

High Costs and Small Benefits: A Field Study of How Users Experience Operating System Upgrades

Francesco Vitale¹ Joanna McGrenere² Aurélien Tabard³ Michel Beaudouin-Lafon¹ Wendy E. Mackay¹

¹LRI, Univ. Paris-Sud, CNRS,
Inria, Université Paris-Saclay
F-91400 Orsay, France

²University of British
Columbia
Vancouver, Canada

³Univ Lyon, Université Lyon 1,
CNRS UMR5205, LIRIS,
F-69622, France

francesco.vitale@inria.fr, joanna@cs.ubc.ca, aurelien.tabard@univ-lyon1.fr, {mbl, mackay}@lri.fr

ABSTRACT

Users must manage frequent software and operating system upgrades across multiple computing devices. While current research focuses primarily on the security aspect, we investigate the user's perspective of upgrading software. Our first study ($n=65$) found that users delay major upgrades by an average of 80 days. We then ran a field study ($n=14$), beginning with in-depth observations during an operating system upgrade, followed by a four-week diary study. Very few participants prepared for upgrades (e.g., backing up files), and over half had negative reactions to the upgrade process and other changes (e.g., bugs, lost settings, unwanted features). During the upgrade process, waiting times were too long, feedback was confusing or misleading, and few had clear mental models of what was happening. Users almost never mentioned security as a concern or reason for upgrading. By contrast, interviews ($n=3$) with technical staff responsible for one organization's upgrades focused only on security and licensing, not user interface changes. We conclude with recommendations for improving the user's upgrade experience.

ACM Classification Keywords

D.2.7 Software Engineering: Distribution, Maintenance, and Enhancement; H.5.2 Information Interfaces and Presentation: User interfaces—user-centered design; K.6.5 Management of Computing and Information Systems: Security and Protection

Author Keywords

Software upgrades; observational study; qualitative analysis

INTRODUCTION

The management of OS upgrades has changed dramatically over the last decades. Although technical administrators used to manage upgrades of shared computing platforms, with the advent of personal computing, it became a job that end users were expected to do on their own.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI 2017, May 06-11, 2017, Denver, CO, USA

© 2017 ACM. ISBN 978-1-4503-4655-9/17/05...\$15.00

DOI: <http://dx.doi.org/10.1145/3025453.3025509>

While minor updates increasingly happen in the background [1], users must still manage a large number of OS upgrades and updates that are distributed directly to their collection of computing devices, from laptops to tablets and smartphones. However, the human effort to manage upgrades and accommodate the changes that they cause (expected or unexpected, positive or negative) is not well understood.

Most of what we know about the upgrade experience comes from the popular press, where negative stories about major upgrades dominate. For example, in 2015 Microsoft released Windows 10 and promoted it aggressively [42], pushing users to install it within one year [15]. Some people ended up not having a choice [6], while others disabled important security updates just to avoid it [4]. In 2014 Apple introduced OS X 10.10 and with it a new version of the Spotlight search engine [35] that started tracking the user's location [5]. Users were not aware and a major controversy followed [36]. Even minor updates can bring unexpected changes that have negative impacts on users [24].

The research literature on upgrades focuses on technical and security aspects [7, 19] and rarely addresses the user experience. We are aware of only two papers, both by Vaniea and colleagues, that specifically address the user experience [39, 40]. In one study [40], they found that people report avoiding updates, but they do not quantify this phenomenon. They also found that users rely on past experiences to decide whether or not to upgrade. A bad experience can result in avoiding future upgrades, irrespective of how important they are for security reasons. In a second study [39], Vaniea et al. modelled the process of upgrading, and found that users evaluate potential costs and benefits in their decisions. However, they do not document these costs and benefits in detail. Additionally, both studies used retrospective methodologies. We build and expand directly upon their findings using complementary methods.

We first ran an online study of 65 OS X users to obtain quantitative data as to the extent to which they delay major and minor upgrades. In complement, we conducted an in-depth study of 14 users across several operating systems in which we observed the participants as they installed an OS upgrade on their own devices. Finally, we captured the costs and benefits experienced through a four-week diary study.

RELATED WORK

The two literature areas most relevant to software upgrades are: (1) how users make decisions on updates in different contexts and with varying degrees of freedom; and (2) how people remember past affective episodes to make future decisions.

User attitudes, beliefs, and decisions about upgrades

When deciding to upgrade, users consider multiple variables. Vaniea and Rashidi [39] found that the update process generally consists of six stages: 1. awareness → 2. deciding to update → 3. preparation → 4. installation → 5. troubleshooting → 6. post state. All stages are critical to the user experience and prone to breakdowns. Their study offers an unprecedented look at the process of updating and brings to light several issues that we also encountered. However, it captures a broad but shallow picture of the upgrade user experience, leaving unexplored many specific issues that arise during the installation process itself. Our work addresses this gap.

Update notifications influence decisions – The design and layout of notifications influence users’ assessment of the importance of an update. Good design can reduce negative emotions such as confusion and annoyance [10]. Design elements such as buttons with simple, clear choices improve trust. The content of notifications also has an impact. Generally, messages from online reviews by other users are more effective than the notification messages that contain a list of permissions required by the application [38]. Notifications also cause problems by interrupting users. The literature on the subject is vast¹ and the issue largely beyond the scope of this paper.

Reasons to avoid or install updates – Common reasons to avoid updates include changes in functions, privacy concerns [38], and uncertainty about their consequences [9, 40]. In some cases, users think that updates are useless [40]. On the other hand, habits, trust, and frequent use are frequently cited for updating software [39].

Update installation rates – Looking at updates of a mobile application over 102 days, Möller et al. [28] show that on average only a minority of users (17%) install an update on the day of its release. Installation rates increasingly drop over the following days and only 53.2% of users, on average, end up installing an update within a week. We expand on these findings by looking specifically at OS upgrade installations.

Forced upgrades – When users are forced to upgrade, often as employees of a large organisation, they do not seem to experience any benefits [21]. In such organisations, management and security staff are often in charge of installing upgrades; and business rather than user needs drive decisions [20]. The main reported consequence of an upgrade is that users must spend time learning the system again. By contrast, Fleischmann et al. [11] argue that users view updates that introduce new features positively. Their results, however, come from a very narrow controlled experiment based on an hypothetical scenario, making it difficult to generalise their findings.

Automatic updates – Automatic updates are more effective than manual updates in keeping users up to date, especially in

browsers. Chrome’s “silent update” has proven effective since its introduction [7]. Firefox also shows that automatic updates boost adoption rates [14]. From a technical standpoint, some concerns have been raised about the security of the shipping mechanism [8]. Vaniea and Rashidi [39] discuss the user perspective on automatic updates in more detail. Suffice it to say that automatic updates can confuse users and diminish their sense of control [41]. Additionally, non-expert users tend not to set automatic updates [17].

Note that we use the words *upgrade* and *update* interchangeably. When necessary, we qualify them as major or minor.

Memory and peak-end effects

Most of the literature on the upgrade user experience uses a retrospective methodology, which suffers from classic limitations related to human memory, such as peak-end effects. According to the peak-end rule [18], the moment of highest intensity and the ending of an affective episode influence how people feel about it when making retrospective judgements. The duration of the episode, by contrast, is neglected [13]. However, duration may influence how people remember episodes when there is a mismatch between expectations and reality [18]. Vaniea et al. [40] already found that users rely on memories of previous update experiences when making future update decisions and that the duration of upgrades seems to have some impact. Therefore, we wanted to better investigate the link between actual experience, memory, and future decisions. In particular, we sought to understand what is the most memorable aspect of the upgrade experience.

STUDY 1: DELAY OF UPGRADES

To quantify the extent to which people delay upgrades, we conducted an online study of OS X users. Previous studies have shown that users avoid [40] or delay [17] updates. However, they do not distinguish between major and minor updates, nor OS and application updates, and they do not quantify the delay. We chose to study OS X because the data for most updates can be transmitted through a single system file that contains reliable installation dates.

Participants – We recruited participants through online forums and newsletters, social networks, and personal networks. We received 65 responses from Canada, Europe and the United States. The participants’ age ranged from 18 to “65 and above”, with 49% between 25 and 34. The sample was overeducated, with 93% of participants having at least a Bachelor’s degree. Occupations included graduate student (26), teacher (17), developer (3), designer (3), consultant (3), engineer (2), researcher (3), system administrator, press officer, head of communication, vice principal, copywriter, coordinator, and 2 unknown. We did not compensate participants.

Data collection – Participants uploaded a system file from their computer: `InstallHistory.plist` (located at `/Library/Receipts`). It contains a log of updates installed through the App Store, including OS X major and minor upgrades, and some third-party ones. We also asked five closed-form questions, including whether the computer they were using was their main one, whether they decided when

¹<http://interruptions.net/literature.htm>

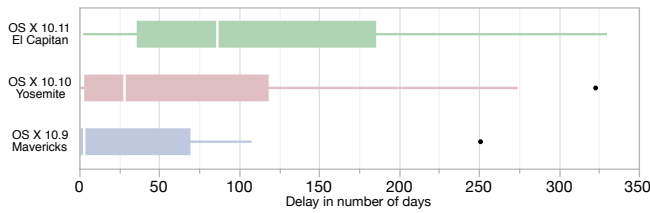


Figure 1. The average installation delay for major OS X releases. (n=65)

to install updates on their own, demographic questions, and a self-assessment of their technical skills (5-point Likert scale).

Data analysis – We extracted the installation dates for 18 major (e.g. 10.9), or minor (e.g. 10.9.4), OS X releases between 2013 and 2016² from the `InstallHistory.plist` file. We then calculated the difference in number of days between the installation date and the official release date³ [16].

Results – Overall, we collected data about 394 upgrades (71 major, 323 minor). 92% of participants reported information about their own computer, with 88% having control over their own updates. We found that participants delayed major upgrades by 80 days on average (median: 46, min: 0, max: 330, SD: 88.8) and minor updates by 16 days on average (median: 8, min: 0, max: 228, SD: 25). We also found that the delay for major upgrades increased over time (Figure 1). We suspect the delay for 10.9 is short (average: 43 days, median: 3, min: 0, max: 251, SD: 74.7, installations: 12/394) as it was the first major free OS X upgrade.

Previous studies suggest that these results should be seen as conservative. 51% of our participants reported having very high technical skills (self-rating of 5/5). This expertise can correlate with faster installation rates [17]. Based on a study of over 8.4M computers, Nappa et al. [30] found that technical users were 50% faster than ordinary users in installing application patches. The median installation time for ordinary users was 45 days, consistent with our results. Overall, these findings suggest that non-experts might delay updates even more.

STUDY 2: THE UPGRADE EXPERIENCE

In our second and main study, we sought to understand how end users experience software upgrades. Our goal was to observe the upgrade process as it unfolds and characterise the main issues users face. We also wanted to document the effects of an upgrade, immediately after the upgrade and in the longer term. Finally, we wanted to see how people remember the experience after some time has passed, to understand how retrospective judgements of the event might play a role in future decisions to upgrade. Our results could shed light on why people avoid or delay upgrades, as found in Study 1.

² 10.9 through 10.9.5, 10.10 through 10.10.5, 10.11 through 10.11.5

³ In cases where users installed beta versions of the software before the official release (7 out of 65), we used only the installation date of the official release. In case the download and installation steps took place on different days, which can happen with major upgrades, we used the download date.

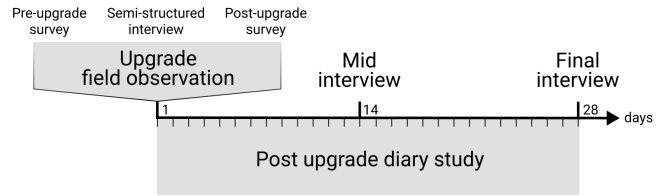


Figure 2. Study procedure: Initial field observation of an upgrade; a 4-week diary study; mid-study and final follow-up interviews.

Participants

We recruited 14 participants, aged 23 to 43 (mean: 28.5, SD: 7.1) through advertisements on university campuses and personal networks in France and Italy. The participants included 8 graduate students, 2 engineers, a video editor, a teacher and a journalist. 12 participants had at least a Bachelor degree, with a background in computer science (7), engineering (2), communication, psychology or educational sciences. We did not compensate participants.

Procedure

Figure 2 gives an overview of our procedure. The study consisted of an *in-situ* observation of an upgrade followed by a four-week diary study (duration determined by piloting). Participants upgraded an operating system of their choice on their own device(s). We asked them to prepare as they normally would (e.g. doing a backup), but did not give any examples. Participants chose where the observation took place: either their home or office, or in some cases our lab or the library. Before the upgrade, they completed a short survey about their expectations. Then, we observed them performing the upgrade and conducted a semi-structured interview. We did not help in cases of difficulty and let them make decisions. If the upgrade was still underway after asking all of our questions, participants were free to perform other activities. At the end of the process, they completed a second, short survey. Two investigators were responsible for all observations. Interviews were conducted in English, French, or Italian.

After the observation, participants started creating daily diary entries. We encouraged them to take screenshots of noticeable changes, to serve as a reminder when creating the entries. Each participant received a daily email reminder at their preferred time. We used a different schedule for two participants, who used the device they upgraded only on specific days of the week. The email contained two links: one to report any changes, the other to indicate that they had nothing to report. (We considered both responses to be an entry.) Half-way through the diary period, we conducted a short, semi-structured interview (four minutes on average) to check in on each participant. At the end of the four-week period, we conducted a final, semi-structured interview (seven minutes on average). In two cases, we skipped the mid-study interview because of scheduling difficulties.

Data collection

Observations – In the pre-upgrade survey, we asked participants (1) how much time they thought the upgrade process

would take (closed question), and (2) how well they expected it to go (7-point Likert scale). We measured how much time the upgrade took, noting the different steps required and relevant prompts or error messages. The semi-structured interview questions focused on the upgrade process, users' expectations, emotions and attitudes, their background and technical expertise, and their previous experiences with updates on other devices. When prompting users to share what they were thinking, we focused on their emotional reactions and interpretation of what was happening in the process. We probed some issues more than others, because we wanted to make sure all participants addressed them, for example, what they thought of the feedback during the installation. We audio-recorded each observation, took extensive hand-written notes, and later transcribed the audio. In the post-upgrade survey, we asked participants closed questions about (1) how much time they thought the process took, and (2) how they thought the process went, as well as 7-point Likert scale questions on (3) the degree of control they felt in the process, and (4) how comfortable they were with having this level of control.

Diary study – In the diary, we asked participants to (1) list any changes in their systems that they believed were related to the recent upgrade, (2) describe any impact of those changes, (3) rate their satisfaction with the decision to upgrade (7-point Likert scale), (4) upload any screenshots, and (5) give additional comments. The mid-study semi-structured interview asked participants for any additional screenshots, and probed for further details about their entries. The final semi-structured interview asked them to further clarify the diary entries, identify the most useful new feature discovered, describe the biggest frustration over the previous month, and assess whether or not the upgrade was worth it. We also asked participants to reflect on the upgrade process to assess peak-end effects in their retrospective evaluations: We asked them to recall the most vivid memory from the observed installation process, as well as the main steps they went through, how they felt during and at the end of the process, how they felt after four weeks, and how they would approach upgrades in the future. We audio recorded and later transcribed the interviews.

Data analysis

We followed a thematic analysis of data collected during the observations and the diary entries. We used open coding to come up with a first set of categories and themes. Then, we grouped macro and sub-categories to create a hierarchy.

Observations – We broke down the observations by chunking participants' statements into short expressions centered around a topic. We used breaks and pauses in their answers to create self-contained expressions. Each expression is categorized by sequence: before, during or after the upgrade. We then created categories of themes (e.g., control, feedback, assessment), and categorised emotions and expectations as either positive, neutral, or negative. Two members of the research team initially coded all the data from the observations separately, then compared the results and discussed the categories. We grouped high-level themes emerging from the data after multiple iterations and discussed them regularly with the rest of the research team.

Platform	Major upgrade	Minor upgrade
OS X	P2, P4, P6, P8, P11, P14	P3
Windows	P5, P7, P13	
Linux	P1, P12	
Android		P3, P6, P9
iOS		P10, P14

Table 1. OS major and minor upgrades performed by the participants. Expertise indications: expert users, *above average users*, *average users*.

Diary study – For each diary entry, we created categories for the type of *change* (e.g., bugs, lost functions, lost settings) and the type of *impact* (e.g., emotional, time saved or lost, change in habits). We gave a valence rating to each change (positive, neutral, or negative). We later framed the changes and the impacts in terms of costs and benefits, based on their valence rating. Emotional impacts had additional clusters of subcategories based on whether they were negative or positive (e.g., negative: frustration, annoyance, irritation; positive: like, appreciation). We referred to Plutchik's model of emotions [33] and Posner and Russel's circumplex model of affect [34] to group emotions into clusters. One member of the research team created the first code book after four iterations on the diary entries. Then, two members of the team discussed a subset of the entries together, focusing on the subset of participants with rich entries and those who were the most difficult to categorise. We used the participants' words as much as possible when creating the categories. We did not infer emotional impacts unless participants were explicit in their entries, for example, "*I am very frustrated with the audio settings*" (P5). We were careful to not over-interpret the valence of changes and the emotions participants reported. We present counts to show the relative frequency, but do not make more general claims based on these numbers.

RESULTS FROM STUDY 2

Participants performed minor and major upgrades on different operating systems (Table 1). The majority (11) performed only one upgrade. P6 performed a major upgrade on OS X and a minor Android-based one, but only reported changes about Android. P14 performed a major upgrade on OS X and a minor on iOS, but only reported changes about OS X. P1 performed a clean install of a new OS. P12 was forced to do a clean install (that we did not observe) after the upgrade failed during the observation. From our data analysis, we ended up with a total of 435 self-contained expressions by participants concerning the upgrade process.

We organise our results around four different periods of the upgrade experience: (1) just before the upgrade, (2) the installation process itself, (3) the four weeks that follow, and (4) memory of the installation.

Before the upgrade: attitudes, expectations, and planning

The data in this section were captured as part of the interview during the installation, when participants reported their attitudes, expectations, and planning going into the upgrade.

Security is not a large concern for users

Given the importance of security as a reason to install software upgrades, we investigated how users think about it. We did not

ask explicitly about their concerns or practices around security. Instead, we checked if they mentioned it spontaneously. To our surprise, only three participants (two of them experts) indirectly mentioned security as a concern, by pointing to expectations for improved security in their devices, and none cited security as the reason for the upgrade. For example, P3 said: “*I do care about security to some extent. I know it’s important. Maybe I should be more concerned.*” None of the other participants mentioned security as a concern or a reason for upgrading.

Users worry about upcoming changes

Most participants (10) worried about upcoming changes in their systems. Four participants hoped there would be no changes in the functions they already knew how to use, suggesting that they did not want to invest time in learning how to use the system again:

I hope the functions will stay exactly the same [...] You have your habits, I have my habits, so I don’t want a big mess [...] I am a little reluctant to have big changes, because they’re uncomfortable. I need to learn again how to use the system that I used in a certain way. (P14)

Negative expectations focused on problems with hardware and drivers; settings; lost data, functions, or applications; the time it would take to adapt to the changes; new, unwanted applications. Positive expectations focused on new, useful functions; new options for surface-level customisation of the system; fixed bugs; a better working OS; improvements in usability, performance; the system staying the same.

Users do not worry about the upgrade installation

On the other hand, the majority (11) of participants had positive expectations about how the installation process would go: only three expected it to be challenging (Likert scale ratings below the midpoint). Consistent with previous findings [40], one of the participants who had negative expectations mentioned a past, negative experience.

Users do not prepare for upgrades

The relative confidence that participants displayed prior to the upgrade process might explain why they did not prepare for it. Only four participants did some lightweight preparation before the upgrade (e.g., checking the storage of the device, checking and saving documents or bookmarks), and only one did a full backup. Others said they did not do anything, for example, “*I never really prepare for any updates*” (P11), or were taken by surprise by the question:

I am not prepared. What do you mean, how did I prepare? Oh, you mean... a backup? That’s a good point. I have not done a backup. Can I have a moment? (P8)

Vania and Rashidi [39] also found that a minority of users prepare before updates, but suggested that users may not recall this activity when telling stories of past experiences. Our results question this interpretation, since the majority of our participants did not even consider preparing for an upgrade. Performing backups may not be a common practice, even among expert users.

Another missed activity was reading the release notes for the upgrade: only three participants did so and had an idea of what the upgrade would do. One participant later regretted not having read some information on the new system. In general, release notes differ across platforms and they are not always part of the actual upgrade process: sometimes users must go to a website or a different application to read them.

Users’ fuzzy mental models lead to wrong expectations

Whether or not they prepared for it, most participants said they did not have a clear mental model of the upgrade process: “*I don’t know how it works*” (P4). Mental models of interactive systems are by definition incomplete and change over time [32]. However, not having one led participants to incorrect assumptions and expectations. For example, three participants thought that they would be able to downgrade the Mac’s operating system: “*I think it’s doing a backup of the last OS. So you can downgrade if there’s a problem.*” (P2). While this is possible on Windows, downgrades on the Mac require reinstalling the OS. P8 also expected other applications would be updated during the process.

The installation process: duration, feedback, and control

The observations lasted two hours on average (min: 25 minutes, max: 6 hours). The distribution of time expectations for the upgrade relative to actual duration was trimodal: four participants overestimated how long the upgrade would take, five underestimated, and five made a correct prediction. However, there was a bigger margin of error when underestimating than overestimating. On average, participants overestimated by 10-15 minutes, and underestimated by two hours or more.

Long waiting times frustrate users

During the upgrade, most participants (11) were particularly annoyed by both the long duration of the overall process and the waiting times with nothing to do. They consistently complained that the installation process took too long and was slower than expected, even during minor updates. The fact that they did not have much to do, besides “babysit” the computer, led to boredom: “*This is boring. Last time we made pizza and ate it while I upgraded*” (P5), frustration: “*You have to wait and you can’t do anything. Maybe they could offer you a movie or trailers without watching this.*” (P4), or anger:

*What the f**k! it’s been four hours for an upgrade! [...] I don’t get it. Why do you need to wait hours for the configuration? Why can’t it ask for everything in the beginning? I need to babysit it step by step! [...] I am bothered! [...] There’s no end in sight. (P14)*

Vania and Rashidi [39] also found that upgrade duration is a common complaint. Our findings are consistent with theirs, but our in-situ data provides more detail, and further shows a link between duration, waiting times, and user frustration. The long duration is the main issue users face when performing OS upgrades and below we report how it affects their memory of the process. In addition to long durations, participants are also frustrated by the the continuous attention that upgrades require.

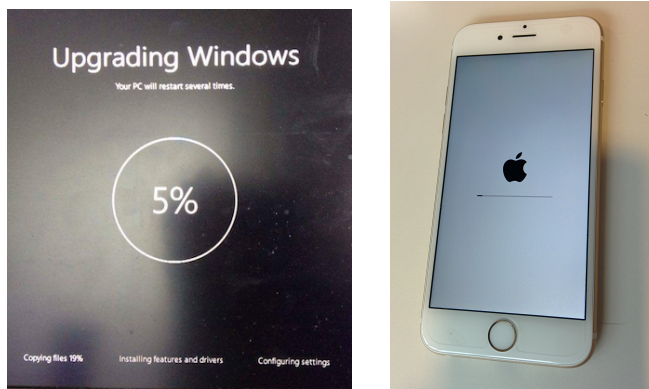


Figure 3. Progress feedback on Windows (left) and iOS (right).

Poor progress feedback compounds users' frustration

Participants were often dissatisfied with the progress feedback during the installation process. More specifically, the accuracy of the estimated time to completion caused frustration and distrust in the system. In particular, progress bars on Apple devices were unreliable and inconsistent across different steps: sometimes they got stuck, gave wrong estimates, for example, jumping from 27 to 57 minutes, or did not provide any timing information at all (Figure 3). As P8 put it: *"This progression bar is bulls**t. It goes from 30 to 7 minutes, it gets stuck... What the h**l is this representing?"* Some participants distrust the progress bars: *"They lie to you. They want you to stay. [...] They always say something that is not true."* (P2). Two participants were positive about having the progress bars, but still regarded them as ineffective.

The feedback was better on Windows, showing an overall percentage of the installation process and separate percentages for the sub-steps (Figure 3). However, users still found it slightly confusing: *"[It's] not very obvious, but it tells me what it's doing."* (P5). *"I just realised 5% is overall and 17% only files. Not clear."* (P7). Additional messages at the end of the process explained that the upgrade was almost over. They provided reassurance and built anticipation: *"It's nice that they have this message, it's reassuring. [...] I'm not bored."* (P5).

Lack of control causes anger in some users

Another issue was the lack of control over which features were changed, removed, or added to the system. During the upgrade to Windows 10, participants faced a confusing screen where they had to choose so-called "Express Settings" or find a less prominent link to customise them. This is a "dark pattern", an interface designed to trick and force users into doing something they do not want [3]. Two participants clicked "Customise settings" and were shocked and angry to discover what they would have agreed to:

They want me to send my contacts?! Advertising ID? Location... to trusted partners?! This makes me really hate Microsoft. (P5)

Apple also insisted on promoting and pushing specific services or applications, for example, iCloud and Mail, even when participants did not want them: *"Because I use the OS I have*

to use Mail? [...] It's like a commercial, they asked me twice. [...] I'm not stupid." (P4).

The majority of participants (9) reported feeling that they had little or no control over the upgrade process, although two were neutral, and three felt that they had some control. Opinions about lack of control were mixed, ranging from clear discomfort to having no problems with it. For example: *"They gave me a bundle and I didn't make any decisions, but of course it's that way. Isn't that supposed to be? [...] Those people might know better. They're the experts."* (P6).

Despite the overall difficulty of the installation process, the majority of the participants (9) had a positive perception and found it smooth (Likert scale ratings above the midpoint). We reflect on how our sample might explain this result in the Discussion. Two participants had a neutral opinion, and three thought it was challenging (Likert scale ratings below the midpoint). Interestingly, there were four flips: two participants went from positive expectations to a negative assessment, and two went from negative to positive.

Post-upgrade chores add to frustration

Participants had to spend additional time at the end of the process to perform post-upgrade chores: installing additional updates and programs (8 participants), checking old applications and documents (4), cleaning up the system from unwanted applications (3), customising the system after losing settings (3). This was a clear source of frustration. Participants wanted to move on: *"What's this stuff? Not now... I just want to use my computer."* (P14), or just go on with their life: *"I was supposed to meet a friend. I have a social life!"* (P8).

These post-upgrade chores overshadowed the potential benefits of new functions or applications. Only four participants discovered new functions or applications immediately after the upgrade. Three explored the system spontaneously, one read the official release notes by accident while clicking on menu items. New functions and applications caused mixed reactions: confusion, disinterest, but also appreciation. Cleanup activities were mostly related to some of the new applications, especially on Windows: *"Candy Crush? Uninstall, thank you!"* (P7).

After the upgrade: high costs and small benefits

The diary completion rate was 81%. We collected 293 entries, 20.9 per participant on average (min: 7, max: 27, SD: 6.3). 91 entries reported one or more changes, 202 had nothing to report. We excluded from our coding seven examples that were user actions, not changes, e.g., *"I made a list of software I need."* (P12). We ended up with a total of 125 changes, 8.9 per participant on average (min: 0, max: 17, SD: 5.2).

Users notice most changes in the first week

Participants noticed the majority of changes (63%) in the first week and 81% by the end of the second week, though five participants also noticed some changes in the final week. After the first two weeks, some participants were unsure if the changes they noticed were related to the upgrade or not. Some mentioned that they got used to any changes that might have occurred: *"It works exactly as before. Either I've gotten used to the changes, or it's not like the changes were very*

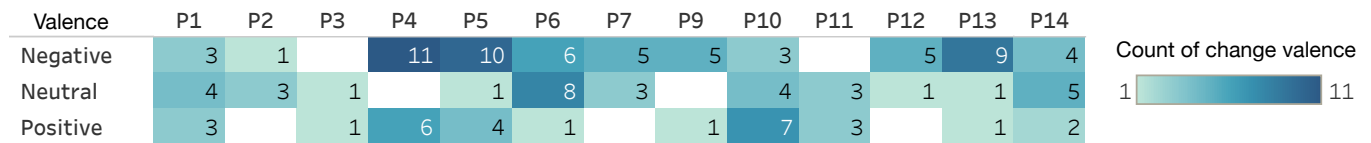


Figure 4. Number of changes (n = 125) with valence rating reported by participants after the upgrade (the higher the count, the darker the color).

Costs		Benefits	
Negative change	Count	Positive change	Count
lost settings	12	new features	13
bugs and crashes	10	better interface	6
worse performance	9	better performance	4
unwanted features	6	better features	2
worse interface	6	better behaviour	2
worse behaviour	5	fixed bugs	1
lost features	4	no changes	1
hardware issues	4	Total	29
notifications	3		
updates	1		
lost data	1		
changed features	1		
Total	62		
Negative impact	Count	Positive impact	Count
lost time	19	saved time	4
change in habits	17		
physical pain	1		
lost data	1		

Table 2. Changes (n=125) and impacts (n=42) reported by participants in the diary study. (Neutral changes and impacts not listed.)

important.” (P14, day 23). Only P8 did not notice any changes in the month after the upgrade: “To be honest, I don’t see what has changed. At all. Neither positive or negative.” (P8). This was surprising, but P8 was one of the two participants who only used the upgraded device a few days per week.

Costs outnumber benefits

Over the course of four weeks, participants experienced more than double the number of negative changes (50%, 62/125) relative to positive ones (23%, 29/125); with neutral changes (27%, 34/125) in between. In general, most participants oscillated between noticing negative, neutral, or positive changes throughout the four weeks after the upgrade (Figure 4). Overall, the costs they encountered after the upgrade outnumbered the benefits (Table 2). Users experienced more negative than positive emotions, both during and after the upgrade (Table 3). In our analysis, negative emotions were more varied and stronger: e.g., anger, fury and hate (16) vs. love (1). There were instances both during the upgrade and in the diary where participants (5) expressed very strong negative emotions in reaction to what was happening.

Changes to settings stood out strongly in the analysis, because all were negative. Participants often felt annoyed, even “furious” because of these changes:

All my documents, videos, pdfs, everything is now formatted to open by default with Windows apps, and not with my old apps. [...] I don’t know where to change them.

Negative emotion	Count	Positive emotion	Count
confusion, disorientation	81	like, appreciation	45
frustration, annoyance	73	anticipation *	5
worry, fear, anxiety	27	convenience §	5
impatience *	27	satisfaction *	4
dislike, disappointment	21	surprise *	3
anger, fury, rage, hate	16	trust	3
distrust *	12	pleasure §	2
disinterest *	8	love §	1
panic *	4	calmness §	1
stress *	4	happiness §	1
reluctance *	3	relief *	1
boredom *	3	confidence *	1
regret *	2	reassurance *	1
guilt *	1		
demotivation *	1		
despondency §	1		
embarrassment §	1		
fatigue §	1		

Table 3. Emotions experienced by participants during and after the upgrade. (*) denotes emotions only present during the observations, while (§) denotes emotions present only during the four-week diary phase.

[...] I think they did it on purpose to lock people in, and I am furious. This is malicious design. (P5)

Surprisingly, only one participant lost some data and one implied feeling physical pain because of a visual change in the interface: “Luminosity too strong, hurts my eyes.” (P4). Notably, not a single participant reported any changes or mentioned anything related to security.

When noticing positive changes after the upgrade, participants often expressed appreciation for better performance or an improved interface: “I noticed the way folders in a list look changed, now there’s the size near the icon. I like that it has the file size near the icon seems like a better place for it.” (P5). They also found some of the new functions to be “cool”, “nice” or “convenient”. Even minor changes could elicit very positive reactions: “I found new emojis. Haha. Those emojis are really cute and I love them.” (P10).

When asked, most participants (9) were unable to come up with an example of the most useful new feature they discovered after the upgrade: “To be honest, I don’t remember.” (P4). The rest mentioned some new functions, but most did not seem particularly convinced: “The finger sliding... the scrolling?” (P7), “Maybe the share... when I find something, a link... I want to share...” (P10). By comparison, most participants (12) expressed their biggest frustration with conviction. Examples include bugs, poor performance, additional updates and installations, and new features. Two participants also mentioned the upgrade process itself.

Remembering the upgrade process

We interviewed participants one month after the installation to see what they remembered; whether it was worth it in the end; and how they might change their future approach to upgrades.

Users remember long duration

When asked to pick the first memory that came to mind, most participants (11) mentioned the duration of the process: “*The time it took?*” (P3), “*It was so long.*” (P7). P13 mentioned talking with other people about the upgrade and advising them to be prepared for a demotivating and long installation process. Three participants also remembered that there was a mismatch between their expectations and the actual time it took. Only two participants mentioned a peak moment, and three each mentioned the beginning or the ending of the process. When asked specifically about the ending, most participants (8) had incomplete recollections. For example, P14 was particularly negative at the end of the actual upgrade, expressing frustration with the slowness, need for additional installations, and losing his background wallpaper. Four weeks later, he had forgotten this and largely discounted everything that he had described before: “[At the end] *Good. I don’t know. I didn’t notice any difference. It worked, so it was good.*” (P14).

Upgrading was worth it

Just over half of the participants (8) thought that the upgrade was worth it, although five thought it was not, and one was neutral. Some of those who thought it was “worth it”, did so even after describing an overall negative experience during or after the process, with negative changes outnumbering positive ones. For example, P12, whose upgrade failed during the installation, said: “*Yeah, sure. It was worth it [...] [But] I don’t think it was a positive experience. I don’t think it has in any significant way improved my system.*”

Users will keep avoiding upgrades

When asked about approaching future upgrades, four participants said they would follow the same approach as before, nine predicted they would change, and one wished to not be prompted about upgrading for a couple of years at least. Of the nine who predicted a change, five said they would delay, completely avoid the installation or arrange it in a way to skip the long waiting times, others would be more careful or worried: “*Now when I see an update on my phone, I am actually way more careful. I haven’t updated my phone since.*” (P6). Only one participant predicted a positive change in approaching upgrades: “*This experience will make me less worried.*” (P5).

AN IT MANAGEMENT PERSPECTIVE

Our results suggest that users do not view security as a major concern; and certainly not as a compelling reason to upgrade. Yet most upgrade literature focuses on security. This led us to consider the perspective of IT administrators: Do their views align more closely with software manufacturers or with users?

To that end, we interviewed three IT managers at a large multi-site research institute in Europe, with 2800 scientists and 850 administrative and engineering staff. We spoke to the chief of Information Technology, who is responsible for the organisation’s computer infrastructure. We also spoke to the head of security, responsible for protecting the organisation’s

network from external threats, and the head of licensing, who tracks the organisation’s computers and software versions, and is also a member of the *HelpDesk* management team.

We conducted a semi-structured interview about how they manage the organisation’s upgrades. We learned that they control the software on the engineering and administrative staff’s computers, but not that of researchers (although they would like to). They are responsible for 2100 machines: as of July 2016, they had identified 1203 with “critical vulnerabilities” and only 735 without any detected vulnerabilities. They explicitly do *not* consider usability issues, and were somewhat surprised that we asked.

We probed further, and asked whether user interface changes resulting from an upgrade might disrupt a user’s workflow. They replied by again stressing the importance of installing updates as soon as possible. The security chief said: “*End computers always need to be at the last version. There are always problems with security. You cannot be late in applying security updates!*”

We also interviewed a member of the administrative staff, who receives upgrades suddenly, usually without warning. If the upgrade involves a major change in functionality, she attends a half- or full-day training session to learn the new features. If she runs into other usability problems, she either adapts to them, or else contacts the *HelpDesk* to solve the problem.

While preliminary, these findings suggest a significant disconnect between the users’ and the IT managers’ perspectives. Users clearly worry that upgrades will cause usability problems, whereas these IT managers focus on security first, followed by breakdowns and licensing issues. Usability issues are never considered, except after-the-fact, when they generate questions directed to the *HelpDesk*.

DISCUSSION

We first reflect on our methodology, then expand on some of our findings. We conclude with design recommendations.

Reflections on our methodology

Recruiting difficulties – The tendency for people to delay or avoid upgrades seemed to influence the number of participants we managed to recruit for the observational study. The recruitment lasted three months. Potential participants declined to participate citing their lack of interest in the new version of their operating system, lack of time, and worry that the upgrade would compromise their work before a deadline. We also realized that many users prefer to upgrade at home, on evenings or weekends, since they can keep an eye on the device while doing other things. Potential participants may be reluctant to be observed during this suite of somewhat awkward activities. Committing to the follow-up four-week daily diary clearly did not help, even when we advertised it as being very lightweight.

Limitations in our sample – One consequence of these recruiting difficulties was the composition of our sample: most participants had above-average education and technical skills. However, previous studies suggest that technical expertise might not have an impact on the upgrade experience, since

expert and non-expert users describe similar experiences [39]. Expertise only seems to influence installation rates: experts install updates faster and more frequently than non-experts [17]. In particular, technical users are 50% faster than ordinary users when installing application patches [30]. The participants' initial expectations about the upgrade process were surprisingly positive, although we do not know whether or not this is an artifact of selecting people who are willing to participate in such a study. We agree with Fagan et al. [10] that we need additional research with a more general population.

Limitations of the study – As with all observational studies, participants may behave differently without an observer present. We designed our methodology to ask questions during the “dead” moments of the upgrade process, to avoid prolonging the study. However, this may have affected the participants' experience. Our presence may have also caused them to pay more attention to the process than they otherwise would have. Without an observer, some participants might have left their computer unattended or perceived the experience differently.

Reflections on our findings

Duration and progress should be better managed

Our results support previous studies that highlight users' frustration with long installation times. Software companies must do a better job of informing users of the duration. Apple avoids predicting OS X and iOS update timing: “*The time it takes to download the update varies according to the size of the update and your Internet speed*” [2]. Google follows a similar approach for Android; only Microsoft offers an explicit estimate:

The time that is required to upgrade to Windows 10 depends on factors such as the age of your device and how it's configured. Most devices will take about an hour to upgrade as soon as the download is completed. [27].

Providing accurate times is clearly a challenge given the variables at play, but surely reasonable estimates must be possible. Users appreciate progress bars, that are the defacto standard [29], but these should be supplemented with estimates.

Duration may hamper users from experiencing the benefits

Participants were exhausted by the extensive waiting times and longer-than-expected durations, especially when followed by post-upgrade chores. This left little time for exploring possible new benefits from the upgrade. Software designers should consider taking advantage of the dead time during installation, perhaps with successive screenshots that describe the changes, as well as their rationale and potential benefits. Better yet would be to let users learn and interact with new features, both to engage them, but also to distract them from the slow installation process.

Interfaces should support learning

Related to learning curves, we did not measure the impact of learning as a consequence of upgrading, nor did we focus on learning strategies. However, these participants clearly worried about the learning required to adapt to the upgrade's changes. Previous studies show how learning represents a cost for users [37] and might defer them from adopting a new or

changed system. Franzke and Rieman [12] demonstrate that early versions of a software application offer a better environment for learning because they usually include relatively few features. Subsequent upgrades should build on those features to reduce learning costs. Unfortunately, new software versions often introduce many new functions at once, contributing to “software bloat” [26]. Users can theoretically fight back complexity by customising and personalising their systems, but most do not [22] since this just takes more time. Another approach is to use multiple interfaces, each available at the same time, with a simple toggle for going back and forth, to ease the transition between versions [25]. Future work should address these opportunities and better measure how learning strategies differ across versions, contexts, and time.

Support reluctant users who worry about changes

Two thirds (10) of the users worried about upcoming changes, with four specifically wanting no changes at all. This begs the question of how we can encourage worried or reluctant users to embrace change. Preparing them to expect changes may be part of the solution (possibly, as mentioned above, by using upgrade time to describe changes). Users should not be expected to find and read release notes or check forums for discussions about changes. Alternatively, it could be interesting to allow users to select which changes to receive. This would require software to be compiled on the fly, based on individual needs, with corresponding implications for maintenance and software infrastructure adaptations.

Duration and peak-end effects

The theory of peak-end effects would predict that the user's retrospective assessment of the installation will be affected by a peak moment in the middle or the quality at the end of the upgrade process, but not the duration of the experience (what researchers call *duration neglect* [13]). However, this is not what we found: the interplay with peak-end effects in the upgrade experience may be more complex than in other situations. Three participants (P5, P7, and P12) experienced clear peak moments with strong negative emotions, and yet P12 was the only one who recalled the moment when asked “What is the most vivid memory you have?”

Similarly, other participants experienced clear end effects (P4, P8, P14) and yet none of those were recalled either. Instead, the great majority of our participants (11) answered this question by reporting about duration – remembering that the upgrade took a long time and in some cases that there was a mismatch with what they expected. What's more, some users said they would approach future upgrades in a way to avoid long waiting times. In this sense duration seems to affect not only the memory of past experiences, but also future decisions.

This may be because users expected shorter durations or could not predict when the process would end. Also, users experience changes from the upgrade over subsequent days, prolonging the overall experience. While preliminary, these results suggest that duration neglect does not necessarily apply to upgrades. However, significant additional research is needed to better understand this phenomenon. Some of the design recommendations below address this finding.

Users deem upgrades “worth it” despite negative experiences

The majority of participants (11 out of 14) reported negative reactions *during* the upgrade process, including a third (5) who expressed extremely negative emotions at specific points. The four weeks following the upgrade were somewhat more balanced, but still at best a mix of positive, neutral, and negative (6 participants had a majority of negative changes). However, a month later, we were surprised by their answers: a slim majority (8) felt the upgrade was worth it.

Why do the majority consider the upgrade to be “worth it” in the end? This may be because they weigh the costs and benefits differently than researchers – counts only provide one way of characterising them. Or perhaps they simply resign themselves to periodically installing upgrades (like going to the dentist), and cope with the result. The choice of upgrading is largely irreversible, and undoing it requires significant time and effort. The easiest option may be simply to adapt to the upgrade [23] and rationalise the choice [43].

Perspectives on security vary, leaving users vulnerable

Users and IT managers appear to have divergent perspectives on upgrades. Users worry about user interface changes and the potentially negative impact on their work practices. By contrast, IT managers worry about protecting computers. While the IT managers we interviewed support users who do not control their own machines (which is different than the participants in both of our studies), we know from Khoo et al. [21] that users in an organization who are “forced” to upgrade experience a learning curve followed by elective benefits. This is altogether not dissimilar to what our participants experienced. Users who manage their own machines are clearly vulnerable to security threats, making upgrades essential. This implies that the user upgrade experience must be significantly improved, to encourage users to upgrade on a regular basis. Even if some are relatively satisfied a month after the process, negative impressions remain, causing users to put off the next upgrade.

Design recommendations

We recommend the following to improve the upgrade experience, opening opportunities for new design solutions. While the first two recommendations are consistent with [39] and [40], the others are derived directly from the study results.

Make upgrades reversible: Alert users at the outset if the process cannot be reversed.

De-couple security upgrades from all other upgrades: Facilitate regular installation of security updates, and let users determine when to install user interface updates.

Provide accurate duration times: Give users time ranges (e.g., “2 to 3 hours”) and let them postpone the upgrade when duration is first revealed.

Inform users of progress: Provide clear feedback on progress during download and installation; clarify that these are estimates; and describe what is happening.

Do not force users to babysit the installation: Ask for all relevant user input in one step near the beginning, and store their preferences for later use.

Decrease upgrade times: Improve efficiency or run upgrade activities in the background.

Leverage “dead time”: Take advantage of times that do not require user input to describe changes or tutor new features.

Let users control feature changes: Provide transparency, giving users informed consent with respect to any substantive changes in the user interface or privacy settings.

Inform users of security improvements: Clarify any important but invisible security benefits, and consider reminding users of these 1-2 weeks post-upgrade, to balance other potentially negative experiences.

CONCLUSION

Our first study shows that Mac users delay major OS upgrades by 80 days on average. Our second study provides the first detailed observation of the upgrade process, with a follow-up longitudinal diary component that together capture the user’s experience of OS upgrades. The study shows that the upgrade process often breaks basic design principles [31], with long durations, poor feedback, and fuzzy mental models, causing frustration and confusion. After an upgrade, the costs outnumber the benefits. While some users deem the effort of upgrading worth in the end, the majority predict a change in their attitudes towards future upgrades because of their experience. Further, our preliminary interviews with IT security personnel suggest that users and IT staff have divergent views on upgrades. IT staff focus on security issues and protecting computers. By contrast, users consistently ignore security and focus on user interface changes. We address these findings with novel design recommendations, that open a design space around upgrades and provide the basis for future research with broader user populations.

Millions of users must manage their own software upgrades. Our design recommendations for improving the user experience of upgrades should lead to better upgrade mechanisms that will set the right balance between user costs and benefits. A better user experience will encourage more people to install upgrades in a timely manner, helping to ensure that more devices are secure and up to date.

ACKNOWLEDGMENTS

This work was partially supported by European Research Council (ERC) grant n° 321135 CREATIV: Creating Co-Adaptive Human-Computer Partnerships. We thank the participants for their time, the anonymous reviewers for their insights, as well as the ExSitu and MUX labs for their help.

REFERENCES

1. Ron Amadeo. 2016. Android N borrows Chrome OS code for “seamless” update installation | Ars Technica. <http://arstechnica.com/gadgets/2016/05/android-n-borrows-chrome-os-code-for-seamless-update-installation/>. (May 2016). (Accessed on August 2, 2016).

2. Apple. 2016. Get help with over-the-air iOS updates - Apple Support. <https://support.apple.com/en-us/HT201435>. (June 2016). (Accessed on August 2, 2016).
3. Harry Brignull. 2013. Dark Patterns: inside the interfaces designed to trick you | The Verge. <http://www.theverge.com/2013/8/29/4640308/dark-patterns-inside-the-interfaces-designed-to-trick-you>. (Aug. 2013). (Accessed on July 28, 2016).
4. Brad Chacos. 2016. Forced Windows 10 upgrades push users to dangerously disable Windows Update. <http://www.pcworld.com/article/3075729/windows/fearin-g-forced-windows-10-upgrades-users-are-disabling-critical-updates-at-their-own-risk.html>. (May 2016). (Accessed on August 1, 2016).
5. David Coburn. 2014. How Apple’s OS X Yosemite tracks you. https://www.washingtonpost.com/posttv/business/technology/how-apples-os-x-yosemite-tracks-you/2014/10/22/66df4386-59f1-11e4-9d6c-756a229d8b18_video.html. (Oct. 2014). (Accessed on August 8, 2016).
6. Shane Dingman. 2016. Considering updating to Windows 10? The choice may be out of your hands - The Globe and Mail. <http://www.theglobeandmail.com/technology/tech-news/considering-updating-to-windows-10-the-choice-may-be-out-of-your-hands/article28533377/>. (Feb. 2016). (Accessed on August 1, 2016).
7. Thomas Duebendorfer and Stefan Frei. 2009. Why silent updates boost security. *TIK, ETH Zurich, Tech. Rep* 302 (2009).
8. Kevin Dunn. 2004. Automatic update risks: can patching let a hacker in? *Network Security* 2004, 7 (2004), 5–8.
9. Michael Fagan, Mohammad Maifi Hasan Khan, and Ross Buck. 2015. A study of users’ experiences and beliefs about software update messages. *Computers in Human Behavior* 51 (2015), 504–519.
10. Michael Fagan, Mohammad Maifi Hasan Khan, and Nhan Nguyen. 2015. How does this message make you feel? A study of user perspectives on software update/warning message design. *Human-centric Computing and Information Sciences* 5, 1 (2015), 1.
11. Marvin Fleischmann, Miglena Amirpur, Tillmann Grupp, Alexander Benlian, and Thomas Hess. 2016. The role of software updates in information systems continuance—An experimental study from a user perspective. *Decision Support Systems* 83 (2016), 83–96.
12. Marita Franzke and John Rieman. 1993. Natural training wheels: Learning and transfer between two versions of a computer application. In *Human Computer Interaction*. Springer, 315–328.
13. Barbara L Fredrickson and Daniel Kahneman. 1993. Duration neglect in retrospective evaluations of affective episodes. *Journal of personality and social psychology* 65, 1 (1993), 45.
14. Stefan Frei, Thomas Duebendorfer, and Bernhard Plattner. 2008. Firefox (in) security update dynamics exposed. *ACM SIGCOMM Computer Communication Review* 39, 1 (2008), 16–22.
15. Samuel Gibbs. 2016. Windows 10: Microsoft launches intrusive full-screen upgrade reminder | The Guardian. <https://www.theguardian.com/technology/2016/jul/04/microsoft-windows-10-full-screen-upgrade-notification-pop-up-reminder>. (July 2016). (Accessed on August 1, 2016).
16. Rob Griffiths. 2016. A useless analysis of OS X release dates | The Robservatory. <http://robservatory.com/a-useless-analysis-of-os-x-release-dates/>. (July 2016). (Accessed on August 1, 2016).
17. Iulia Ion, Rob Reeder, and Sunny Consolvo. 2015. “... no one can hack my mind”: Comparing Expert and Non-Expert Security Practices. In *Eleventh Symposium on Usable Privacy and Security (SOUPS 2015)*. 327–346.
18. Daniel Kahneman, Barbara L Fredrickson, Charles A Schreiber, and Donald A Redelmeier. 1993. When more pain is preferred to less: Adding a better end. *Psychological science* 4, 6 (1993), 401–405.
19. Moazzam Khan, Zehui Bi, and John A Copeland. 2012. Software updates as a security metric: Passive identification of update trends and effect on machine infection. In *MILCOM 2012-2012 IEEE Military Communications Conference*. IEEE, 1–6.
20. Huoy Min Khoo and Daniel Robey. 2007. Deciding to upgrade packaged software: a comparative case study of motives, contingencies and dependencies. *European Journal of Information Systems* 16, 5 (2007), 555–567.
21. Huoy Min Khoo, Daniel Robey, and Srinivasan Venkoba Rao. 2011. An exploratory study of the impacts of upgrading packaged software: a stakeholder perspective. *Journal of Information technology* 26, 3 (2011), 153–169.
22. Wendy E. Mackay. 1991. Triggers and Barriers to Customizing Software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)*. ACM, New York, NY, USA, 153–160. DOI: <http://dx.doi.org/10.1145/108844.108867>
23. Wendy E Mackay. 2000. Responding to cognitive overload: Co-adaptation between users and technology. *Intellectica* 30, 1 (2000), 177–193.
24. Andrew Martonik. 2015. Hey Google, don’t just change fundamental parts of my launcher without warning | Android Central. <http://www.androidcentral.com/google-dont-just-change-my-launcher-without-warning>. (September 2015). (Accessed on September, 10 2016).
25. Joanna McGrenere, Ronald M. Baecker, and Kellogg S. Booth. 2002. An Evaluation of a Multiple Interface Design Solution for Bloated Software. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. ACM, New York, NY, USA, 164–170. DOI: <http://dx.doi.org/10.1145/503376.503406>

26. Joanna McGrenere and Gale Moore. 2000. Are we all in the same "bloat"? In *Graphics interface*, Vol. 2000. 187–196.
27. Microsoft. 2016. Upgrade to Windows 10: FAQ. <https://support.microsoft.com/en-us/help/12435/windows-10-upgrade-faq>. (July 2016). (Accessed on August 2, 2016).
28. Andreas Möller, Florian Michahelles, Stefan Diewald, Luis Roalter, and Matthias Kranz. 2012. Update behavior in app markets and security implications: A case study in google play. In *Proc. of the 3rd Intl. Workshop on Research in the Large. Held in Conjunction with Mobile HCI*. Citeseer, 3–6.
29. Brad A. Myers. 1985. The Importance of Percent-done Progress Indicators for Computer-human Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '85)*. ACM, New York, NY, USA, 11–17. DOI: <http://dx.doi.org/10.1145/317456.317459>
30. Antonio Nappa, Richard Johnson, Leyla Bilge, Juan Caballero, and Tudor Dumitras. 2015. The attack of the clones: a study of the impact of shared code on vulnerability patching. In *2015 IEEE Symposium on Security and Privacy*. IEEE, 692–708.
31. Jakob Nielsen. 1995. 10 Heuristics for User Interface Design. <https://www.nngroup.com/articles/ten-usability-heuristics/>. (Jan. 1995). (Accessed on August 12, 2016).
32. Donald A Norman. 1983. Some observations on mental models. *Mental models* 7, 112 (1983), 7–14.
33. Robert Plutchik. 2001. The Nature of Emotions Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist* 89, 4 (2001), 344–350.
34. Jonathan Posner, James A Russell, and Bradley S Peterson. 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology* 17, 03 (2005), 715–734.
35. John Siracusa. 2014. OS X 10.10 Yosemite: The Ars Technica Review | Ars Technica. <http://arstechnica.com/apple/2014/10/os-x-10-10/9/#spotlight>. (Oct. 2014). (Accessed on August 1, 2016).
36. Ashkan Soltani and Craig Timberg. 2014. Apple's Mac computers can automatically collect your location information - The Washington Post. <https://www.washingtonpost.com/news/the-switch/wp/2014/10/20/apples-mac-computers-can-automatically-collect-your-location-information/>. (Oct. 2014). (Accessed on August 1, 2016).
37. Franck Tétard and Mikael Collan. 2009. Lazy user theory: A dynamic model to understand user selection of products and services. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*. IEEE, 1–9.
38. Yuan Tian, Bin Liu, Weisi Dai, Blase Ur, Patrick Tague, and Lorrie Faith Cranor. 2015. Supporting Privacy-Conscious App Update Decisions with User Reviews. In *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices*. ACM, 51–61.
39. Kami Vaniea and Yasmeen Rashidi. 2016. Tales of Software Updates: The Process of Updating Software. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3215–3226. DOI: <http://dx.doi.org/10.1145/2858036.2858303>
40. Kami E. Vaniea, Emilee Rader, and Rick Wash. 2014. Betrayed by Updates: How Negative Experiences Affect Future Security. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2671–2674. DOI: <http://dx.doi.org/10.1145/2556288.2557275>
41. Rick Wash, Emilee J Rader, Kami Vaniea, and Michelle Rizor. 2014. Out of the Loop: How Automated Software Updates Cause Unintended Security Consequences.. In *SOUPS*. 89–104.
42. Lance Whitney. 2016. Microsoft's Windows 10 upgrades are getting even more sneaky-pushy - CNET. <http://www.cnet.com/news/microsoft-windows-10-upgrade-s-get-more-sneaky-pushy/>. (May 2016). (Accessed on August 1, 2016).
43. Timothy D Wilson and Daniel T Gilbert. 2005. Affective forecasting knowing what to want. *Current Directions in Psychological Science* 14, 3 (2005), 131–134.