Short-Term Methodology for Long-Term Usability

David G. Novick, Baltazar Santaella, Aaron Cervantes, Carlos Andrade
Department of Computer Science
The University of Texas at El Paso
El Paso, TX 79968
+1 915-747-5480

novick@utep.edu, bsantaella@utep.edu, arcervantes@miners.utep.edu, candrade@gmail.com

ABSTRACT

Approaches to understanding usability of computer interfaces over the long term typically rely on longitudinal studies, which are limited in scope to the period of the experiment. In this study, we explore whether a non-longitudinal, cross-sectional approach can reliably detect useful differences in usability between novices and experts. Our approach takes a "snapshot" of usability problems and behaviors across a heterogeneous sample of users, ranging from novice to expert. Our analysis suggests that a cross-sectional methodology can distinguish between less experienced and more experienced users with respect to the kinds of applications that cause frustration, frequency of use of help, and whether the problem was solved. Our analysis also suggests that the method is poor at distinguishing causes of frustration and the overall distribution of types of solutions tried. The data also suggest that three months of use of an application is the most useful point at which to distinguish less-experienced from more-experienced users.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Evaluation/methodology, training, help, and documentation.*

Keywords

Usability, time, methodology

1. INTRODUCTION

In the last five years, the research community has turned increasing attention to the issue of usability over time, based on the insight that the problems encountered by experienced users of a user interface may not be the problems discovered by the new users who typically serve as subjects in usability tests. A special interest group at CHI 2007 [13] on capturing longitudinal usability was followed by a panel presentation at CHI 2008 [15] on the methodology of longitudinal usability data collection. CHI 2010 hosted a special interest group on best practices in longitudinal research [2]. And this year brought a workshop on theories, methods and case studies of longitudinal research [4]. Longitudinal studies have shown that, as users gain experience with a computer program, they encounter different kinds of problems and try to solve these problems in different ways. And longitudinal methodologies address significant shortcomings relative to other methodologies for understanding usability over time. But they are inherently limited in scope, so to speak, to no more than the length of each study's period [16]. How, then, can people looking at longitudinal issues break past the time limits of longitudinal methodologies to study users' experiences over much longer periods of time?

In this paper, then, we review the development of methodologies for studying usability over time, looking at both longitudinal and cross-sectional methods. We review analytical frameworks for the study of long-term usability and discuss objections to cross-sectional approaches. We propose a methodology that uses a cross-sectional design and contemporaneous evaluation. We test this "snapshot" methodology using the natural experiment produced by people using computers in their everyday lives. We analyze the results of the study, with particular attention to seeing if the snapshot method can find differences between less-experienced and more-experienced users with respect to interaction behaviors such as use of help and abandonment of tasks. Finally, we discuss the implications of our results for understanding the time period in which users change from novices to experts and discuss the study's methodological limitations.

SIGDOC'12, October 3–5, 2012, Seattle, Washington, USA.

2. BACKGROUND

Longitudinal studies, which take place over time, contrast with cross-sectional studies, which take place at one point in time but with participants with different levels of experience. In other words, longitudinal studies can be seen as a within-subjects methodology and cross-sectional studies as a between-subjects methodology. With respect to longitudinal methodologies, two different methodological paradigms have been proposed. One paradigm [16] distinguishes: *micro* studies, which are typically short-term usability tests; *meso* studies, which look at users over a period of weeks or months; and *macro* studies, which look at users over years or even the program's entire life-cycle. An alternate paradigm [5] differentiates methodologies in terms of when the data are collected relative to the users' experience: repeated sampling studies, which use a pre-and a post-test; longitudinal studies, which collect multiple user experiences as they occur; and retrospective studies, which collect multiple user experiences from memory at the end of the study's period.

2.1 Medium-term studies

Researchers have reported successful several meso studies in recent years. The modern line of longitudinal research into usability of user interfaces begins with Mendoza and Novick's 2005 research on usability over time [7], which addressed the issue that usual usability testing may actually reveal problems of novice users and

of learnability than of underlying problems that would frustrate experienced users. The study examined usability issues among middle-school teachers creating Web sites, for both the use of documentation and the underlying software, tracking the causes and extent of user frustration over eight weeks. The authors found that, over the eight weeks, the users' level of frustration dropped, the distribution of causes of frustration changed, and the users' responses to frustration episodes changed. These results suggested that the sorts of errors that are most prominently featured in conventional usability testing may not be significant over longer periods of time.

Subsequent meso studies included iPhone usability over a period of five weeks [3] and Web-based homework application over a semester [6]. The results of these studies reinforced the conclusion that that users' initial experiences with an application may not reliably predict their experiences in prolonged use.

2.2 Longer-term studies

But what about usability over longer periods of time? The necessary limitations of time and resources mean that few if any macro studies have been conducted [16]. Repeated sampling studies would be impractical, as researchers would have to conduct pre- and post-tests years apart, and would have data for only two time points for their efforts. Longitudinal studies would be equally impractical, as the researchers would have (a) to support a cohort of subjects over years, while the subjects may want to change to a different version or a different program, and (b) to continue the single study for years, perhaps with negative results. And retrospective studies, despite some evidence of effectiveness for medium-term time frames [16, 5, 12], are either relatively short-term or suffer from the distance of memory to the extent that the study seeks longer-term results. A study of the relative reliability of usability observation methodologies suggested that that retrospective approaches, such as interviews and surveys, provide less reliable views of users' problem-solving behaviors than do contemporaneous approaches such as participatory evaluation and direct observation [9].

Two proposed approaches attempt to address the problem of creating short-term methodologies for long-term usability. The Always-On+Adoption approach [1] in effect aims to reduce the time needed for a longitudinal study by speeding up the rate of initial use. While likely a useful technique for many applications, Always-On+Adoption can only reduce the overall time of the longitudinal study to the extent that it shortens the time of the users' early use. As a result, it may not produce enough speed-up to be practical for years-long macro studies, and results from this approach have not yet been reported. A second approach, UX Curve [11], is a retrospective methodology in which the user draws a curve that describes how the program's user experience has changed over time; while drawing the UX curve, the user explains the reasons behind the changes in the curve. UX Curve is similar to critical incident analysis but focuses explicitly on the temporal progression of the user experience. Using this approach enabled researchers to obtain qualitative experience data in two hours for six months of use of a program. But like critical incident analysis, UX Curve is limited by the effectiveness of the users' recall. Salience of memory may reflect the importance of incidents, but retrospective techniques appear to be less reliable than contemporaneous techniques [9] and, for really long-term studies, users may not be able to recall incidents that were truly critical from a design standpoint.

2.3 The "Snapshot" Approach

Given the impracticality of traditional longitudinal methodologies for macro-scale studies of usability, and given the relative reliability of contemporaneous methodologies over retrospective methodologies, a practical methodology for the study of long-term usability would be a cross-sectional methodology with contemporaneous recollection.

Cross-sectional methodologies have been criticized as vulnerable to under-controlled variation among users, leading to false attributions of effects to the variation in time rather than variation among users [5]. This criticism relies principally on the work of Prümper *et al*. [10]. While Prümper *et al*. observed that changes in the definition of experience (e.g., overall experience with computers vs. specific experience with a program) led to different usability results, their study nevertheless showed that it would be possible, given a particular definition of experience, to derive meaningful results with a cross-sectional methodology. And in fact, Prümper *et al*. relied on such a methodology to report the substantive results of their study. This suggests less that cross-sectional methodologies are inherently unreliable and more that researchers should take care to specify the experience perspective through which they interpret their results.

Actually, longitudinal studies suffer from exactly the same problem identified by Prümper et al. because the studies' subjects do not remain constant with respect to use characteristics. That is, it is true that the subjects gain experience over time with respect to the application that is the target of the study, but it is also true that they gain experience with computers generally, too. Along the same lines as Heraclitus's observation that no-one steps into the same river twice, one can note that no longitudinal survey samples the same users twice. Moreover, longitudinal studies are inherently subject to this problem, because the researcher cannot control for the experience perspective.

Longitudinal studies track the same individuals, across common stages of experience, while cross-sectional studies look at different individuals over different stages of experience. Longitudinal studies track use of the same application, while cross-sectional studies can look at use of multiple applications. Indeed, one can view the whole world of computing as a natural experiment using a cross-sectional design. That is, at any moment in time, for almost any given application there are many—possibly millions—of users, each of whom has his or her own level of experience with that application. This level of experience could range from a few minutes to several decades, which would enable macro-scale studies. So instead of tracking users over time and waiting for years, one could take advantage of the world's natural experiment to gather data on a set of users who, right now, have different levels of experience.

Accordingly, we implemented and used such a technique, which effectively takes a "snapshot" of use by multiple users of different levels of experience, with their frustration episodes reported through contemporaneous participatory evaluation rather than retrospection. This method is between-subjects and asks the users to report frustration episodes at the time the episodes occur. In this way, we suggest, researchers looking at longitudinal issues can break past the time limits of longitudinal methodologies to study users' experiences over much longer periods of time. In the sections that follow, we describe our approach, report results of an initial study, and discuss whether this technique for looking at usability over longer periods of time usefully adds to the understanding of usability issues.

3. METHODOLOGY

The key idea of our approach is that users self-report (via Web-based surveys) their experiences with computer technology as these experiences occur. Subjects fill out an initial survey with demographic and experiential information, and then fill out incident surveys when they encounter frustration with computer programs.

Thus we asked users first to fill out an initial survey online about themselves and their experience with technology. This survey included information about varieties of experience with computers and programs, along the lines studied by Prümper *et al*. Pre-testing suggested that the initial survey took less than ten minutes to complete.

We solicited subjects via social networking and asked subjects to recruit their acquaintances. Over a period of 25 days, 71 subjects (excluding apparent duplicates) completed the demographic survey. The subjects lived in ten U.S. states and two other countries, with most subjects from the area of El Paso, TX, other areas of Texas, and the Pacific Northwest. As indicated in Table 1, subjects' median age was 30-39.

Table 1. Age of subjects.

N	Age
1	18 to 24
4	25 to 29
32	30 to 39
14	40 to 49
13	50 to 59
5	60 to 69
1	70 +

The subjects' experience with computers ranged from 8 to 40 years, with a median of 20 years. Their self-evaluated proficiency ranged from our scale's minimum (1) to the maximum (5), with a mean self-evaluated proficiency of 3.93.

Subjects were then asked to fill out at least one, and possibly more, incident reports during the next three days, each time they encountered a frustrating experience in using computer technology, using a Web-based form. The incident report form included questions about the user's specific extent of experience with the program involved in the report. Pre-testing suggested that the experience report also took less than ten minutes to fill complete. Each demographic survey had a unique 13-digit identifier that was the only way to connect demographics to incident reports, thus guarding anonymity of the subjects.

The experience report comprised an introduction and seven main questions. The categories of solution methods in Question 7 were based on the twelve categories reported in [8], although as noted in Section 4, our analysis collapsed these to four categories. The experience report asked these questions:

Below please describe a frustrating experience that you had recently in using a computer. We will combine this information with that of others to aid us in designing and improving future technology help systems for users like you.

Now think about a recent experience using a computer that frustrated you.

In your own words please share the experience.

1) What software or application were you using that caused the frustration?
   - Approximately for how long (days, months or years) have you been using this application?
   - On average, how many days a month do you use this application?
2) Please explain what you were attempting to do.
3) What caused your frustration?
4) Could you identify the problem?
   - Yes
   - No
   - Unsure
   - Do not remember

   Please explain.
5) Were you able to solve the problem?
   - Yes
   - No
   - Unsure
   - Do not remember

   Please explain.
6) Could you identify other possible solutions?
   - Yes
   - No
   - Unsure
   - Do not remember

   Please explain.
7) What tools or options did you use in trying to solve the problem?
   - I read the printed manual
   - I used the electronic manual
   - I used the help feature in the application/program
   - I searched the web for a solution
   - I searched the online knowledge base of the application/program
   - I asked someone I know for help
   - I asked for help online
   - I found a workaround or alternative method
   - I found a solution via trial and error
   - I talked with the technical support department for the application/program
   - I gave up
   - Other:

Do you have any additional comments on the frustrations you have experienced with technology or help systems?

Do you have any comments on this survey?

4. RESULTS

The study's subjects reported 41 unique frustration episodes. The subjects reported experience levels with the application that ranged from two days to eleven years. Figure 1 presents the distribution of experience levels with the application. The x-axis is the upper limit of the time interval, in years, for the episodes, and the y-axis is the number of episodes for that time interval. For example, there were six episodes where the subjects had more than four and up to eight years of experience with the application.

We now examine the time-series results for the kind of software the subjects used, the nature of the frustration episodes they reported, the solution methods they tried, and the extent to which they used help resources.

4.1 Software Involved in Frustration Episodes

As the study's subjects reported frustration episodes that occurred with the software they were using, we first examined differences between less-experienced and more-experienced users with respect to the kind of software they used and which, because of the study's design, led to the frustration episode. We defined less-experienced as having three months or less experience with the software. As indicated in Figure 2, the less-experienced users experienced problems with different kinds of software than did experienced users, and this difference in software was significant (ChiSq < 0.001).

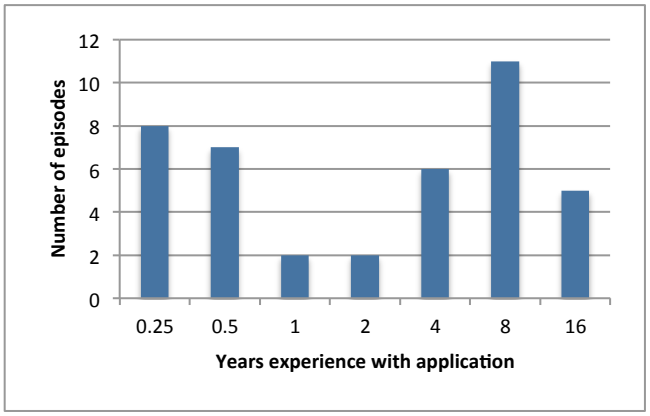

Figure 1. Distribution of self-reported experience levels with the application involved in users' frustration episodes.

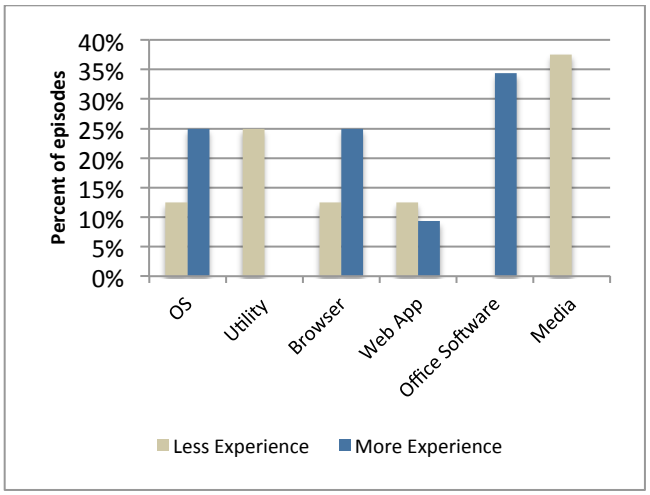

Figure 2. Differences between less-experienced and more-experienced users with respect to software involved in frustration episodes.

Less-experienced users had more problems with utility and media software, and more-experienced users had more problems with operating systems, browsers, and office software.

4.2 Nature of Frustration Episodes

Earlier studies (e.g., [8]), reported changes over time in the relative frequencies of different kinds of frustration episodes. These kinds of results are perhaps the primary goal of studies of usability over time: what caused the users' problems as a function of experience.

While the data from our study do show changes over time, the trends are not as clear as those reported in [8]. Figure 3 shows the changes in relative distributions of frustration causes as a function of time. The y-axis is the percentage of episodes of a frustration type for a particular level of experience with the software, and the x-axis is the experience level ("0" is the interval of experience from none to less than three months, etc.). The lack of a clear overall trend may be due to the heterogeneous software used by the subjects. Some of the kinds of frustration causes did appear to have interesting trends, though. Figures 4 and 5 show the trends for "App behavior" and "Freeze/crash," respectively.

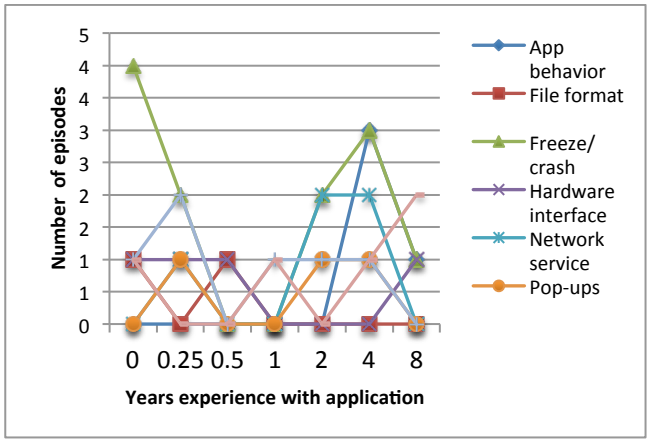

Figure 3. Relative frequencies of frustration causes as a function of experience with the software.

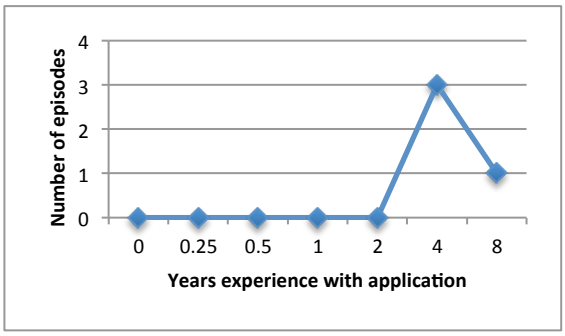

Figure 4. Absolute frequency of frustration episodes caused by "Freeze/crash" as a function of experience with the software.

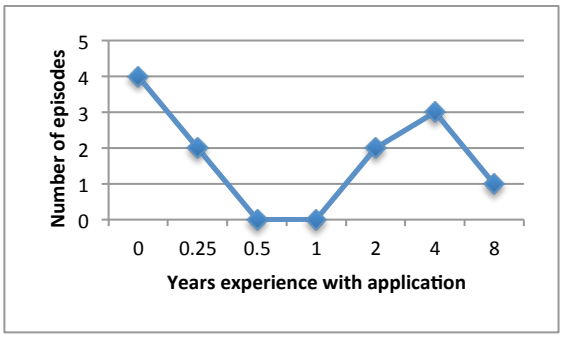

Figure 5. Absolute frequency of frustration episodes caused by "App behavior" as a function of experience with the software.

The results for "Freeze/crash" seem counterintuitive, in that the more-experienced subjects reported greater use of office software

and lower use of media software, in that office applications would seem less likely to freeze or crash than media applications. The results for "App behavior" may suggest that novice users do have a greater frequency of usability problems with an application's user interfaces, which trends down until the point when, as the users gain more than a year's experience, they begin tackling more advanced tasks, leading again to usability problems with the interface. In other words, some of the differences between less- and more-experienced users may be due to the more-experienced users trying to accomplish more with the application: attempts at tasks with greater complexity lead to greater frustration. The actual problems reported by the users tend to support this conjecture. For example, experienced users of office applications had problems generating a table of contents in Microsoft Word or creating a function in Microsoft Excel to select unique numbers in a column.

We also looked at the relative frequencies of causes of frustration episodes not as a time series but rather distinguishing between less-experienced users (zero to three months of experience with the software leading to the frustration episode) and more-experienced users (more than three months of experience). This analysis is shown Figure 6. Again, no clear relationship is evident, which was confirmed by statistical analysis (ChiSq > 0.57).

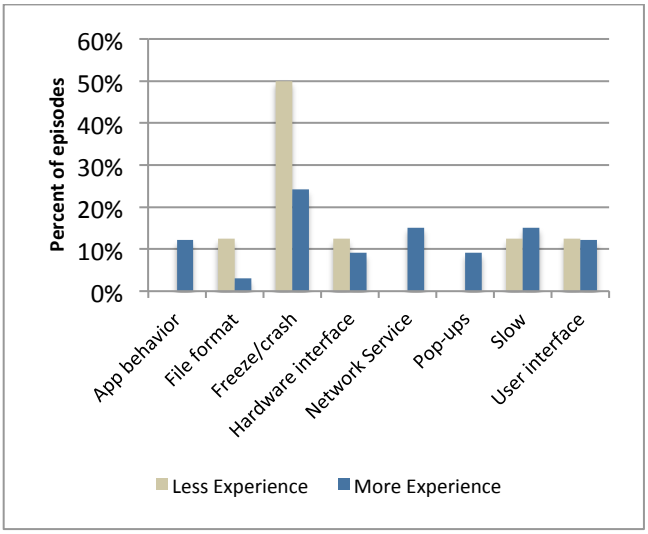

Figure 6. Relative frequencies of frustration causes as a function of less or more experience with the software.

4.3 Solutions Tried

For each episode, the users reported all of the methods with which they tried to solve (or abandon) the problem. The questionnaire used the twelve categories (plus "other") reported in [8], but for purposes of analysis we collapsed the responses into four categories:

- Using help from a manual, on-line help, or Web pages
- Asking someone, such as a colleague, a help desk, or a stranger via the Internet
- Using a workaround, which produces the appearance of a solution, or trial and error exploration of the interface
- Rebooting the computer or giving up.

Subjects could report more than one solution method for a frustration episode (e.g., used help system, asked someone I know, gave up). The distributions of the subjects' reported solution methods are show in Figure 7, distinguishing less-experienced users (zero to three months of experience with the software leading to the frustration episode) and more-experienced users (more than three months of experience). Analysis of these data suggests that there is not a clear overall relationship between level of experience and choice of solution method (ChiSq > 0.29). In only one episode did a subject report using a printed manual.

However, relative greater numbers of more-experienced users than less-experienced users actually solved their problem (ChiSq < 0.05), as could be surmised from Table 2, where less experience means zero to three months of experience with the software leading to the frustration episode and more experience means more than three months of experience.

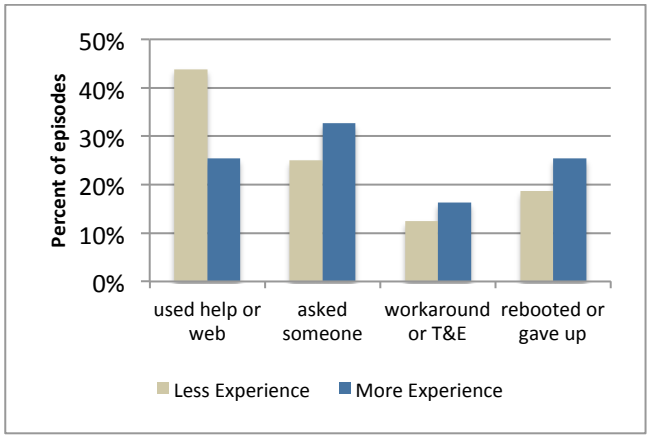

Figure 7. Distribution of solution methods as a function of experience.

Ironically enough, the experienced users achieved this success while using help significantly less than less-experienced users, as shown in Table 3, where "used help" means used help from a manual, on-line help, or Web pages, or asked someone, such as a colleague, a help desk, or a stranger via the Internet. This difference was significant (ChiSq < 0.001).

Table 2. Problem solution as a function of experience.

Solved problem	Inexperienced	Experienced
Yes	0	12
No	6	9

Table 3. Use of help as a function of experience.

	Less Experience	More Experience
Did not use help	1	14
Used help	11	6

These results are consistent with those reported in [8], which suggest that people using help are no more likely to achieve task success than people not using help, and that use of help tends to fall off as function of experience, even when people encounter novel problems.

5. DISCUSSION

In the course of our analysis, we tried distinguishing less-experienced and more-experienced subjects with different break points. That is, we tried a series of analyses where more-experienced meant more than three months, more than six months, and equal to or more than four years of experience with the software with which the user had a frustration episode. These analyses

suggest that the greatest changes in users' experiences come early. For example, we looked at the significance of the chi-square test for whether subjects used help. As indicated in Figure 8, Chi-square increases, and thus significance falls off, as the break-point between experience levels increases.

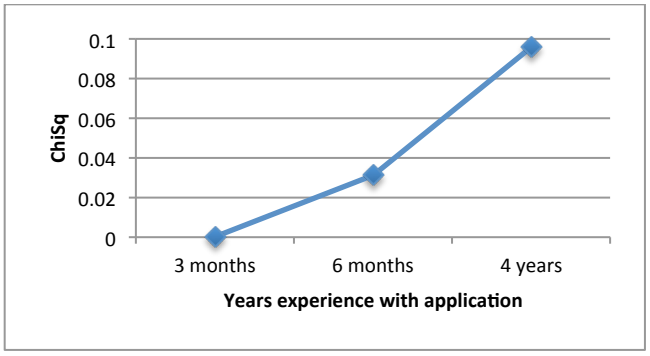

Figure 8. Chi-square of test of use of help as a function of break-point between experience levels.

We found a similar result in looking at the kinds of software reported by subjects in their frustration episodes as a function of experience. As indicated in Figure 9, again Chi-square increases, and thus significance falls off, as the break-point between experience levels increases.

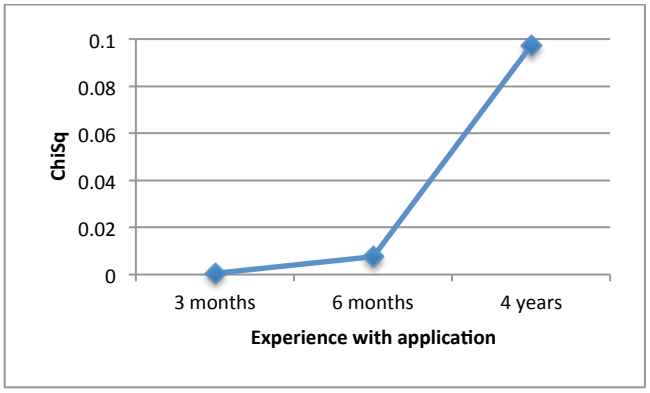

Figure 9. Chi-square of test of type of software used as a function of break-point between experience levels.

This analysis suggests that macro-longitudinal studies beyond perhaps a year may not be particularly useful for many kinds of applications, as most of the observable differences in use appear to occur within three to six months of initial use. Because we had only eight experience reports with application use less than or equal to three months, we were not able to explore whether an even lower break-point for experience would be useful.

6. CONCLUSION

This paper explored whether a non-longitudinal, cross-sectional approach could reliably detect useful differences in usability between novices and experts.

Our results suggest that it is possible to study long-term usability with a methodology that combines a cross-sectional approach with contemporaneous evaluation. This "snapshot" methodology avoids potential problems with retrospective evaluation, while, as we have argued, being no worse off than longitudinal methodologies with respect to variability among subjects. The snapshot methodology enables researchers to look at usability over much longer periods of time than would be practical with longitudinal methods.

Our analysis suggests that a cross-sectional methodology can distinguish between less experienced and more experienced users with respect to the kinds of applications that cause frustration, frequency of use of help, and whether the problem was solved. Our analysis also suggests that the method is poor at distinguishing causes of frustration and the overall distribution of types of solutions tried. This latter result is frustrating, so to speak, because understanding these differences is a principal goal of studies of usability over time, and the results of the longitudinal studies discussed in Section 2 suggest that these differences across time do, in fact occur. It may be that the heterogeneous collection of software used by our study's subjects contributed to this problem and that applying the snapshot approach to a single application would yield more useful results. It is also possible that clearer results might be produced by collecting more than 41 experience reports.

The data also suggest that three months of use of an application is the most useful point at which to distinguish less-experienced from more-experienced users.

6.1 Limitations

While the study did yield useful results with respect to kinds of applications that cause frustration, frequency of use of help, and whether the problem was solved, with only 41 experience reports, it seems clear that a greater number of reports would be highly useful.

More fundamentally, our results may reflect a self-selection bias among the subjects who participated in the study. This possible bias may be seen, for example, in the distribution of experience for the software used, which is weighted toward multi-year use, and in the subjects' overall experience with computers, which had a median value of 20 years. Thus our results should be interpreted as reflecting the usability experiences of relatively sophisticated users of computers.

Along similar lines, the extent of experience with applications may be a function of when the applications became available. Users may have some number of years experience with an application because it was launched that many years ago; other applications, more recently launched, will not have any users with longer experience. As a result, trends for causes of frustration may conflate newness of the application and lack of experience with the application.

The relatively high experience levels of our subjects also meant that we had fewer low-experience frustration reports than we had hoped to obtain. Consequently, our analysis could not reliably distinguish experience levels below three months.

7. ACKNOWLEDGMENTS

This work was supported in part by an endowment from AT&T. The authors thank the study's participants for their time. The authors also thank Nigel Ward and the conference's reviewers for their helpful comments, which markedly improved this paper.

8. REFERENCES

[1] Jacobsson, M. and Nylander, S. (2012) Always-On + Adoption – a method for longitudinal studies. CHI 2012 Workshop on Theories, Methods and Case Studies of Longitudinal HCI Research, May 5, 2012, Austin, TX.

[2] Jain, J., Rosenbaum, S., and Courage, C. (2010). Best practices in longitudinal research, *Proceedings of the 28th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI '10),* April 10-15, 2010, Atlanta, GA, 4791-4794.

[3] Karapanos, E., Hassenzahl, M., and Martens, J.-B. (2008). User experience over time: An initial framework. *Proceedings of the 26th International Conference on Human Factors in Computing Systems (CHI '08)*, Florence, Italy, 3561-3566.

[4] Karapanos, E., Martens, J.-B., and Hassenzahl, M. (2012). CHI 2012 Workshop on Theories, Methods and Case Studies of Longitudinal HCI Research, May 5, 2012, Austin, TX, http://longitudinalusability.wikispaces.com/CHI2012Workshop.

[5] Karapanos, E., Martens, J.-B., and Hassenzahl, M. (2012). On the retrospective assessment of users' experiences over time: Memory or actuality? In: CHI 2012 Workshop on Theories, Methods and Case Studies of Longitudinal HCI Research, May 5, 2012, Austin, TX.

[6] Khanlarian, C., 2010, A *Longitudinal Study of Web-Based Homework*, Ph.D. Dissertation, University of North Carolina at Greensboro.

[7] Mendoza, V., and Novick, D. (2005). Usability over time, *Proceedings of SIGDOC 2005*, Coventry, UK, September 21-23, 2005, 151-158.

[8] Novick, D., Andrade, O., Bean, N., and Elizalde, E. (2008). Help-based tutorials, *Proceedings of SIGDOC 2008*, Lisbon, Portugal, September 22-25, 2008, 1-8.

[9] Novick, D., Elizalde, E., and Bean, N. (2007). Toward a more accurate view of when and how people seek help with computer applications, *Proceedings of SIGDOC 2007*, El Paso, TX, October 22-24, 2007, 95-102.

[10] Prümper, J., Zapf, D., Brodbeck, F.C., and Frese, M. (1992). Some surprising differences between novice and expert errors in computerized office work, *Behaviour & Information Technology* 11(6), 319-328.

[11] Roto, V., and Kujala, S. (2012). Studying six months in two hours. In: CHI 2012 Workshop on Theories, Methods and Case Studies of Longitudinal HCI Research, May 5, 2012, Austin, TX.

[12] Szóstek, A., and Walo, K. (2012). Prospective Day Reconstruction method: a way to validate service design concepts. In: CHI 2012 Workshop on Theories, Methods and Case Studies of Longitudinal HCI Research, May 5, 2012, Austin, TX.

[13] van Lumig, C.J.L. (2009). *The influence of user expertise on the usability experience: Interfaces for different users at Vodafone call centers*. Master's thesis, Eindhoven University of Technology.

[14] Vaughan, M., and Courage, C. (2007). SIG: Capturing longitudinal usability: what really affects user performance over time?, *CHI '07 extended abstracts on Human factors in computing systems*, April 28-May 03, 2007, San Jose, CA, 2149-2152.

[15] Vaughan, M., Courage, C., Rosenbaum, S., Jain, J., Hammontree, M., Beale R., and Welsh, D. (2008). Longitudinal usability data collection: art versus science?, *CHI '08 extended abstracts on Human factors in computing systems*, April 05-10, 2008, Florence, Italy, 2261-2264.

[16] von Wilamowitz Moellendorff, M., Hassenzahl, M., and Platz, A. (2006). Dynamics of user experience: How the perceived quality of mobile phones changes over time. User Experience - Towards a unified view, Workshop, *4th Nordic Conference on Human-Computer Interaction*, October 14 - 18, 2006, Oslo, Norway, 74-78.