"No Good Reason to Remove Features"
Expert Users Value Useful Apps over Secure Ones

Steve Dodier-Lazaro, Ingolf Becker, Jens Krinke(✉), and M. Angela Sasse

University College London, London WC1E 6BT, UK
s.dodier-lazaro@cs.ucl.ac.uk, j.krinke@ucl.ac.uk
http://www0.cs.ucl.ac.uk/

Abstract. Application sandboxes are an essential security mechanism to contain malware, but are seldom used on desktops. To understand why this is the case, we interviewed 13 expert users about app appropriation decisions they made on their desktop computers. We collected 201 statements about app appropriation decisions. Our value-sensitive empirical analysis of the interviews revealed that *(a)* security played a very minor role in app appropriation; *(b)* users valued plugins that support their productivity; *(c)* users may abandon apps that remove a feature – especially when a feature was blocked for security reasons. Our expert desktop users valued a stable user experience and flexibility, and are unwilling to sacrifice those for better security. We conclude that sandboxing – as currently implemented – is unlikely to be voluntarily adopted, especially by expert users. For sandboxing to become a desirable security mechanism, they must first accommodate plugins and features widely found in popular desktop apps.

Keywords: Value-Sensitive Design · Security · Productive security · Sandboxing · Apps · Appropriation

1 Introduction

Sandboxes are security mechanisms that execute processes in a fully controlled and isolated environment. They are typically used to isolate apps from one another on operating systems (OSs). They protect users both against malicious apps and against exploits targeting legitimate apps. Sandboxes have become an essential building block of modern OSs [2,3,12,19]. However, sandboxes impact how app features can be implemented, and sometimes prevent the implementation of features found in apps, because the methods used to implement those features are also useful for malware writing. Therefore, sandboxed and unsandboxed versions of the same app can differ slightly in behaviour or affordances.

The security benefits of sandboxes are tangible. On Mobile OSs, all apps are sandboxed, which prevents malware-ridden and malicious apps from affecting other apps on the system. On desktop OSs, however, sandboxes are only partially deployed. Desktop developers struggle to make their apps compatible with

T. Tryfonas (Ed.): HAS 2017, LNCS 10292, pp. 25–44, 2017.
DOI: 10.1007/978-3-319-58460-7_3

sandboxing without sacrificing important features and plugins. Many ultimately opt out from supporting this security feature [8,13,15,20,26]. Plugin infrastructures (which allow third-party developers to augment an app with additional features or user experience improvements) and features such as emulating keyboard input, screen sharing, audio recording, inter-process communication and bulk file processing are forbidden in sandboxes to prevent malicious behaviours, but they are sometimes too critical for apps to abandon [22,31]. These incompatibilities are not, *per se*, technological constraints that cannot be overcome. They are design decisions made by sandbox designers. Instead, designers could have chosen to complicate sandboxed apps' security policies to support those potentially dangerous features.

On Windows, many popular apps like Dropbox, Steam, iTunes, Google Drive, VLC, Office, Photoshop, etc. are not sandboxed, or only in rudimentary versions with missing features [14]. Tech reporters argued that sandboxed apps are rarely downloaded and used on Windows, as they lack critical features and degrade productivity [6]. After five years, the adoption of sandboxing stagnates on Windows, and even dwindles on OS X where developers have publicly announced abandoning the Mac App Store [20,26,31]. On Linux desktops, sandboxed app stores exist [7,9,10], but none have a substantial user base. Consequently, desktop users are not currently taking advantage of the security benefits of sandboxes, despite being exposed to phishing attacks, malware, ransomware, etc. Still, many productive activities such as software development, complex information work, data science, etc. require the use of desktop OSs.

Moreover, assuming sandboxing meets usability requirements, users still need to either abandon their current apps in favour of new, sandboxed apps. How users will arbitrate such decisions about app adoption or retainment has not been addressed in past research.

We hypothesise that developers refuse to support sandboxing because it would degrade what makes their apps valuable to their users. Our analysis of developer discussions on sandboxing revealed two main issues: some types of features cannot be implemented in sandboxed apps, and sandboxed apps cannot have plugins. If the consequences of sandboxing upset users or make apps useless, it would explain why developers are reluctant to support it.

To answer these questions, we interviewed 13 expert users to explore the values they seek to fulfil when they make choices about apps. We aim to unveil the *de facto* requirements that sandboxed apps must meet in order to entice user adoption, support app adaptation needs, and prevent app abandonment.

We show that our users struggle with explaining and accepting feature loss, and may choose to abandon apps that remove features – especially for security reasons. We show that plugins are useful and valuable to expert users, and are a crucial way to improve their productivity. We also show our participants do not consider security as a prime factor in their decisions related to app appropriation.

We also make the following contributions: we perform a value-sensitive analysis of app adoption, adaptation via plugins and abandonment. We find that different values underpin each of these processes, and that the values recruited to

think about content consumption and production apps differ. We identify shortcomings in past usable security research: temporal aspects of appropriation (e.g. use of plugins, which address issues that were experienced in use and reflected upon by users) can only be studied in-the-wild; and participants' appreciation of security must not be distorted by priming.

We first present relevant research. Next, we explain our study design and research questions. Then, we present our value analysis of three aspects of app appropriation. We continue with a detailed analysis of participants' reactions to feature loss. We finish with a list of limitations, and conclude with a summary of our findings and open problems.

2 Background and Related Work

Usability evaluations of security mechanisms are mostly restrained to their user interfaces. We argue there is more to technology adoption than usable interfaces. If a tool does not perform a function that is useful to users, or if this function conflicts with other valued artefacts, the tool may be ignored. This is why Smetters and Grinter [27] have called for usable security to ensure that designed systems are useful. Likewise, Mathiasen and Bødker [17] examine usable security from the lens of experience-driven design [18]. They "concern [themselves] with how, on the one hand, the use experience is determining the security technology, while on the other hand, the security technology resists, constrains and directs the use experience". By framing sandboxing as an appropriation problem rather than a usability one, we can focus on the compositional and spatio-temporal aspects of user experience, which are usually ignored in usable security.

2.1 The Usability of Sandboxes

Only two usability studies of sandboxes exist [23,25]. Both had participants perform scripted scenarios in a lab, emulating basic app interactions. These studies do not model the impact of introducing sandboxes on the complex app ecosystems of the real-world. Expert users may rely on features that are more demanding on security policies, or sometimes not possible to formulate safely with current app sandbox models. These differences in technological needs are masked by seemingly successful usability studies, but it remains unclear if users would be able to appropriate a fully sandboxed OS.

2.2 Value-Sensitive Design

We did not want to just document participants' preferences, but understand *why* they held such preferences. Value-Sensitive Design (VSD) [11] is a methodology that reveals values involved in user behaviours and the frictions between them. It combines three forms of analysis. Conceptual analysis is used to identify stakeholders, their goals and potential value tensions between them. Empirical analysis reveals tensions in studied environments where technologies are deployed.

Technical analysis probes how artefact designs position themselves with regards to values and value conflicts. We used a VSD conceptual analysis to design the interview we report on, and an empirical study to model the values involved in app appropriation and relate them to security, which we report on here.

3 Study Design

We aim to identify how sandboxes clash with the needs of expert users. We performed semi-structured interviews with 13 users about the apps they use.

3.1 Research Questions

Feature loss and plugin loss are externalities of sandboxing that developers expect and dislike, and thus focus most of our investigation on these aspects. However, other tensions might yet have to be uncovered. We hence explore the relationship between users and their apps more thoroughly, including situations like app adoption and abandonment which are have been ignored in past studies. We treat plugin usage as acts of app adaptation, and thus include their use in our value analysis. If the presence of features emerges as an important value for users, and if plugins play a distinct and important rules in users' practices, it would corroborate developers' worries about these two aspects of apps that conflict with sandboxing.

We first investigate what users *value* and prefer in their apps, and the relation between these values and security. Our research questions are:

RQ1: Which values drive app appropriation? Is security one such value?
RQ2: How much do expert users rely on plugins? What value do plugins provide to expert users?

After that, we turn to how users relate to and *react to* feature removal in their apps. We discuss their own experiences and beliefs, and then explore how they make sense of feature removals motivated by security reasons.

RQ3: Is feature loss acceptable? How does it impact users' choices of apps?
RQ4: How does security-motivated feature loss differ from other types of loss with regard to acceptance and reaction?

3.2 Data Collection and Coding

We performed semi-structured interviews centred around participants' usage of apps, how they manage and value their apps, and about their information management and security strategies. The interviews lasted 40 min to 1:50 h (median 1:14 h), and we collected 81 to 227 statements per participant (median 140).

We coded our data separately for the value analysis and questions about feature loss. In the next section on value analysis, we allocated all participant's statements for each topic to characteristics of the apps that they relate to

(we call those *app traits*), e.g.: apps being slow or costly, or the fact that an app offers new features. We re-coded previous answers and refined app traits as we went along, until all participants answered could be unambiguously classified. We then mapped these app traits to the value they support, to enable a value-sensitive empirical analysis of participants' behaviours. In the section on feature loss, we used Grounded Theory's open coding [30] to identify themes in participants' answers, e.g. how they made sense of feature loss statements or the expected compensations for feature loss.

Self-reported data suffer from accuracy issues. To eliminate potential demand traits biases [21], we only retained strong statements – which participants justified or supported with prior experiences. We eliminated 18 hypothetical, vague or contradictory statements, and used 201 in our findings.

3.3 Recruitment and Demographics

We advertised our study on a Reddit community dedicated to Linux. We used Linux users because participants were recruited as part of a larger field study, parts of which include deploying software components that cannot be written for closed-source OSs. Linux is for this reason the *de facto* standard OS in systems research. We paid participants £20 for participating to the interview this paper is based on, out of a total of £100 for participating to the whole project.

We recruited 13 Xubuntu users from 7 EU countries and from the USA, aged between 18 and 54, representative of desktop Linux users for age, occupation, gender and degree of Linux proficiency. Most were expert Linux users, except P6 and P12 (beginners), and P3 and P10 (IT professionals). P10 and P13 are security experts, and P12 attends security classes. Our participants include a Web developer, two (adult) high school students, two tech support representatives, a musician, a consumer retail employee, a student teacher, a sales engineer and four computer science students. 8 of them write code, 7 perform information work, and 7 produce media content (e.g. graphics, audio, video, photos).

3.4 Use of Deception

We told participants the study focused on their multitasking habits, to avoid non-respondent bias from participants with limited motivation to engage with security, and social desirability biases and demand trait biases during the study. We chose multitasking to attract participants who have a need for productivity, as opposed to leisure users of computers. We revealed the deception to participants near the end of the interview. Unless when mentioned otherwise, all the data we use was obtained before we revealed the deception. This study was approved by the UCL Research Ethics Committee under identifier 6079/001.

4 Value-Sensitive Analysis of App Appropriation

Sandboxes can make an impact in terms of everyday security only if they are *used*, rather than merely *usable*. To this end, we aim to determine how

sandboxing interplays with three aspects of app appropriation: adoption, adaptation and retainment. Sandboxes may conflict with users' ability to obtain features and may incur a performance penalty. If users' adoption and abandonment behaviours are driven by the presence or absence of features and by performance considerations, then sandboxing will conflict with users' main decision factors. This could lead to sandboxed apps being adopted less often, or apps being abandoned after they become sandbox-compatible.

Besides, sandboxes prevent apps from providing plugins. Plugins are part of how apps can be adapted to better suit workflows. Users of plugins must compare the benefits afforded by plugins with the sandbox's benefits and decide whether to adopt or circumvent the sandbox based on such a cost/benefit analysis. We aim to find out where plugins are used, and what value they provide.

4.1 Method

We classified participants' statements on how they appropriate apps and on the plugins they use, based on the app traits they relate to (e.g. "Ad-blocking" or "Access to content" for plugins; "Unresponsive UI" or "Privacy Issues" for app abandonment). For plugins, we paid attention to their *reported purpose*, e.g. P11 uses a VPN service to access foreign media rather than for security. When participants added or replaced components of their desktop Environment (DE), we recorded those events as DE plugins.

Next, we categorised traits into values: *usefulness*, *security & privacy*, *usability*, *productivity*, *credibility*, *affordability*, *mobility*, *stability* and *flexibility*. We chose values to highlight known tensions in the usable security literature (*security* vs. *usability* [1], *usefulness* [27] and *productivity* [4]) and to capture concerns identified in our conceptual analysis (*usefulness* and developers' *credibility*).

We classified apps into categories: browsers, communication apps (email and messaging), file sharing apps (cloud storage and torrent), media consumption apps (e.g. music and video players, news aggregators, etc.), media and document editors (e.g. Office, audio, video, image editors), code editors, DEs and security apps. When a statement refers to an app's feature or to a past experience with an app, we assign it to the category that fits the app.

4.2 App Adoption and Abandonment

We look at the values governing app adoption and app abandonment, in order to discover potential challenges during the transition to sandboxed apps. When developers port their apps to a sandbox, externalities can include features being incompatible, loss of plugins or performance degradation. They must decide if those changes will put users off from adopting or continuing to use their app. Hence, we asked participants what would convince them not to try a new app, and what would convince them to abandon an app they are using.

Losing Interest in Potential Apps. We recorded 20 statements of interest loss. P4 gave no answer, and P2's answers were too weak to be included.

As Fig. 1 shows, half of our 12 respondents stopped considering an app because it lacked a feature. Feature loss is a possibility when porting an app to a sandbox, either because the feature relied on privileged operations (e.g. bulk file processing, access to hardware, IPC) or on libraries that are themselves not compatible with the sandbox. Thus, if an app developer removes a key feature because of sandboxing, fewer users will adopt their app in the future.

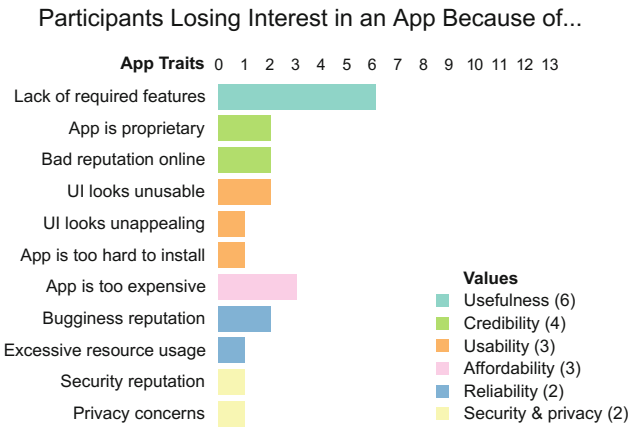

Fig. 1. Participants decided not to install potential new apps primarily because they lacked a required feature. Other reasons revolve around *Credibility* and alleged *Usability* and *Reliability*.

P10 mentioned avoiding apps that have a reputation for "breaking other programs somehow" or "security stuff". He also avoids apps that are hard to install. Apps with such a reputation might benefit from being sandboxed owing to the benefits of app stores. Ultimately however, sandboxes appear more detrimental than beneficial to adoption for our cohort.

Abandoning a Current App. We also analysed what reasons participants have to stop using their current applications, to identify the impact of sandbox introduction for the current users of an app. 11 participants provided 21 statements on app abandonment. P2's data was again removed.

Figure 2 shows that *Reliability* is the primary factor for app abandonment: participants stopped using apps because they became too slow, buggy, or used too much RAM. *Usefulness* follows in users' reasons for app abandonment. It is by changes in apps or in user needs. Two participants no longer needed an app, and two had a better replacement available. Five abandoned an app because it was missing a feature (in four cases, it was lost to an update; in one case, it was

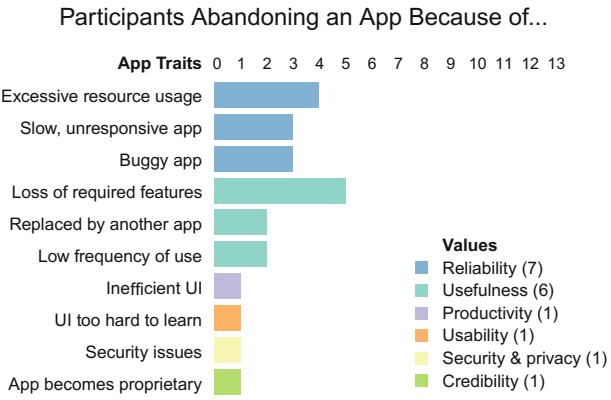

Fig. 2. Participants stopped using applications primarily because of *Reliability* issues: bloated apps, unresponsive or buggy UIs. Apps also fell out of use, or lost required features after an update.

only partially implemented). *Security* was mentioned only once spontaneously as a good reason to abandon an app. Two other participants stated security was a good reason after we accidentally primed them.

4.3 Using Plugins to Customise Apps

Expert users commonly install plugins on their apps to improve them. Plugins are routinely found on browsers, but also code editors, media editors, information work apps, communication apps, media players, etc. They are written by third-party developers, and are banned from the Windows App Store, the OS X App Store (partially) and on Mobile platforms. Browsers run unsandboxed in order to retain the ability to provide plugins.

Our participants reported using 73 plugins (2 to 9, average 5), for all app categories except media consumption apps (46 for browsers; 14 for code editors; 2 to 4 for communication apps, document editors, DEs and security apps). When asked, seven participants mentioned 11 additional plugins they would like to have. Participants plausibly had more plugins installed than they recalled, as many Linux productivity apps and media players are distributed with some plugins enabled by default. If all Linux apps were sandboxed, participants would resultingly miss out on a significant part of their user experience. In this section, we document the role of plugins to understand how users would be affected if they chose to adopt sandboxed apps. This informs us on the values that security mechanisms compete against when they compromise the ability to have plugins.

Desired Plugins and Features. We asked participants to imagine an additional feature or plugin they would like to have, to check if specific types of features are in demand, or if plugins are wanted for specific app categories. Plugins were desired for browsers, communication apps, code editors and DEs.

We found that the 73 installed plugins and 11 desired plugins and features were similar in terms of the values they support and concerned similar app categories. Consequently and for space reasons, we discuss 'installed plugins' and 'desired plugins' together in this paper.

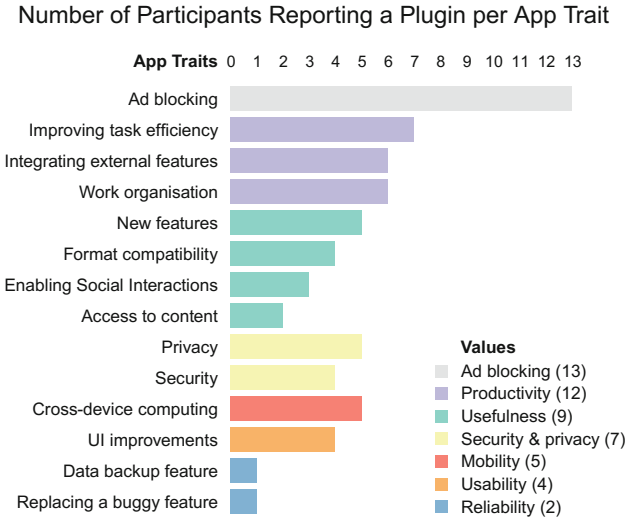

Fig. 3. The plugins installed and wanted by our participants primarily support *Ad-blocking*, *Productivity* (task efficiency, external features, work organisation) and *Usefulness* (new features, format compatibility, access to content, social interactions).

The Role of Plugins. Plugins were predominantly used for browsers, but also for content production apps such as code or image editors and for communication apps. The features provided by plugins supported a variety of app traits, e.g. making an app compatible with a new format. Our classification aims to show what exactly participants would lose if plugins were removed. Some types of users or some apps' userbases may be more affected than others. We highlight the app traits for which sandboxes may be able to replace plugins with other techniques. We counted how many participants mentioned each trait and assigned traits to values, as shows Fig. 3. The Ad-blocking trait was mentioned by all participants and not classified into a value due to its unique nature.

Plugins mostly support the *productivity* value, with three traits relating to it. Firstly, plugins help participants perform small tasks faster, e.g. code snippets or tools to accelerate browsing through Web pages. Secondly, they integrate features normally found in other apps to make them faster to access, e.g. image processing features in browsers or source version control in code editors. Thirdly, plugins help participants organise their work sessions in browsers, DEs and code editors, e.g. tools to manage tabs or improve window placement.

Plugins also support *Usefulness*, with traits such as the compatibility with new document formats, enabling new social interactions, granting access to copyrighted content, and with the introduction of new features. *Security* plugins consisted of script and Flash blockers, HTTPS Everywhere, and a password manager. *Privacy* plugins comprised end-to-end encryption for instant messaging and email apps and of plugins to prevent user tracking on the Web and on Facebook. Sandboxes can partially emulate some features of network security plugins, albeit without proper integration into apps' UIs. They cannot compensate for the loss of plugins in the *Usefulness* category.

Accounting for Productivity Apps. Our participants used plugins for code editors and document and media editors, as well as DEs and browsers. We call both editor categories 'production apps' – apps used in productivity contexts. Browsers, DEs and communication apps are hybrid, relevant to all sorts of use contexts. Media consumption apps (music and media players, online social networks, news aggregators, etc.) are, themselves, rarely ever useful in productivity contexts. Even though plugins are available for most of the media consumption apps mentioned by our participants, none of them used plugins for this category. Thus, plugins are particularly in demand for production apps. This is especially true for code editors where 6/8 participants used plugins. The *Productivity* value also accounted for 7/15 plugin mentions for the code editor category. Therefore, users of code editors are particularly dependent on plugins to boost their productivity. They would be more affected than others by plugin loss.

4.4 Values Driving Appropriation over Time

We recorded other value statements that are not specific to adoption, abandonment or plugins. Two values were frequently mentioned: stability and flexibility.

6 participants expressed, in 8 statements, discontent when their user experience is disrupted by changes in apps, therefore preferring *stable* experiences. P7 and P5 expressed disbelief about feature removal. P5 said: "If there is a need and there something covering this need, if you remove it it's really hard to explain to your users that it's just not there any more". Three participants were attached particularly to a specific feature (e.g. the ability to browse books or albums by their cover for P5, or the reopening of documents at the page they were last closed for P10) while we discussed their work habits. Finally, P13 expressed not wanting to change the apps he was habituated to, and disliking when those apps' UI changed after an update.

4 participants also praised, in 6 statements, software that is *flexible* and can be adjusted to their needs. P4 and P12 told us how they take advantage of settings and plugins to speed up keyboard-driven workflows. P4, P5, P12 and P13 mentioned customising applications like their document editors or DE. P5, for instance, says "I have been able to basically make my own toolbars with everything that I use. That's really flexible. [...] And it's pretty much the same idea in all applications".

4.5 Summary of Findings

RQ1: Which values drive app appropriation behaviours? Is security one such value? We found apps are:

adopted if they are *useful*, appear *usable* and *affordable*, and have a reputation of *reliability*, *security* and *credibility*
adapted with plugins to boost *productivity* and *usefulness* and sometimes to provide *security* and *ad blocking* capabilities
abandoned when they lose their *usefulness* or *reliability*

Users also valued a *stable* user experience, and *flexible* apps that can be adjusted to their needs.

RQ2: How much do expert users rely on plugins? What value do plugins provide to expert users? All our participants used plugins – for browsers, DEs and all types of editors, but not for media consumption apps. Plugins mainly provide usefulness and productivity. They also provide ad-blocking in browsers, and security for Internet-facing apps. Few of the benefits provided by plugins could be replaced by other mechanisms, if plugins were to become unavailable.

Productivity plugins were more prevalent for productivity apps and DEs, and our participants were in demand for more productivity plugins than they already had. Thus, people who use computers for productive work, and specifically users of some types of apps, would see their productivity decrease if they no longer had access to plugins.

4.6 Implications for Sandboxing

Sandboxing threatens *usefulness* by preventing the implementation of some features, *reliability* by degrading performance and resource usage, and *stability* by causing developers to transform or drop some features. Sandboxes thus conflict with the values recruited by participants when they decide to adopt and abandon apps. Owing to their effects on plugins, sandboxes further threaten *productivity* and *usefulness*, the main values supported by the use of plugins. Developers who chose to drop features and plugins to support sandboxing will be confronted to loss of users and potential new users, according to our value-sensitive analysis.

Our participants' liking of *stability* suggests sandbox designers shouldn't expect user experience sacrifices as a prerequisite to sandbox adoption. Mobile OSs never had plugin infrastructures, and so their users have adopted what was available. Android and iOS are dominated by media consumption apps [28,29], and since there is no plugin demand for consumption apps, plugins are not as crucial for Mobile OSs as they are for desktops. Users might refuse to switch to sandboxed versions of desktop apps if this means losing plugins they have already integrated into their work practices.

Plugin loss will particularly affect users with productivity goals, and some demographics e.g. users who write code (and expectedly, over demographics that were not represented in our cohort). When productivity is put in competition

with security, users respond by implementing "shadow security" practices, which involve disengagement from sanctioned, verified security mechanisms, even if they do value security [16]. It is advisable that plugins be supported by sandboxes, especially since there is no technical barrier to distributing plugins on the Windows and Mac App Stores, just like standalone apps.

5 Feature Loss

We learnt that usefulness is a major driver of appropriation decisions, and we know that sandboxes conflict with usefulness by forbidding some features. We now explore the value arbitrations made by participants when they are confronted with feature loss. We query how they explain feature loss in an app they use, and how they react to it, especially if "security reasons" motivate it.

5.1 Method

We asked participants, if a feature was removed from an application, what they would do and how it would affect them. We also asked them what good and bad reasons a developer could give to justify this change. When possible, we asked participants about features they mentioned during the interview. Otherwise, we would ask about "a feature" or "the ability to have plugins" for an app they mentioned using. Most participants responded with hypothetical scenarios based on apps they used.

We formulated the security question as such: we asked participants what they would think if a developer were to remove a feature or plugin "for security reasons". P12 spontaneously mentioned security as a valid reason for removing a feature, obviating the security question. P5 and P9 were mistakenly asked about justifications to feature removal after we had revealed the security deception.

We refer to answers based on participants' features as "own experiences", and answers to the security question as "security reasons". As the interviews were semi-structured, some participants did not answer, especially P3 and P11.

5.2 Justifying Feature Removal

We wanted to know what determined whether users would accept the disappearance of a feature. If a specific reason makes sense to users, they will be less incredulous and suspicious when a feature is removed for that reason. Inversely, if users are told a feature is removed for a reason they do not understand, they might deplore the developer's decision and be more prone to switch apps.

We collected 18 reasons which participants thought were acceptable (see Fig. 4) and 8 unacceptable (Fig. 5) to justify feature removals. 5 participants recalled actual experiences of feature loss, showing it is a commonplace experience, though overall participants did not find it easy to answer those questions.

Maintainability was seen as the most valid reason to remove features, by 3 participants, with 2 mentions from P4. This included removing code that was

too difficult to maintain or not stable enough, or making plugins temporarily unavailable after an update. However, one of the "feature loss" app abandonment reasons we discussed in the previous section was justified with maintainability: P4 abandoned the GNOME DE because its plugins would often stop working after an update. So the reason is not unanimously accepted.

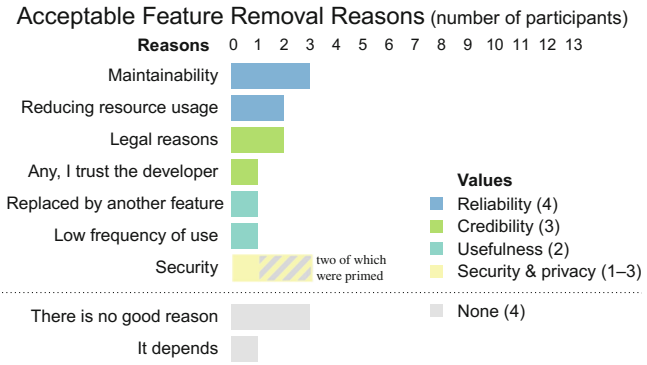

Fig. 4. Number of participants citing a reason as acceptable to justify feature removal.

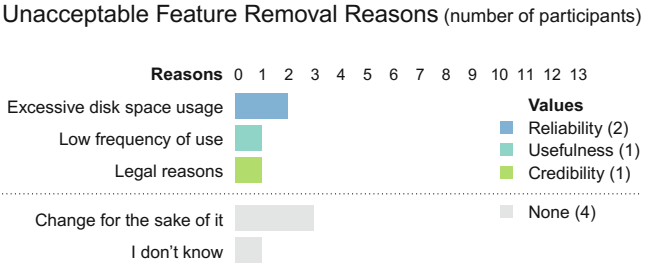

Fig. 5. Number of participants citing a reason as *not* acceptable.

Security was mentioned thrice, albeit two times by participants whom we accidentally primed to think about security beforehand, as we forgot to ask the question about feature removal until right after revealing the security topic of the study and before discussing security practices. Legal reasons were mentioned both as a good and as a bad justification. So was reliability, with participants claiming that excessive CPU or RAM usage were valid reasons, but excessive disk usage wasn't. Likewise for usefulness: P6 mentioned not caring about a feature he did not use, whereas P12 strongly opined that developers should not remove a feature used only by a minority of users.

Participants could conceptualise why feature are removed (maintainability, legal issues, reliability, and security), but none of the enumerated reasons seem to be *always* justified. Besides, three participants thought feature removal to be

inexcusable, no matter the reason. Therefore, there is no *blanket rationale* that developers can invoke to explain away a decision to remove a feature.

5.3 Security Reasons

We asked eleven participants (except P3 and P11) what they would think of a scenario where a feature is removed for "security reasons".

Are Security Reasons Really Accepted? Eight participants considered security to be a good reason to remove a feature, once we asked them. The three others did not answer the question, but described how they would analyse the feature loss instead. None found it explicitly unacceptable. Yet, only P12 mentioned security spontaneously – as well as P5 and P9 right after we primed them. Security might be a positive value to our participants, but it is not something they think about when features are affected by updates.

Making Sense of "Security Reasons". Even though participants agreed security was an acceptable justification, they sounded negative about it. We had expected them to state that they would no longer use the insecure software. Instead, they showed us they would attempt to understand the announcement and to decide for themselves if they should be concerned and adjust their practice.

Participants were mostly defiant because of how they made sense of "security reasons". They understood security as *incident response*, rather than the *anticipation* of risks that have not yet materialised, or *compliance* with external constrains. Yet, sandbox feature constraints derive from risk management considerations rather than security vulnerabilities.

Three participants clearly expressed the idea that the security risk had resulted in exploitation, using words such as "malware", "breach" or "security exploit". Three more talked of a "vulnerability" or "security hole" and wondered if their data could be compromised as a result. Only P8 pondered that the feature itself might have represented a danger, without mentioning the existence of a fault attributable to the developer.

5.4 Deciding What to Do About Feature Removals

How many users would abandon an app if its developers decided to remove an important feature from it? The answer to this question is relevant to developers who must decide whether to adopt feature-degrading sandboxes or not. We thus asked our participants how they would react to the loss of a feature they had previously mentioned to us, or to the loss of plugins. We sometimes asked participants about more than one feature. Figure 6 presents the 20 reactions we collected from 11 participants for feature loss in general (some participants answered for several features, P9 gave weak answers, P11 was not asked). It also shows the 11 reactions collected for security-induced feature loss from 9 participants (P1 gave two answers, and P3, P9, P11 and P13 gave none).

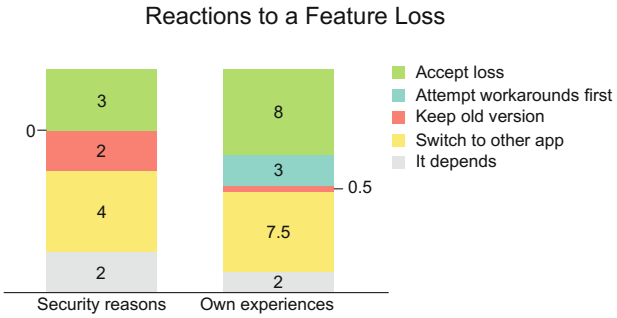

Fig. 6. Participants are more likely to accept an update that induces feature loss for reasons other than security. Some will deploy workarounds to emulate or replace the lost feature, before seeking a replacement app. Over a third of participants would abandon an app that lost a feature and seek another one with an equivalent feature either way.

For updates motivated by security reasons, participants decided to stay on the old, insecure version of the app in 2/10 cases. In 4/10 cases, they preferred switching to another app. 2/10 said their reaction would depend on the feature or the developer's attitude. This leaves only 3/10 cases where participants would accept the update. This reaction contradicts our finding that nearly all participants agreed security is a valid reason to remove features. We hypothesise this discrepancy is due to usefulness taking precedence over security in driving participants' choices. Another possible conjecture is that our expert users have become *prejudiced* against security announcements, owing to dissonance between alleged and perceived security benefits in past security experiences.

The cost of feature loss was viewed as higher than the security benefits in our security question. In contrast, when we asked about feature removal in a generic update, participants rooted for the imagined benefits of the update more often: they would use the new version in 11/21(52%) cases – including 3/21(11%) cases where they would attempt to emulate the lost feature with the new version, but would switch back to the old one or to a new app if their coping mechanism fails to satisfy them. Security is, after all, a secondary goal [24,32], so it comes after features which support a primary goal. Our value analysis corroborates this: factors like usefulness, productivity or reliability trump security in participants' decisions. P10 would either switch to another app or stay on the old version. In 7.5/21(36%) cases in total, participants would switch to another app. In 2/21(10%) cases, participants said it depends on the feature.

5.5 Getting Something Out of the Loss

In both conditions, three participants expected lost features to be re-introduced after some time. When a disruption is temporary, participants might tolerate it as a necessary evil. P1, P10 and P13 also expected the app to be improved in some way (e.g. reducing RAM usage, speeding up the UI, or integrating popular

plugins into the app) in the general case. This desire for *compensation* was not seen in the security condition, as a security benefit was already communicated.

P5, P10 and P13 wanted developers to explain what the vulnerability was that had been fixed. Other participants sought to convince themselves of the well-foundedness of the security reason. P7 stated he expected to be told "how much time has there been a security breach, why have they not warned me beforehand, and what happens now". P12 said "they'd have to justify it pretty well". P2, P8 and P10 said they would look into the issue to decide if they should feel concerned. Overall, those participants had untrusting attitudes towards developers who announced security updates.

5.6 Summary of Findings

RQ3: Is feature loss acceptable? How does it impact users' choices of apps and practices? Feature removal has a substantial impact on users: over a third may abandon an app when a feature they used disappears. Half won't consider updating an app with a missing feature, and they may also abandon an app that loses a feature. A forth of participants expected feature loss to be temporary, and a forth also expected it to be compensated with improvements.

There is no consensus among participants over what constitutes good reasons to remove a feature. Maintainability, reliability and legal issues were mentioned, although as bad reasons too. Security was mentioned spontaneously by one participant, and after security priming by two more participants. Given the prevalence of *stability* in user values, we find feature loss hard to justify overall.

RQ4: How does security-motivated feature loss differ from other types of loss with regard to acceptance and reaction? When asked, our participants claim security is a valid reason to remove features. Yet, they are four times more likely to ignore a security update than a non-security update that removes features. This illustrates how security is a secondary goal to users.

Participants view security-motivated feature removals as incident response rather than a preventative measure. They expect developers to explain why a security risk existed and the consequences if it. Thus, developers' credibility may paradoxically suffer when they announce security improvements.

5.7 Implications for Sandboxing

Sandboxes restrain the ability to implement some features as a form of *risk management*, rather than because these features cause systematic vulnerabilities. As our participants understand security as *incident response*, they are likely to attribute a sandbox-related feature loss to a fault on behalf of app developers. Besides, we've seen that there is no *blanket rationale* that developers can invoke to explain away a decision to remove a feature, which all participants would believe is legitimate. Therefore, the task of explaining a sandbox-motivated feature loss to users seems particularly strenuous and hazardous for developers.

Feature removal can lead to user base attrition. As we've seen, this is more so the case when feature loss is justified by security. In competitive app ecosystems

where many apps provide similar features, having to remove features from one's app may act as a deterrent for developers to consider sandboxing. We argue that the current restrictions on features and plugins place an unfair burden on app developers, and that sandbox designers must review those decisions rather than wait out for developers to finally 'get it' and adopt sandboxing. Presently, there are valid incentives in place for app developers to stay away from sandboxing.

6 Limitations

6.1 Cohort Size

The field study we are running involves sustained interactions with participants, forcing us to keep a small cohort. We thus have too few participants to provide statistical significance for our results. We provide quantitative data as much as possible to allow for our results to be aggregated to future studies on this topic. Besides, we view the presentation of our method as a contribution in itself, relevant to security designers who need to study barriers to the adoption of security technologies in their app ecosystems.

6.2 Deception

We ensured the validity of our data by using deception. This means less data was available as we could not incite our participants to detail their mental models of security without drawing their attention to our actual topic of interest.

6.3 Method of Report

App appropriation events are rare, and participants sometimes struggled to recall details of their past experiences. We helped them recall past events by using diary data to discuss the apps which we knew they used, and we eliminated statements where participants sounded hesitant or were inaccurate.

6.4 Linux Users

We recruited Linux users. They are reflective about technology and often have experience with multiple OSs. This is not a threat to validity, but reduces the scope of our findings to experienced and reflective practitioners. Many Windows and OS X users are experts, too – including developers, digital artists, researchers, etc. Linux users prefer software that is open-source. Thus, our data likely overstates the importance of the app traits related to proprietary licenses.

7 Implications for Usable Security Research

Some of our findings would not have been possible to make if we had stuck to the methods used in previous sandbox usability research [23,25]. We derive methodological implications for future usability evaluations of security mediators.

7.1 Productive Security Is Achieved over Time, Not in the Lab

Beautement et al. [5] argue that the cost of security might be accepted during initial interactions, but rejected over time as users wear out their "compliance budget" – their ability to comply with security when the cost of it exceeds its benefits. When newly introduced security artefacts disrupt stability (e.g. with feature loss) or flexibility (e.g. by removing plugins), these artefacts cannot be declared usable solely on the basis of one-off interactions in a lab setting. Those values are fulfilled over time, and so the impact that changes in users' practices have on them must be studied over time too.

Previous usability studies of sandboxing [23,25] failed to study how participants ultimately react to the cumulative frustrations caused by a degraded user experience, or how they can improve their productivity once sandboxes hinder apps' flexibility. Ergo, sandboxes must be introduced in-the-wild and their impact on practice monitored until they are completely appropriated or rejected by participants. Otherwise, researchers may falsely conclude that sandboxes are usable, when participants' compliance budget is exhausted in superficial interactions settings and their interaction would not have been sustained in-the-wild.

7.2 Deception Is Necessary to Discover Actual Behaviour Drivers

Participants overwhelmingly agreed that security is an acceptable reason to remove a feature, when we asked them. Yet, they would be less likely to continue using an app that lost a feature for security, rather than for other types of improvements. We conclude from that that querying participants directly about their attitude to security can mislead researchers into thinking that security is sufficiently valued to influence user behaviour. We've shown that explicit attitudes towards one value are not the proper measure for drivers of behaviour. Instead, researchers should focus on building value hierarchies and identifying the main values that users recruit in making decisions that impact security. This means that study designs must include deception to avoid non-respondant and social desirability biases, and to produce valid value hierarchies.

8 Conclusion

Sandboxes do not provide support for several types of features, and for plugins, resulting in second-class apps. Sandboxes also decrease app performance slightly. Sandbox adoption is low on desktop OSs, and some developers even forsake sandboxed versions of their apps. We investigated how expert desktop users arbitrate different values in apps, and how they cope with feature loss, to understand how they arbitrate between usefulness, productivity and security, and how likely they are to adopt or retain apps that sacrifice features for security improvements. If users are likely to abandon newly sandboxed apps, it would explain developers' reluctance to support sandboxing.

We built a model of values involved in three desktop app appropriation processes: adoption, adaptation, and abandonment. We found that lack of features was the primary reason for users to reject a potential app, and one of two reasons (along with reliability) for users to abandon an app they're using. We also found that users like to adapt and customise their apps, primarily to meet productivity goals, especially for browsers and productivity apps like code editors. Besides, feature loss is a seldom understood phenomena that is poorly accepted by users. A non-negligible portion of our participants would abandon an app that removes a feature they use, especially if justified by security improvements.

Sandbox designers must identify the features threatened by the changes sandboxing brings about, and they must improve support for the relevant APIs so that these features survive sandboxing. They could support plugins by distributing them on app stores and subjecting them to the same security checks as apps. These corrections are essential to avoid putting security in competition with usefulness and security. Indeed, our value analysis clearly shows that security will not be privileged by expert users, and thus, that sandboxed apps are less likely to be adopted than their insecure counterparts.

In future work, we will continue to investigate how app sandboxing and our participants' digital lives fit together. We will assess the fitness of app sandboxing for the information management strategies of our participants using qualitative and quantitative data we collected, and we will investigate how many of the apps they used contain features typically threatened by sandboxing.

References

1. Adams, A., Sasse, M.A.: Users are not the enemy. Commun. ACM **42**(12), 40–46 (1999)
2. Apple Inc.: App Sandboxing, September 2016. https://developer.apple.com/app-sandboxing/
3. Apple Inc.: iOS Security iOS 9.3 or later, May 2016. https://www.apple.com/business/docs/iOS_Security_Guide.pdf
4. Beautement, A., Becker, I., Parkin, S., Krol, K., Sasse, A.: Productive security: a scalable methodology for analysing employee security behaviours. In: SOUPS 2016. USENIX Association (2016)
5. Beautement, A., Sasse, M.A., Wonham, M.: The compliance budget: managing security behaviour in organisations. In: NSPW 2008. ACM (2008)
6. Chacos, B.: And the study says: Windows 8 users rarely touch Metro apps, May 2013. http://www.pcworld.com/article/2039445/
7. Canonical: Ubuntu Core Documentation - Security and Sandboxing (2016). http://docs.ubuntu.com/core/en/guides/intro/security
8. Counsell, D.: Not on the Mac App Store, November 2015. https://www.dancounsell.com/not-on-the-mac-app-store/
9. Docker Inc.: Overview of Docker Hub (2016). https://docs.docker.com/docker-hub/
10. Flatpak: Flatpak - the future of application distribution (2016). http://flatpak.org/
11. Friedman, B.: Value-sensitive design. Interactions **3**(6), 16–23 (1996)

12. Google: Android: application security, September 2016. https://source.android.com/security/overview/app-security.html
13. Hoffman, C.: Why the Mac App Store Doesn't Have the Applications You Want, March 2015. http://www.howtogeek.com/210605/
14. Hoffman, C.: Why Desktop Apps Arent Available in the Windows Store (Yet), March 2016. http://www.howtogeek.com/243559/
15. Paul, I.: The 10 most glaring Windows Store no-shows, April 2013. http://www.pcworld.com/article/2033876/
16. Kirlappos, I., Parkin, S., Sasse, M.: Learning from shadow security: why understanding non-compliance provides the basis for effective security. In: Workshop on Usable Security, USEC 2014, February 2014
17. Mathiasen, N.R., Bødker, S.: Threats or threads: from usable security to secure experience? In: NordiCHI 2008. ACM (2008)
18. McCarthy, J.C., Wright, P.: Technology as Experience. MIT Press, Cambridge (2004)
19. Microsoft: Windows 8 Security Overview, June 2013. https://technet.microsFFoft.com/en-us/library/dn283963(v=ws.11).aspx
20. Dzhumerov, M.: Mac App Store: The Subtle Exodus, October 2014. http://blog.helftone.com/mac-app-store-the-subtle-exodus/
21. Nichols, A.L., Maner, J.K.: The good-subject effect: investigating participant demand characteristics. J. Gen. Psychol. **135**(2), 151–165 (2008)
22. Cohen, P.: The Mac App Store and the trouble with sandboxing, April 2014. http://www.imore.com/mac-app-store-and-trouble-sandboxing
23. Potter, S., Nieh, J.: Apiary: easy-to-use desktop application fault containment on commodity operating systems. In: USENIX ATC 2010 (2010)
24. Sasse, M.A., Brostoff, S., Weirich, D.: Transforming the 'weakest link' a human/computer interaction approach to usable and effective security. BT Technol. J. **19**(3), 122–131 (2001)
25. Schreuders, Z.C., McGill, T., Payne, C.: Empowering end users to confine their own applications: the results of a usability study comparing SELinux, AppArmor, and FBAC-LSM. ACM Trans. Inf. Syst. Secur. 14(2): (2011)
26. Sketch: Leaving the Mac App Store, December 2015. http://bohemiancoding.tumblr.com/post/134322691555/leaving-the-mac-app-store
27. Smetters, D.K., Grinter, R.E.: Moving from the design of usable security technologies to the design of useful secure applications. In: NSPW 2002. ACM (2002)
28. Statista: Most popular Google Play app categories in February 2014, by device installs, February 2014. http://www.statista.com/statistics/279286/
29. Statista: Most popular Apple App Store categories in June 2016, by share of available apps, June 2016. http://www.statista.com/statistics/270291/
30. Strauss, A., Corbin, J.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage Publications Inc., Thousand Oaks (1998)
31. Streeting, S.: Between a rock and a hard place our decision to abandon the Mac App Store, February 2012. http://blogs.atlassian.com/2012/02/between-a-rock-and-a-hard-place-our-decision-to-abandon-the-mac-app-store/
32. Yee, K.P.: Aligning security and usability. IEEE Secur. Priv. **2**(5), 48–55 (2004)