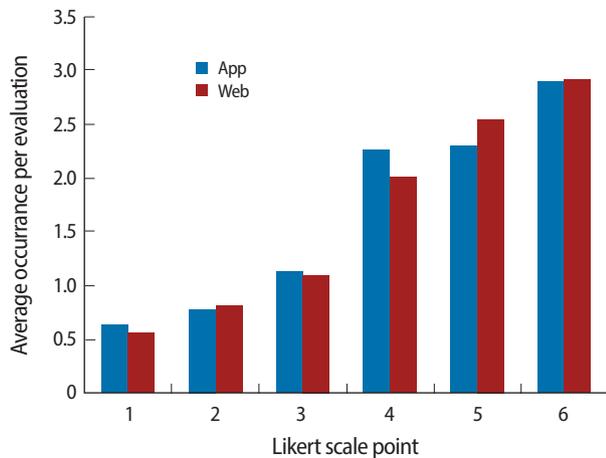


Fig. 3. Distribution of responses on the Likert scale, averaged over all questions answered per evaluation.

ing the evaluations in the two sub-populations according to the term in which they were submitted, and counting the number of evaluations in each sub-population submitted during the relevant final seven days when submission was possible. The P-values for the differences between the groups were calculated using a chi-square test using STATISTICA 64 (StatSoft Inc.); $P < 0.05$ was chosen as the significance level.

RESULTS

In the autumn term of 2012, 1,734 course evaluations were submitted; in the spring term of 2013, 1,606. The study thus comprises a total 3,340 course evaluations. At the change of term, 49% ($n = 175$) of the students surveyed answered that they had an iPhone. The investigation obtained responses from 357 (56%) of the 638 students enrolled during the terms covered by the study. In the period in question, 22.8% ($n = 761$) of evaluations were submitted using the app and 77.2% ($n = 2,579$) using the web-based tool. The median evaluation score was 4.50 (with an interquartile range of 3.70-5.20) for evaluations submitted using the app, and 4.60 (3.70-5.20) for the web-based tool ($P = 0.24$) (Table 1).

The average frequency of free-text comments for all the submitted evaluations was 50.1% ($n = 1,672$). Of evaluations sub-

mitted using the app, 50.5% ($n = 384$) included comments, while the corresponding figure was 49.9% ($n = 1,288$) for the web-based tool ($P = 0.80$). The proportion of complete evaluations submitted using the app was 98.9% ($n = 753$), against 96.1% ($n = 2,479$) for the web ($P < 0.001$). In total, the proportion of evaluations submitted in the final week was 24% ($n = 801$). This proportion was 16.2% ($n = 123$) for the app and 26.3% ($n = 678$) for the web ($P < 0.001$) (Table 1). The distribution of the scores on the Likert scale was similar for the app and the web-based tool (Fig. 3). Of the enrolled students, women comprised 56.5% of the total in the autumn term of 2012 and 57.7% in the spring term of 2013. Of the evaluations submitted via the app, respondents identified themselves as women in 60.2% ($n = 458$) of cases, compared to 56.2% ($n = 1,450$) for the evaluations submitted using the web-based tool ($P = 0.16$).

DISCUSSION

We developed an iPhone app as a complement to a web-based tool for evaluating clinical placements. In the two terms covered by the present study, just under half of the enrolled medical students had an iPhone. Evaluations submitted using the app made up nearly a quarter of the total. When it came to average scores or the number of free-text comments, the evaluations submitted using the app were no different from those submitted using the web. Furthermore, course evaluations submitted using the app were more often complete than those submitted using the web. Compared to the transition from paper to web-based evaluations, the challenges posed by the introduction of app-based system are somewhat different, because it is designed to be used in parallel with a web-based system. This means that there is an even greater need for consistent results. In this study, we found that the evaluation scores in evaluations submitted using the app and the web were very similar, as were the distributions of scores on the Likert scale. In the app-based responses there are signs that the end-point scores were being avoided—the so-called central tendency [14]. However, this avoidance is such a weak effect, that, taking the other findings into account, we judge it not to have any significant influence on the quality of the course evaluations.

In addition to a high response rate, it is best if evaluations

are complete. If respondents tend not to answer all the questions, the reliability of the results will be invisibly reduced, exactly as if the response rate for the course evaluation as a whole were low. Obviously, there can be legitimate reason for students to choose not to answer all the questions; however, there is a danger that this kind of reduced reliability might arise by mistake. For example, respondents do not realize that they have left questions unanswered. Students may not answer because they become bored of filling in the questionnaire. For example, the evaluation is too long or it lags when used. Something similar was observed in a study of the use of a short messaging service-based mobile phone tool [8], for although the course evaluation in that study only involved five questions, 20% of respondents left the last three questions unanswered. In our study, we did not observe respondents failing to answer the later questions. On the contrary, we have shown that app-based responses were more often complete than those submitted using the web. It should be pointed out that both groups had a very high proportion of complete evaluations.

In the transition from paper to web-based course evaluations, it was feared that students would submit fewer free-text comments; however, the opposite proved to be the case [6,7]. Similar apprehensions were expressed by both students and faculty before the app-based option was introduced. One study of evaluations submitted to a web-based evaluation system from mobile devices has to some extent confirmed these fears, for it was found that evaluations sent in from mobile devices had somewhat fewer 'meaningful comments' than did evaluations submitted using computers [15]. In addition, it should be noted that a study of the use of short messaging services as an evaluation tool also found a reduction in the number of responses to questions requiring free-text answers [8]. Thus there was a suspicion that using a mobile phone might predispose respondents to be unwilling to provide comments. Our results show that there is no difference in the frequency of free-text comments made with the app or the web-based tool. However, we have not investigated the length of the comments, nor have we attempted to judge their quality.

It is reasonable to suppose that evaluations submitted by respondents directly after completing a course unit are more likely to be accurate and specific. It has previously been shown that the earlier they are submitted, the greater the positive effect on what is judged to be the quality of comments [16]. However, with our current system, it is impossible to record how soon an evaluation was submitted after a clinical placement without compromising anonymity. As a substitute, we chose to look at the proportion of course evaluations submitted during the final week of the evaluation period, as it was thought likely that these would have been submitted with a longer delay following the end of the course unit in question. Our re-

sults show that fewer evaluations were submitted using the app than the web in the final week of each evaluation period. This implies that app-based evaluations were submitted with less delay after the course element they referred to had ended.

In terms of evaluation score and the frequency of free-text comments, we find that evaluations submitted using an app are no different from those submitted using the web, and the distribution of scores is similar for both tools too. Thus the app used in the present study meets the requirement for high consistency with the web-based system it is intended to complement. We would therefore argue that the introduction of an app alongside an existing web-based course evaluation tool provides a number of clear advantages, especially in the form of increased accessibility, and no obvious disadvantages. The app continues to be developed, and we shall monitor how it is used in future. An Android version of the app was introduced for the autumn term of 2013.

The present study suffered from a number of limitations. We do not know the actual number of iPhones owned and used by respondents. The sampling at the halfway mark reached 56% of those enrolled in the courses in question. This means that the true proportion of students with iPhones could, theoretically, lie between 23% and 72%. In addition, the number of iPhones almost certainly varied with time. During the first term of the study, a new iPhone model went on sale, so there is reason to believe that more students owned an iPhone at the halfway mark (after Christmas) than had the preceding autumn term. At sampling, the gender of the iPhone users was not considered; therefore, it is therefore hard to draw conclusions about whether the larger proportion of women respondents reflects a preference for the app, or just iPhone ownership. In addition, we chose to investigate the frequency of free-text comments, but not their quality, largely because of the difficulty of measuring this. We decided that using a parameter such as the length of the comments would be inadequate—a long comment is not necessarily better than a short one. A tool does exist for the qualitative analysis of free-text comments [17], but the drawbacks are that it is not used in any other course evaluation context at the medical school and so has no proven track record, while it is not available in a Swedish-language or Swedish-localized version.

In conclusion, the introduction of an app to complement a web-based course evaluation system met with rapid adoption. Course evaluations submitted using the app did not differ significantly from those submitted using the web-based tool when it came to evaluation scores or the incidence of optional free-text comments. Course evaluations submitted using the app were more likely to be complete than were those submitted using the web-based tool. We conclude that apps promise to be useful tools for course evaluations.

ORCID: Jesper Hessius: <http://orcid.org/0000-0002-7596-6577>;
Jakob Johansson: <http://orcid.org/0000-0003-4248-4421>

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

Data from the autumn term of 2012 from this study was presented at the 2013 Association of Medical Education in Europe conference as an ePoster titled “Using a mobile application as a supplemental tool for student evaluations: adding ievaluation to e-evaluation” and is available for viewing online at http://ieposter.com/eposter_viewer/340/using-a-mobile-application-as-a-supplemental-tool-for-student-evaluations-adding-ievaluation-to-e-evaluation.html. We would like to thank all the students who participated in the course evaluations of their clinical placements, not only for helping this study, but also for making the Uppsala University Medical Programme’s clinical teaching even better in the process.

SUPPLEMENTARY MATERIAL

Audio recording of abstract.

REFERENCES

1. Aleamoni LM. Student rating myths versus research facts from 1924 to 1998. *J Pers Eval Educ*. 1999;13:153-166.
2. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? *Assess Eval High Educ*. 2008;33:301-314. <http://dx.doi.org/10.1080/02602930701293231>
3. Carini RM, Hayek JC, Kuh GD, Kennedy JM, Ouimet JA. College student responses to web and paper surveys: does mode matter? *Res High Educ*. 2003;44:1-19.
4. Avery RJ, Bryant WK, Mathios A, Kang H, Bell D. Electronic course evaluations: does an online delivery system influence student evaluations? *J Econ Educ*. 2006;37:21-37. <http://dx.doi.org/10.3200/jece.37.1.21-37>
5. Leung DY, Kember D. Comparability of data gathered from evaluation questionnaires on paper and through the internet. *Res High Educ*. 2005;46:571-591. <http://dx.doi.org/10.1007/s11162-005-3365-3>
6. Anderson HM, Cain J, Bird E. Online student course evaluations: review of literature and a pilot study. *Am J Pharm Educ*. 2005; 69:34-43. <http://dx.doi.org/10.5688/aj690105>
7. Donovan J, Mader CE, Shinsky J. Constructive student feedback: online vs. traditional course evaluations. *J Interact Online Learn*. 2010;9:283-296.
8. Struthers J, Irvine P, Jackson C. Can text messaging be used as an effective method for collecting quality and detailed evaluation data from students on clinical placements? *Med Teach*. 2013;35: 678-683. <http://dx.doi.org/10.3109/0142159X.2013.801549>
9. Coulby C, Hennessey S, Davies N, Fuller R. The use of mobile technology for work-based assessment: the student experience. *Br J Educ Technol*. 2011;42:251-265. <http://dx.doi.org/10.1111/j.1467-8535.2009.01022.x>
10. Farago P. iOS and android adoption explodes internationally [Internet]. Flurry Insights Blog. 2012 [cited 2015 Aug 27]. Available from: <http://blog.flurry.com/bid/88867/ios-and-android-adoption-explodes-internationally>
11. Payne KB, Wharrad H, Watts K. Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC Med Inform Decis Mak*. 2012;12:121. <http://dx.doi.org/10.1186/1472-6947-12-121>
12. Kihlberg P, Perzon M, Gedeberg R, Blomqvist P, Johansson J. Uniform evaluation of clinical teaching: an instrument for specific feedback and cross comparison between departments. *Hogre Utbildning*, 2011;1:139-150.
13. Lewiss N, Hellgren L, Johansson J. Quality improvement in clinical teaching through student evaluations of rotations and feedback to departments. *Med Teach*. 2013;35:820-825. <http://dx.doi.org/10.3109/0142159X.2013.802302>
14. Berk RA. The secret to the “best” ratings from any evaluation scale. *J Fac Dev*. 2010;24:37-39.
15. Champagne MV. Student use of mobile devices in course evaluation: a longitudinal study. *Educ Res Eval* 2013;19:636-646. <http://dx.doi.org/10.1080/13803611.2013.834618>
16. Johnson TD. Online student ratings: will students respond? *New Dir Teach Learn*. 2003;2003:49-59. <http://dx.doi.org/10.1002/tl.122>
17. McNulty JA, Gruener G, Chandrasekhar A, Espiritu B, Hoyt A, Ensminger D. Are online student evaluations of faculty influenced by the timing of evaluations? *Adv Physiol Educ*. 2010;34:213-216. <http://dx.doi.org/10.1152/advan.00079.2010>