UNITED STATES DISTRICT COURT
WESTERN DISTRICT OF WASHINGTON
AT SEATTLE

FEDERAL TRADE COMMISSION, Plaintiff, v. AMAZON.COM, INC. Defendant.	Case No. 2:14-CV-01038-JCC

EXPERT REPORT OF ANDREW L. SEARS, PH.D.

December 7, 2015

TABLE OF CONTENTS

Page

I. Qualifications

1. I am the Dean of and a Professor in the College of Information Sciences and Technology at The Pennsylvania State University. I have been working in the field of computer science with a focus on human-computer interaction since 1988. I earned a Bachelor's Degree in Computer Science from Rensselaer Polytechnic Institute in 1988 and a Ph.D. in Computer Science from the University of Maryland, College Park, in 1993. The chair of my doctoral dissertation committee at the University of Maryland was Professor Ben Shneiderman. Jennifer King relies upon Professor Shneiderman in her expert report and identified him as one of the "[l]eading researchers" in the field of human-computer interaction. I concur that he is one of the leading researchers in the field of human-computer interaction. Indeed, Professor Shneiderman is responsible for many innovative ideas, methods, and tools widely accepted today, such as direct-manipulation interface design, advances in information visualization, and his well-respected "Eight Golden Rules of Interface Design" first explained in his leading treatise "Designing the User Interface: Strategies for Effective Human-Computer Interaction" published in 1986. Professor Shneiderman also is well known for contributing to the development of the research area of universal usability, a concept pertinent to my expert opinions set forth below.

2. My research has explored many aspects of human-computer interaction, including the use of mobile devices, touchscreen-based interfaces, web-based interactions, and accessibility in the context of computing technologies. My research has been funded by various government agencies, foundations, and corporations, including the National Science Foundation, the National Institute on Disability and Rehabilitation Research, the National Institute of Standards and Technology, the Verizon Foundation, IBM, and Motorola.

3. In addition to my academic research and teaching, I have advised a variety of companies and organizations in the areas of human-computer interaction and user-interface design, including design for mobile devices and touchscreen-based interactions.

4. I served as a founding Editor-in-Chief of the Association for Computing Machinery's (ACM's) journal "Transactions on Accessible Computing" before becoming a member of that journal's editorial board. I also served on the editorial boards of several additional journals,

including ACM's "Transactions on Computer-Human Interaction," the "European Journal of Information Systems," the "International Journal of Human-Computer Studies," and "Universal Access in the Information Society."

5. I served as Conference and Technical Program Co-Chair of the premier conference on human-computer interaction: ACM Conference on Human Factors in Computing Systems (CHI 2001) and as both General Chair and Program Chair for the premier conference on accessibility in the context of computing: ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2004 and ASSETS 2005). I served as the Chair of the ACM Special Interest Group on Accessible Computing and as a member of the ACM Council. I currently serve as a member of the Board of Directors for the Computing Research Association. I was named an ACM Distinguished Scientist in 2010.

6. I have received numerous awards in recognition of my service to both the ACM Special Interest Group for Computer Human Interaction (1998, 1999, 2001) and the ACM Special Interest Group for Accessible Computing (2004, 2005). As a doctoral student, I was selected to receive a fellowship by NASA, which supported my PhD studies, and I was also selected to participate in the Doctoral Consortium at CHI 1992.

7. I was first contacted by Amazon.com attorneys on October 23, 2015, and retained to review and, if appropriate in my opinion, respond to Ms. King's expert report. I am being compensated for my work on behalf of Amazon at the rate of $1000 per hour. My compensation is not contingent upon the outcome of my opinions or of this litigation.

II. Introduction

A. Assignment

8. I was retained in late October 2015 by counsel for Amazon to serve as an expert witness for purposes of consultation and potential expert testimony in the case of *FTC v. Amazon.com, Inc.* Specifically, I was asked to analyze the expert report of Jennifer King and to opine on the following:

a. Whether the methods and practices described and used by Ms. King in her expert report in this case are consistent with accepted and reliable methods and practices in

academic and commercial fields related to human-computer interaction and user-interface design.

b. Whether Amazon's Kindle Fire tablet interfaces, notifications, and descriptions associated with in-app purchases were so far below contemporary, accepted business practices that they would have unfairly confused parents about the presence of and opportunity to make in-app purchases, about the presence of and opportunity to enable Parental Controls to restrict in-app purchases, or the availability of and means to request a refund for any accidental or unwanted in-app purchases.

c. The veracity of Ms. King's opinions, including her conclusions "that Amazon did not effectively convey to consumers downloading an in-app charge app (an app containing in-app charges) from the Amazon Appstore that children could incur in-app charges"; "that Amazon did not effectively convey to consumers downloading an in-app charge app from the Amazon Appstore that they would have to change their device settings to prevent children from incurring in-app charges without parental involvement"; "that Amazon did not effectively convey to consumers who incurred unauthorized in-app charges that refunds were available for those charges from Amazon"; and "that Amazon did not effectively convey to consumers who incurred unauthorized in-app charges how to request a refund for those charges from Amazon."

B. Information Considered

9. My opinions are based on more than twenty-five years of knowledge, skill, experience, training, and education in the field of human-computer interaction and user-interface design. The list of specific materials I considered and relied upon in forming my opinions in this report is available at Appendix B.

C. Summary of Opinions

10. Usability is a subjective concept that evolves with time and experience and must be viewed with great care and consideration and in context. Context includes users, user goals, prior

experience, state-of-the-art, environment, system capabilities, and system goals. As a result, what is considered well designed at one point in time may be considered poorly designed at another point in time. Current understanding of usability may appear obvious in hindsight, but it is often the result of significant research or extensive trial and error unavailable at launch.

11. There is significant risk by imposing liability for negative results during usability trial and error or while exploring new options in an effort to improve upon the status quo. Unless liability is limited to situations in which the design was objectively unreasonable at launch or made in bad faith—and not simply sub-optimal, particularly in hindsight—the risk of liability will deter positive efforts and hamper innovation.

12. Understanding usability requires knowledge of the intended users and the various tasks those users may want to complete. Tasks are often prioritized; compromises in design are often necessary and recommended. Time and experience routinely redefine usability. Changed understanding of user ability and technological capability often create new opportunities for improved human-computer interaction and user-interface design.

13. It is common and accepted practice for organizations to release a product, observe how the product is used, gather feedback from users, and revise the product to address user concerns and improve the user experience. Amazon's in-app-purchase innovations for usability are consistent with this accepted approach.

14. Heuristic usability evaluations or "inspections" such as Ms. King undertook here have significant limitations, particularly where they are performed by evaluators with limited formal training or usability experience. Even when performed by experts, results from heuristic evaluations typically differ from one individual to the next; thus, it is strongly discouraged to rely on a single, subjective evaluation to categorically establish usability or "effectiveness."

15. There is no single accepted set of heuristics, and isolated heuristic evaluations are prone to overemphasizing the severity of perceived problems, including identifying false-positives—items that would not actually create problems for users interacting with a live system. Reliable heuristic evaluations require the use of multiple evaluators to independently review a system and systematically aggregate the results while carefully assessing the existence and severity of potential problems.

16. The methods and practices described and used by Ms. King in her expert report are inconsistent with accepted and reliable methods and practices in academic and commercial fields related to human-computer interaction and user-interface design. Ms. King's goal of identifying major flaws is particularly susceptible to the shortcomings of heuristic evaluation, and her opinions were not informed by more reliable user testing or surveys.

17. Ms. King often draws sweeping, unqualified conclusions about wide swaths of tablet users and interfaces without statistical support. Ms. King erroneously assumes that Amazon's primary task is always to address in-app purchasing and Parental Controls; she fails to recognize the variety of tasks at issue, the various users, the context of the marketplace and the device; and she applies a standard of perfection or best practice, with the aid of hindsight, that is uncalled for in the field and inappropriate where, as here, the question is whether Amazon's practices were unfair, not whether they were imperfect.

18. Ms. King identifies a collection of potential or candidate problems of uncertain reliability and unknown severity, and her solutions, where suggested, do not adequately consider context, show little or no appreciation for competing design objectives, and could introduce unaddressed and unintended consequences.

19. Contrary to Ms. King's opinion, I conclude that Amazon's initial design and refinement process were reasonable and consistent with the practice of the industry. In several ways it was superior to the state-of-the-art, as it provided more detailed information about in-app purchasing, immediate notification of purchases, and Parental Controls to give customers more options with respect to their children's activities. Amazon designed for parents to exercise control over their account-connected devices, and Amazon's launch-and-learn approach and constant refinement is a preferred approach in the industry to improving usability.

20. Amazon's Kindle Fire tablet interfaces, notifications, and descriptions associated with in-app purchases did not fall so far below contemporary, accepted business practices that they would have unfairly confused parents about the presence of and opportunity to make in-app purchases, about the presence of and opportunity to enable Parental Controls to restrict in-app purchases, or the availability of and means to request a refund for any accidental or unwanted in-app purchases.

III. Analysis

A. Usability Must Be Viewed In Context

21. Usability is largely a subjective concept that evolves with time and experience. There is no single definition of usability. What is usable for one group of individuals may be poorly designed for another group of individuals. What is well designed for one set of tasks may be poorly designed for another set of tasks. What is considered well designed at one point in time may be considered poorly designed at another point in time.

22. It is therefore critical that context be considered when evaluating usability. Context includes not only the users of the system and the goals of those users but also the users' prior experience, both with the system and with other systems that may have existed at the time, and the overall environment in which the system will be used.[1]

23. Understanding usability requires knowledge of the intended users and the various tasks those users may want to complete. The more diverse the population of potential users, the more challenging it can be to design solutions that effectively address the goals and abilities of all users. Designing for diverse tasks also introduces challenges. Tasks are often prioritized, based on the current understanding of which tasks users consider most important or which tasks they will complete most often. It has been long accepted that compromises are necessary for well-designed applications, and designs often must necessarily focus on key segments of the target population and frequent or important tasks.[2] As a result, some features that are important for a specific segment of the population may be less obvious or accessible than other features, which were considered more relevant or important for a larger subset of the population.

24. Since the concept of usability evolves with time and experience, a system may be well designed when introduced only to be critiqued or redesigned later because the community's understanding of usability or design has changed. While the new standard for what constitutes "well designed" may appear obvious in hindsight and with evolving goals, it is often the result of significant research or extensive trial and error, which was not available when the system was

[1] See, for example, Hartson, R., & P.S. Pyla, The UX Book: Process and guidelines for ensuring a quality user experience, Elsevier (2012).

[2] Ibid.

first introduced. Many interfaces have been designed with the best of intentions and in such a way that was consistent with or superior to what had been state-of-the-art at the time, only to have aspects of the user interaction redesigned after the product was released because it was determined that a redesigned solution would provide a superior user experience.[3]

1. Time and Experience Redefines What Is Usable

25. Time and experience routinely redefine usability and what both users and designers consider well designed. The evolution of touchscreen-based interfaces and keyboards serves as an example of how time and experience affects what is considered usable. In the mid- and late 1980s, it was widely believed that touchscreens could only be used to select relatively large objects. At the same time, the state-of-the-art was to design interfaces such that selections were made based on where the user's finger initially contacted the screen.[4] That was the norm, and systems were designed leveraging this understanding.

26. While the original strategy, known as "land on," allowed for selection of large objects, researchers developed new interaction techniques, including what became known as the "lift-off" strategy. The lift-off strategy allows users to reposition their finger before lifting it from the screen.[5] The location where the finger is lifted from the screen is activated rather than the location where the finger first touches the screen. Even with the lift-off strategy, however, the size of the objects an individual could select remained limited. In the late 1980s, researchers demonstrated that carefully designed touchscreen-based interactions could allow users to select smaller targets than previously believed.[6]

27. This changed understanding of how user ability (for finger location) and technological capability (for smaller-area selections) combine created new opportunities for improved human-computer interaction and user-interface design. And although those practices are now accepted

[3] Ibid.

[4] Potter, Richard L., Linda J. Weldon, and Ben Shneiderman. "Improving the accuracy of touch screens: an experimental evaluation of three strategies." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 1988.

[5] Ibid.

[6] Sears, Andrew, and Ben Shneiderman. "High precision touchscreens: design strategies and comparisons with a mouse." International Journal of Man-Machine Studies 34.4 (1991): 593-613.

standards in the touchscreen market, no reasonable human-computer-interaction expert would conclude that the early touchscreen designs were ineffective or were so far below any accepted standard to be considered unusable or unfair at the time they were introduced.

28. The evolution of touchscreen-based keyboards may be even more dramatic.[7] Land-on selection has been replaced by lift-off selection. Larger keys have replaced by smaller keys. Even small lift-off-based touchscreen keyboards have evolved significantly. For example, gesture-based interactions have been integrated, allowing users to drag their finger from one letter to the next, entering a complete word in a single action, and predictive technologies allow users to enter complete words without having to enter each individual letter.[8] Dictionary-based techniques, combined with knowledge of common errors, allow touchscreen keyboards to automatically correct some errors.

29. Some improvements are driven by changes in the underlying technologies, while others are a result of trial and error, a desire to improve upon the status quo, and exploring new options. There is significant risk by imposing liability for negative results during such trial and error or while exploring new options in an effort to improve upon the status quo. Unless liability is limited to situations in which the design was objectively unreasonable at launch or made in bad faith—and not simply sub-optimal, particularly in hindsight—that risk of liability will deter such positive efforts and hamper innovation.

30. A touchscreen keyboard that only allowed individual letters to be entered with each touch—using the land-on or lift-off strategy—would have been state-of-the-art in the late 1980s. In many contexts, this same keyboard would likely be considered poorly designed by today's standards, particularly given users' exposure to and current understanding of auto-correct algorithms, predictive algorithms, and gesture-based interactions in the context of text entry. The time at which an interface was developed must be considered when analyzing whether it was a reasonable solution when it was introduced.

---

[7] Sears, Andrew, et al. "Investigating touchscreen typing: the effect of keyboard size on typing speed." Behaviour & Information Technology 12.1 (1993): 17-22.

[8] See, for example, Zhai, Shumin, and Per Ola Kristensson. "The word-gesture keyboard: reimagining keyboard interaction." Communications of the ACM 55.9 (2012): 91-101.

31. As a result, analyzing the usability of a system that existed at some point in the past, with the goal of understanding how well a system was designed, requires great care and consideration of the context in which the system was produced. Importantly, this requires consideration of what was known and what was considered state-of-the-art at the time. A system may have been well designed when introduced, based on what was known at the time, only to be less than ideal at the present. Without that understanding and careful evaluation, there otherwise remains significant risk that reliance on hindsight will produce merely differences of opinion on optimization, not objective assessment of usability or reasonableness.[9]

2. State of the Art for In-App Purchases in November 2011

32. The evolution of how users were informed about in-app purchases is another example. When Amazon introduced in-app purchases in November 2011, both Apple and Google had already been providing apps that offered in-app purchases. The user experience provided by Apple and Google, who were the largest providers of apps that include in-app purchases, was state-of-the-art at the time. To be sure, Apple and Google are skilled, respected consumer companies with particularly strong reputations for user-interface design and ease of use. Apple began offering apps with in-app purchasing in October 2009.[10] Yet according to the FTC's complaint against Apple, even by November 2011 Apple was not providing any up-front disclosure regarding the existence of in-app purchases.[11] Nor was Apple providing any immediate notification to the device owner after an in-app purchase was completed. Similarly,

[9] Ms. King often ignores the importance of context and time. For example, she identifies one study finding that confusion between apps being advertised as free but, in fact, not free to use has led to consumer disappointment and lower consumer ratings. Expert Report of Jennifer King at 20. Yet the results of that study were reported in May of 2015, nearly four years after Amazon launched in-app purchases. Amazon could not have designed its products in 2011 based on the insights provided by an article published four years later.

[10] Expert Report of Ravi Dhar, October, 16, 2015.

[11] *In re Apple Inc.*, Dkt. No. C-4444, Complaint, Federal Trade Commission (Mar. 2014) (FTC_AMZ_00000001). Contrary to the representation in the FTC's complaint against Apple, it may be that by November 2011 Apple included on its app-description page a limited note about in-app purchases. (I understand from Amazon's counsel that the FTC refused to provide information to Amazon in this case about Apple's or Google's in-app purchase disclosures and interfaces.) Even if so, Amazon still met and exceeded the state-of-the-art because it too included a description, its description was detailed, and it specifically identified Parental Controls.

even by November 2011, Google was not providing any up-front disclosure regarding the existence of in-app purchases.[12]

33. When Amazon launched in-app-purchasing opportunities in November 2011, Amazon went beyond then-accepted state-of-the-art. Amazon included a description of the availability of in-app purchases on an app's description page. That note began with large, all capitalized letters stating "PLEASE NOTE." It described in-app purchasing, including that such purchases were paid for using "actual money." And it explained that users could configure Parental Controls.[13] That these details had been excluded by the other market participants supports the conclusion that Amazon was operating above contemporary, accepted business practices with regard to making this information available to all users including parents.

34. Moreover, unlike Apple, in November 2011 (and through today), Amazon *immediately* emailed an order-confirmation receipt to the account holder after each in-app-purchase. That email receipt included detailed information about each purchase, including the name and cost of the product and the order number. It also included links to the user's account page as well as to Amazon help pages.[14] That Amazon included those details and links in an immediate confirmatory email further supports the conclusion that Amazon's in-app purchasing practices exceeded those of its largest competitors and then-standard practices.

3. Design, Release, and Redesign Is An Accepted Practice

35. Usability is not a formula-driven concept. It is based on the experiences of the users. As a result, it is common for organizations to release a product, observe how the product is used, gather feedback from users, and revise the product to address user concerns and improve the user experience. It is common for this to be an interactive process involving multiple releases as the system is improved.[15] This is consistent with the approach Amazon employed here. Design,

[12] *In re Google Inc.*, Dkt. No. C-4499, Complaint, Federal Trade Commission (Dec. 2014) (FTC_AMZ_00000025).

[13] Expert Report of Jennifer King (October 16, 2015) at 25, Fig. 4; Expert Report of Ravi Dhar (October 16, 2015) ¶¶ 54-55.

[14] Expert Report of Jennifer King (October 16, 2015) at 43, Fig. 12; Expert Report of Ravi Dhar (October 16, 2015) ¶ 36 & Appendix A.

[15] Hartson, R., & P.S. Pyla, The UX Book: Process and guidelines for ensuring a quality user experience, Elsevier (2012).

release, and redesign is not only supported by the usability community and literature but also has distinct advantages in terms of obtaining the most pertinent feedback, ensuring earlier entry of competitive alternatives, and furthering innovative approaches.

36. As already noted, Amazon's initial release of in-app purchasing in 2011 went beyond what was state-of-the-art by providing both an up-front notification about in-app purchases and immediate email updates whenever in-app purchases were made. This was followed by multiple revisions to how in-app purchases were introduced and how in-app purchases were completed. Passwords were introduced for certain in-app purchases in 2012 in response to insights gained based on customer behaviors. The Key Details feature was introduced in 2013, providing users with an additional way to learn about in-app purchases. Which purchases required a password continued to evolve based on user experiences, with a new requirement in May 2013 that a password be entered prior to the first in-app purchase regardless of the cost. This same dialog box informed users that they could require a password for all future in-app purchases by turning on Parental Controls. Building on knowledge gained from user experiences, the current interface was introduced in June 2014. This interface requires users to explicitly choose between requiring and not requiring a password for future purchases.

37. An iterative approach involving product launch, user review and feedback, and refinement—such as that used by Amazon—is a preferred approach that allows for continuous improvement of the user experience. Usability and innovation in design will suffer if companies avoid addressing known interface issues or refrain from exploring ways to improve user design because they fear that making such improvements will invite litigation based on the (incorrect) assumption by others that the changes indicate their original designs were unreasonable.

B. Heuristic Evaluations Are Less Reliable and Prone to Error if Applied Incorrectly

38. Even at a fixed point in time, usability is a subjective concept. What is ideal for one individual may be less than ideal for another. What is usable for one individual may be less than clear to another individual. There is no "formula" based approach that can be applied to ensure that a system is usable. As a result, various methodologies have been developed to evaluate

usability in the context of information technologies. When used properly, these methodologies can help identify and address usability problems.

39. Many methodologies are available for evaluating usability. The website Usability.gov, as Ms. King mentions in her report, is one useful resource that summarizes a number of methodologies. However, Usability.gov does not provide a comprehensive understanding of the issues involved and should not be relied on in isolation.

40. Some usability evaluation methodologies are based on direct input from users, often based on interactions with some version of the system in question. These are important techniques because, as highlighted by usability expert Dana Chisnell, on whom Ms. King also relies, there is no substitute for observing users interacting with a system if the goal is to understand the user experience.[16] Other methodologies rely on input from usability experts, who provide insights based on their own personal review of or interaction with some version of the system. Due to the subjective nature of usability, an effective evaluation—whether user- or expert-based—relies on input from multiple individuals rather than on the input of just a single individual.

41. Ms. King's identified method for evaluating the "effectiveness" of Amazon's in-app-purchasing process is not a usability test but a "usability inspection." But a usability inspection is a collection of techniques, not a single methodology.[17] Later, Ms. King clarifies that she conducted a "heuristic evaluation." Stated differently, she "review[ed] an application interface for compliance with an accepted set of heuristics."[18]

42. The history of heuristic evaluation within the human computer interaction community is typically traced back to the early 1990s and research reported by Jakob Nielsen. As detailed below, however, extensive subsequent research now reveals that there are important limitations with a heuristic evaluation. Heuristic evaluations are inherently subjective, and they are even less reliable when the evaluator has limited formal training or experience in usability concepts.

---

[16] Chisnell, Dana. "What you really get from a heuristic evaluation." UX Magazine, Feb. 19, 2010. http://uxmag.com/articles/what-you-really-get-from-a-heuristic-evaluation

[17] See, for example, Mack, R. L., & Nielsen, J. (Eds.). (1994). Usability inspection methods. New York, NY: Wiley & Sons.

[18] Expert Report of Jennifer King (October 16, 2015) at 13 & n.3.

Indeed, even experienced heuristic evaluators typically produce different results when evaluating the same subject. There is no single "accepted set of heuristics." Isolated heuristic evaluations are also prone to overemphasize the severity of perceived problems and identify false-positives—items that would not actually create problems for users interacting with a live system. The more reliable, accepted method for heuristic evaluation requires the use of multiple usability experts to independently review the system, followed by careful analysis involving multiple individuals that assesses the severity of the potential problems. This severity analysis reduces false-positives and better assesses the severity of potential problems. Ms. King's much more informal process does not account for the synthesis of information that would reduce the presence of false-positives and better assess the severity of potentially problems, and therefore her conclusions are subject to those flaws.

1. Heuristic Evaluations Are Subjective

43. Critically, studies confirm that usability is a subjective concept and different individuals applying this same technique will often arrive at different conclusions. Expertise (or lack thereof) affects those conclusions, and heuristic evaluation is more effective when the individuals performing the evaluation have formal training or usability experience. Importantly, even when the individuals are usability experts, results typically differ from one individual to the next. Because variation in usability opinions is inevitable, it is strongly discouraged to rely on a single, subjective evaluation to categorically establish usability or "effectiveness."[19]

2. Reliable Application of Heuristic Evaluation Requires the Use of Multiple Evaluators

44. Also critical to an effective heuristic evaluation is the use of multiple evaluators to independently assess and then evaluate others' assessments. The need for multiple evaluators has been a well-documented guiding principle for at least 20 years.[20] Even Usability.gov, on which Ms. King relies, emphasizes this requirement.[21] Jakob Nielsen, also on whom Ms. King relies

[19] http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation; Nielsen, Jakob. "Finding usability problems through heuristic evaluation." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92), ACM, New York, NY, USA, 373-380.

[20] Ibid.

[21] http://www.usability.gov/how-to-and-tools/methods/heuristic-evaluation.html.

and who is often credited as one of the scholars who introduced this approach to the community, similarly cautions about the importance of using multiple evaluators.[22] To use techniques like heuristic evaluation properly, several usability experts should independently review the system. This initial review should be followed by a systematic approach to aggregate evaluations, synthesize the inputs gathered, and assess the existence and severity of potential problems.

3. Heuristic Evaluations Can Overestimate the Severity of Potential Problems and Identify False-Positives

45. Because heuristic evaluations, which are designed to uncover potential problems, are subjective comparisons of an interface to a set of guidelines or heuristics, such an evaluation often identifies problems that do not actually exist in practice. Accordingly, any item identified by an individual evaluator should be considered a potential or candidate problem.[23] The potential usability problems identified via inspection-based techniques including heuristic evaluation will vary in importance. Some of the items identified will correspond to minor usability problems, causing little more than an occasional annoyance. Other items will correspond to severe usability problems, which may interfere with an individual's ability to complete a task. Importantly, items identified using heuristic evaluation have also been shown to be false-positives. In other words, some of potential problems identified by individual evaluators, including evaluators who are usability experts, will not represent real problems that need to be addressed.[24]

46. A systematic process of reviewing the existence and severity of the potential problems, which also needs to include several individuals, is vital to address and eliminate false-positives while also distinguishing between issues that may produce little more than an occasional irritation and those that may cause severe problems for users.[25] Unlike usability tests, which rely on users interacting with the system and encountering problems, heuristic evaluation relies on

[22] http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation.

[23] Cockton, Gilbert, and Alan Woolrych. "Understanding inspection methods: lessons from an assessment of heuristic evaluation." People and Computers XV—Interaction without Frontiers. Springer London, 2001. 171-191.

[24] Sears, Andrew. "Heuristic walkthroughs: Finding the problems without the noise." International Journal of Human-Computer Interaction 9.3 (1997): 213-234.

[25] http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/.

individuals reviewing a system and identifying things they *think* may be problems. The nature of heuristic evaluation makes the problem severity rating process even more important.

47. The concern of false-positives has been in the forefront since heuristic evaluation was introduced. One of the original papers introducing the heuristic technique highlighted the potential for false-positives.[26] Other researchers subsequently confirmed the problem of false-positives and the importance of a separate process for reviewing problem severity in the context of heuristic evaluation.[27] Researchers have also demonstrated that heuristic evaluation can produce numerous potential problems that are not identified or corroborated through user testing.[28] Finally, Jakob Nielsen continues to highlight the critical importance of an effective severity rating process, stressing the importance of having several people independently review the severity of each potential problem.[29]

4. Heuristic Evaluations Are Not a Substitute for Usability Testing

48. Ms. King states that the "overriding goal" of her heuristic evaluation is to "identify major flaws."[30] She also says "a usability inspection [such as she performs] can provide similar insights to those generated through user testing, particularly when reviewing interfaces for conformance with basic principles."[31] Both of these statements raise serious concerns.

49. First, heuristic evaluations are *not* known or designed to focus on "major flaws." As described above, heuristic evaluations identify a mixture of potentially severe problems, minor problems, and false-positives. Most often, the number of minor problems identified is far greater than the number of severe problems identified.

---

[26] Nielsen, Jakob, and Rolf Molich. "Heuristic evaluation of user interfaces." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 1990.

[27] Sears, Andrew. "Heuristic walkthroughs: Finding the problems without the noise." International Journal of Human-Computer Interaction 9.3 (1997): 213-234.

[28] Law, Effie Lai-Chong, and Ebba Thora Hvannberg. "Analysis of strategies for improving and estimating the effectiveness of heuristic evaluation." Proceedings of the third Nordic conference on Human-computer interaction. ACM, 2004; http://www.nngroup.com/articles/usability-problems-found-by-heuristic-evaluation/.

[29] http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/.

[30] Expert Report of Jennifer King (October 16, 2015) at 14.

[31] Expert Report of Jennifer King (October 16, 2015) at 14.

50. Second, the results produced by heuristic evaluations are not as reliable as those produced through user testing and do not necessarily produce similar insights.[32] In fact, the reference Ms. King uses to support the assertion about the similarity of those results supports the *opposite* conclusion. Ms. King relies on Dana Chisnell's 2010 article titled "What you really get from a heuristic evaluation" for her similarity position,[33] but that article actually concludes otherwise.[34]

51. Chisnell makes several statements that are in direct conflict with Ms. King's interpretation of Chisnell's work. Chisnell states, for example, that "[u]nfortunately, the request [by a client to do a heuristic evaluation of a product] usually suggests that a heuristic evaluation can substitute for usability tests." But Chisnell confirms that heuristic evaluation is "an inspection, *not* an evaluation. It is *not* about the user experience."[35] Chisnell further verifies that any such heuristic review that "claims to answer" questions about a user's actual experience "is just guessing." Chisnell also confirms that heuristic evaluations such as Ms. King's are likely to identify design problems that are not actually problematic: "Worse, they may identify things that don't comply with the heuristics that should *not* be changed."[36]

52. Chisnell explains why Ms. King's individual, subjective heuristic evaluation inspection is unreliable in evaluating Amazon's in-app purchasing process: "Heuristic evaluation may help a team know whether their UI [user interface] complies with someone else's guidelines. But observing people using a design in a usability test gives a team primary data for making design decisions for their users using their design . . . ." Notably, Chisnell identifies touchscreen devices and online connectivity (the interfaces at issue here) as particularly susceptible to error from a heuristic evaluation: "[user testing is superior,] especially in a world evolved far beyond command line entry and simple GUIs [graphic user interfaces] to options like touchscreens, social media, and ubiquitous connectivity." As Chisnell concludes, "For me, observing people using a design will always trump an inspection or audit for getting solid evidence to determine a

[32] http://www.nngroup.com/articles/usability-problems-found-by-heuristic-evaluation/.

[33] Expert Report of Jennifer King (October 16, 2015) at 14.

[34] Chisnell, Dana. "What you really get from a heuristic evaluation." UX Magazine, Feb. 19, 2010. http://uxmag.com/articles/what-you-really-get-from-a-heuristic-evaluation

[35] Emphasis added.

[36] Emphasis added.

design direction." Heuristic evaluations can be useful, when applied properly, for determining if a system conforms to guidelines. But as Chisnell notes, the fact that something does not conform to a predefined set of guidelines does not mean that there is a problem that needs to be addressed.

C. Ms. King's Heuristic Evaluation Is Incomplete and Unreliable

53. Ms. King's report contains questionable conclusions drawn from an incomplete and unreliable heuristic evaluation. By conducting an individual, subjective evaluation in a vacuum with an "overriding goal" to "identify major flaws," Ms. King uses an approach that is particularly susceptible to overemphasizing the severity of perceived problems and identifying false-positives. Ms. King's conclusions were not informed or corroborated by user testing or user surveys, and she does not consider the fact that the overwhelming majority of customers used in-app purchasing without complaint.[37] Yet, Ms. King often draws sweeping, unqualified conclusions about wide swaths of tablet users without statistical support—opining that "many" customers would not understand in-app purchasing, that "many, if not most" customers would never see Amazon's in-app-purchasing notices, that "many users were not likely to know that a free app would have any additional costs attached" (even though no additional costs were ever necessary to use the free apps at issue here), and that "in the majority of cases" Amazon's badging alone must effectively convey the presence of in-app purchasing. Ms. King presents her own personal, broad opinions on these topics while discounting the judgments of Amazon's team of web and tablet designers, who have considerable experience and real-world success improving user interfaces and user experiences.

54. Moreover, Ms. King's analysis seems to erroneously assume that Amazon's primary task is always to address in-app purchasing and Parental Controls. She often fails to recognize the variety of tasks at issue on each screen and how the various groups of potential users would perceive these tasks, particularly for a multimedia device designed to provide a multitude of services to a variety of customers. She fails to recognize that most Amazon customers prefer a frictionless purchasing experience unencumbered by additional dialog boxes and password prompts.[38] She often ignores her own opinion that adding more text likely means that users will

[37] Expert Report of Ravi Dhar (October 16, 2015) ¶¶ 103-07 & Table 2.

[38] Expert Report of Donna Hoffman (October 16, 2015) ¶¶ 55-64, 73-76.

skim the text rather than read it. And, Ms. King's analysis focuses almost exclusively on the most uninformed user who is unwilling to make any effort to understand the purpose and ability of the purchased tablet device. Indeed, she goes so far as to conclude that the term "real money" only "subtly" conveys the use of actual currency and that the term "future in-app purchases" is "vague."

55. Ms. King similarly ignores the fact that Kindle Fire tablets are media-consumption devices, which necessarily implies some purchasing ability via that device—an understanding that is reinforced by the fact that device owners must first link the tablet to an Amazon credit-card account before making any purchase on the device. In other words, Ms. King assumes that the average Kindle Fire account owner is unaware that his or her device, once enabled for credit-card charges, could be used by another user for credit-card purchases. That is not a reasonable assumption because ownership of such an account-connected device should and likely would prompt at least of a modicum of vigilance by tablet owners, particularly parents. In my opinion, it was reasonable for Amazon to proceed on the understanding that users, particularly parents, would not simply hand an Internet-connected and credit card-enabled device to children without any supervision over purchases. It was and remains a reasonable design choice to build into such devices parental choice and regulation, which is what Amazon provided via its notices and Parental Controls (and later via FreeTime).[39] Ms. King ignores this context and reaches a conclusion that is based on the far less reasonable assumption that Kindle Fire account owners would exercise almost zero supervision or responsibility over how others, particularly children, used their devices once enabled for purchases over the Internet.

56. Nor does Ms. King clearly identify to what standard she holds Amazon. She appears to demand a level of perfection or "best" practice from inception of the product, with the aid of hindsight, that is uncalled for in the industry or in the academic literature. No single evaluator can effectively and reliably determine a best practice, particularly in a developing market with a developing interface. The field of human-computer interaction does not recognize a single standard for best practices or even for complete "effectiveness." Ms. King's approach is

[39] Business planning documents for the Amazon Appstore indicate that Amazon early on identified Parental Controls as a tool available for parents who would supervise their child's use of a device. Amazon_265968-70, Amazon_265973, Amazon_265968, Amz_FTC_0037659.

inappropriate here where the question is whether Amazon's practices were unfair, not whether they were imperfect. Indeed, Ms. King's unqualified conclusions provide an inaccurate and inappropriate description of the tablet market, and her analysis fails to consider the state-of-the-art or industry practice when Amazon launched opportunities for in-app purchasing and introduced various interfaces over time. She similarly disregards whether Amazon's approach was reasonable even if not "best," whether it was part of a refinement process that employed data and information from actual user experience and feedback with the goal of continuously improving the user experience, or whether it was made in good faith and not intended to deceive.

57. As a result, Ms. King produces a collection of potential or candidate problems of uncertain reliability and unknown severity, and her solutions, where suggested, do not adequately consider context, show little or no appreciation for competing design objectives, and could introduce unaddressed and unintended consequences. Some of these candidate problems are likely false-positives. Even those items that are not false-positives simply correspond to items that deviate from a set of guidelines, and those "deviations" do not necessarily mean that changes to the system are appropriate or that users would experience difficulty as a result of the issues identified. The following examples highlight why multiple, professional evaluators are necessary when applying such usability inspection techniques.

1. Prominence of In-App Purchase Disclosures

58. Beginning less than a week after Amazon first started offering opportunities for in-app purchasing, Amazon included on the app-description page information about in-app purchases allowing users to understand the meaning of in-app purchasing and to understand that Parental Controls were available for use in the context of in-app purchases. This information was appended to the end of the app description, highlighted by the uppercase text "PLEASE NOTE." The note itself informed users that the "app contains in-app purchasing, which allows you to buy items within the app using actual money." That text clearly and in understandable terms states that purchases can be made from within the app and that these purchases are made using real or "actual" money. The note continues by informing users that they "can configure parental controls from the device Settings menu by selecting Parental Controls." That text also clearly and in

understandable terms informs users that they have the ability to enable Parental Controls via the settings menu.

59. In my opinion, and contrary to Ms. King's view, in the context in which it was encountered, the phrase "in-app purchase" was sufficiently clear. Moreover, because the device is connected to the account holder's Amazon account, which includes a method of paying for purchases such as a credit card, it is reasonable to expect users to understand that these "in-app purchases" would result in charges to their Amazon account.

60. The use of uppercase to highlight "PLEASE NOTE" makes the in-app purchase note more prominent than it would be otherwise. The fact that the app developer could also use uppercase within the app description does not change this fact or completely eliminate its usefulness, as Ms. King suggests. There can be multiple important pieces of information within the app description including the in-app purchase note and other items identified by the app developer. The use of uppercase would help to attract the users' attention to these items that were considered important.

61. Nor did Ms. King identify any instance in which uppercase was used so often that it was no longer a distinguishing feature. And had Amazon made the entirety of the in-app-purchase note uppercase (as opposed to the attention-grabbing first two words), it may have been more difficult for users to read and they may have then, as Ms. King notes, only skimmed or completely ignored the text.

62. Ms. King opines that "in the majority of cases," users would not see or understand the in-app-purchase description, yet she performed no user testing or statistical measurement to draw such a conclusion. And there is simply no support for Ms. King's unqualified position that because other text may have also been displayed in uppercase, the in-app-purchase text was "difficult, if not *impossible*, to notice" and was categorically "ineffective."[40]

63. Finally, Ms. King ignores the simple fact that not all notices and information can be prominent. Not everything can be distinct, else everything becomes indistinct. Tradeoffs regarding importance and placement must be made. Everything cannot always appear "above the

[40] Expert Report of Jennifer King (October 16, 2015) at 26 (emphasis added).

fold," nor should it. It was reasonable and fair and consistent with industry norms for Amazon to conclude that the most important information on the app-description page, and that which appeared at the top of the screen, describe the purpose and use of the app itself. In-app purchasing is optional and supplemental; its prominence should be commensurate with that significance. And Amazon reasonably adjusted that prominence as the significance of in-app purchasing also evolved in the market.

2. Prominence and Clarity of Key Details Badge

64. The Key Details badging was added in June 2013 as an additional means of highlighting important or useful information. As Ms. King notes, the Key Details were "Above the fold," making them even more visible than other text when the app description is initially displayed. The bulleted list makes it clear that this is a list of topics and not the complete details. The Key Details label is presented in a color and font similar to other text on the screen, and the words effectively convey that this section includes important information.

65. Ms. King asserts that the title suggests that the section contains information the user "may wish to know" but that it "does not communicate to the user who is concerned about the cost of the app that it contains information that he or she must know regarding potential charges prior to downloading or using the app." I disagree. The title "Key Details" is a useful way to highlight important details, particularly as users would understand that the term "Key" is synonymous with "important." The purpose of the title was not, as Ms. King suggests, to specifically "communicate to the user who is concerned about the cost of the app that it contains information that he or she must know regarding potential charges prior to downloading or using the app." The purpose of the title was to highlight the availability of important information or "key details."

66. I also disagree with Ms. King's analysis of the text provided for the In-App Purchases key detail. This text clearly and in understandable terms conveys that "actual money" is used, that purchases can be made from within apps, and that you can configure these purchases from within Parental Controls. I disagree with her assertion that the notice does not convey the "fact that IAPs have real costs associated with them."

67. Ms. King states that users must already be familiar with the term "in-app purchasing" for "the element to be immediately effective." I disagree. The choice of words makes it such that the phrase clearly refers to a "purchase" in the context of an "app" or gives the reasonable parent enough information to inquire further about the type of purchasing at issue.

68. Ms. King does not offer alternatives or consider the unintended consequences of implementing those alternatives. With respect to the Key Details badging and implicit throughout Ms. King's report is the contention that Amazon should have added more text and more notices to describe in-app purchasing and Parental Controls in greater detail. But even she admits that "people prefer to skim, rather than read lengthy amounts of text" and that additional "task interruption [that diverts attention from the primary task, e.g., selecting and using the app] is a powerful disincentive."[41] Ms. King makes no effort to reconcile those inconsistent positions or address the balance required to most efficiently give all users, including parents and nonparents, the information and tools to make informed decisions about their devices and the individuals they permit to use those devices.

3. Clarity of In-App Purchase Disclosures

69. Ms. King asserts that it is important to disclose "in-app purchases in a brief and concise statement." She also asserts "given the fact that in-app purchases can quickly become quite costly, the importance of communicating their cost is crucial." However, her two assertions can be in conflict since it may be difficult to "communicat[e] their cost" in a "brief and concise statement" due to the potential for multiple, different in-app purchases with different costs within a single app. And she often criticizes Amazon for using brevity and conciseness in its notices. Given the trade-off, it is important that the message be brief; otherwise, as Ms. King has noted, people may be less likely to read to notice.

70. Ms. King argues that the use of the phrase "parental controls" was "likely to introduce confusion or an additional learning barrier for novice parents or users who are not parents attempting to familiarize themselves with the system." She draws her conclusion from her assertion that the term "parental controls" was originally introduced in the context of content

[41] Expert Report of Jennifer King (October 16, 2015) at 24, 38.

blocking and "purchase controls did not emerge until the late 2000s." Thus, according to Ms. King, for some unidentified portion of the user population "pre-existing familiarity with the term may trigger primary associations with content restriction rather than purchase restrictions." Ms. King's evaluation of the understanding of the term "parental controls" (particularly as late as November 2011) is unnecessarily and incorrectly narrow. Reasonable users would understand the term "parental controls" to refer to *restrictions* set up by parents—a restriction to content or a restriction to purchasing, or both. Much like the definition of the term "in-app purchasing" is self-evident, so too is the term "parental controls," particularly in this context.

71. Ms. King appears to ignore common sense while focusing exclusively on only the most uninformed users who are unwilling to make any effort to understand the purpose and capabilities of the device they purchased. There is of course the possibility that "some" users will not understand the term "parental controls," just as there is the possibility that some users will not understand any term or notice. Any interface design and any choice of words can result in "some proportion of the user population" misunderstanding some details; that possibility cannot be avoided, especially where a company must balance the needs and desires of many varied customers. Ms. King demands a level of perfection that is both unjustified and nonexistent. That *some* users *may* not have understood that commonly used and widely understood term does not make its use "ineffective."[42]

4. Usability of In-App Purchase Flow and Disclosures

72. In May 2013, even though only a small percentage of users were experiencing any misunderstanding about in-app purchases, Amazon introduced a password prompt on each device for all first-time in-app purchases. That dialog box not only required entry of the account holder's password to complete the purchase but also further explained in-app purchases; reiterated that if the account holder wanted to require a password for future purchases, he or she

[42] Ms. King also ignores that Amazon and its Appstore provide unfiltered customer reviews on every app-detail page. Those reviews, which are a widely recognized source of useful information, include commentary (both positive and negative) about the availability of in-app purchasing and Parental Controls. *See* Expert Report of Michael Callahan (October 16, 2015) at 10. By making available not just many reviews but all reviews (critical and otherwise), Amazon's design choice helps train customers to use those reviews as a source of information. Ms. King fails to consider that context.

should set up Parental Controls; and provided a direct hyperlink to enable those parental-control settings.

73. Despite this effort by Amazon to again convey both the opportunity for in-app purchases and the ability to disable them (creating additional friction unwanted by most of its customers in order to help those customers who may have needed additional assistance), Ms. King identifies "at least three" issues with the dialog box. She first observes that "the call to action (entering a password) dominates the focus of this prompt and suggests the other text is less important." Ms. King ignores the fact that the password prompt immediately above the space for entering the password says, "To complete your *purchase*, enter your Amazon password." The wording of the prompt conveys to the account holder that someone is making a purchase on the device, and this is the primary purpose of the dialog box. Moreover, the title of the dialog box, which is presented in larger font size than any other text, also says, "Confirm In-App *Purchase*," prominently verifying for the account holder that password entry will result in a purchase. These are reasonably the most important aspects of the dialog box, and they are made sufficiently clear to the user.

74. Second, Ms. King suggests that users would not understand the phrase "real money" because it appears "without mentioning the dollar amount of the particular charge or using a dollar signs or other signals to emphasize that there is a financial transaction." I strongly disagree with Ms. King's analysis. The lack of dollar signs or a specific dollar amount on this dialog box would not cause the reasonable user (or even an uninformed user) to misunderstand that he or she would be spending money by completing the transaction.

75. Finally, Ms. King opines that the phrase "future in-app purchases" is vague such that "even users familiar with in-app purchases are unlikely to understand that after entering their password a child would be able to incur additional in-app charges without password reentry." I disagree. She asserts that users who "read the full sentence" may "fail to understand that they must change their device settings to prevent children from incurring additional in-app charges without password entry" because the "If you'd like" at the beginning of the sentence characterizes the information as optional. I disagree. I believe this sentence presents the user with clear instructions in the form widely accepted formula: if GOAL then ACTION. If the user

would like to require passwords for future in-app purchases, they must "turn on Parental Controls." The standard by which Ms. King evaluates this dialog box, as well as other prompts, is unreasonably demanding, unsupported by academic literature, and implies that no phrase would ever be sufficient.[43]

5. Prominence and Clarity of Purchase Confirmation Screen

76. Ms. King critiques the "Purchase Confirmation Screen," asserting that "When Amazon required password entry, some parents may not have seen this screen after entering their password and handing the device back to a child." Her critique misses the point. By presenting a confirmation dialog box, Amazon provided the opportunity for users to confirm their transaction. It is always possible that someone may not notice such a confirmation if they quickly hand the device to someone else, but this does not negate the fact that Amazon provided users with an opportunity to confirm their transaction.

77. Ms. King also asserts that "Other parents may have seen the screen, but may not have read the text [about Parental Controls] given that the primary call to action is the close button in the top right." Again, this is always a possibility when a screen is designed to address multiple goals. In this case, the primary goal is to confirm a transaction. A secondary goal was to convey the status of the Parental Controls, providing users with yet another opportunity to verify and modify the status of these controls. Ms. King's analysis seems to assume that the primary task is always to deal with the Parental Controls related issues, failing to recognize the variety of tasks supported by this dialog box and which tasks were considered most important for the majority of users.

6. Ms. King's Analysis of Customer-Service Contacts Is Flawed

78. Ms. King's analysis of customer-service contacts is likewise flawed and suffers from the same defect of searching for and highlighting information that supports a theory rather than evaluating usability objectively. In particular, Ms. King began with over 152,000 "records of

[43] Notably, Ms. King highlights Nielsen's position that it is best practice to disclose additional fees as soon as possible. Expert Report of Jennifer King (October 16, 2015) at 20-21. But she ignores that Amazon in multiple locations disclosed—before even downloading an app and long before making any in-app purchase—that opportunities for in-app purchasing were available and optional. Amazon also disclosed every specific fee before the user made a purchase.

individual consumer complaints corresponding to certain Amazon complaint codes." But this entire set of records was *predisposed* to represent situations where customers raised concerns, thus deliberately reinforcing the existence of her allegedly uncovered "themes." Importantly, Ms. King provides no information or analysis regarding the total number of transactions during this period of time, limiting the ability to understand the prevalence of concerns based on this source of information alone.

79. Moreover, Ms. King pruned those records to approximately 28,000 records of chat sessions and emails that included the phrase "Accidental Order – Child." By limiting the sample to only those contacts with the term "Accidental Order – Child," Ms. King necessarily restricted her analysis to records predisposed to support her "themes" about accidental orders made by children—to the exclusion of contacts made by all Amazon customers about in-app purchases let alone the total number of in-app purchases that occurred during this period.

80. Taking a random sample of that focused subset of records, Ms. King manually reviewed 400 records, identifying a smaller subset of 219 records she deemed "informative" because the customer *explicitly noted* that the basis of the complaint was an unintentional order for an in-app purchase by a child."[44] It should come as no surprise that Ms. King discovered "themes" about accidental child purchases when her analysis *began* from a highly filtered set records consisting of just over 0.1% of the total customer contacts.

81. Further, it is unreliable to draw meaningful conclusions based on approximately 20 self-selected quotes at the end of this process. Moreover, she presented the quotes without context. Ms. King provides no analysis of the outcomes for these interactions or the prior online and tablet shopping experiences for those customers. Understanding outcomes and customer history is critical because it speaks to the user experience and the satisfaction of the user after they contacted Amazon.

82. Nor does Ms. King identify how many of the 400 randomly selected records supported each of her "themes." As presented, no theme was supported by more than five quotes and two were supported by just two quotes. Approximately 2.28% of the 219 selected records were presented in support of the most common themes, but a more appropriate analysis would focus at

[44] Emphasis added.

least on the 400 randomly selected records, in which case only 1.25% of the records supported the most common themes. Even then, however, those records were carefully selected by Ms. King from the broader set of over 152,000 records and this full collection of over 152,000 records represents just a fraction of the total in-app purchases during this period.

83. Finally, different conclusions can be drawn from even those limited quotes Ms. King selects. For example, Ms. King relies on the quote, "my 5 year old was playing a free game on my kindle and apparently unlocked sections of the game that you have to pay for" to support her conclusion that "some customers assumed that a free app would not have a component that allowed for paid purchases." But that selected quote can just as easily support the inference that the consumer already understood that free games could have "sections of the game that you have to pay for" when unlocked and chose to disregard the available controls to prevent accidental purchases when providing the device to a child.

D. Amazon Designed for Parents to Exercise Control Over Their Devices

84. Ms. King emphasizes the need for parents to be able to control their children's use of credit-card-enabled devices, yet she ignores the design objectives and strategies that Amazon employed when creating in-app purchasing for Amazon customers. The earliest consideration of design objectives on the part of Amazon Appstore engineers and business leaders acknowledged that parents would want the functionality to effectively control their children's ability to make purchases on an account-connected device, while simultaneously wanting the freedom to allow their children access to such devices. Those customer needs were fully anticipated, and the Amazon designers and planners included those objectives in their plans.

85. For instance, as early as May 2011, months before in-app purchasing was first introduced, Amazon's design team circulated a Business Requirements Document that set forth the proposed roadmap for in-app purchasing and the design objectives that had been identified.[45] Under the heading "Customer Needs," Amazon's designers and planners observed that "Parental Control" would be desired by parents, stating, "Adults buying for children (or enabling their children to buy for themselves) will want the ability to guard against large, unexpected charges -

[45] Dep. Ex. 47, Deposition of A. Paleja (July 14, 2015).

e.g. through addition of a PIN requirement or spending limits."[46] Ms. King ignores the fact that Amazon built into its design the tools necessary for responsible parents to exercise the supervision that they desired by "opting in" to enable Parental Controls.[47]

E. Dr. Rosenberg's Usability Test

86. As discussed above, a heuristic evaluation is not a substitute for usability testing, which has the advantage of directly observing and evaluating users' interaction with an interface. It is my opinion that a usability test would provide relevant information about the conclusions I have reached.

87. On December 4, 2015, after I had formulated my opinions discussed above, Amazon's counsel informed me that another expert, Craig Rosenberg, Ph.D., had performed a usability test in this case. I requested the opportunity to review it and did so.[48] Based on my review of the Rosenberg test, I have the following additional observations and opinions regarding how it relates to the opinions I independently reached:

a. Dr. Rosenberg's test is the type of usability test that I would find useful when evaluating a user-interface design.

b. Findings in Dr. Rosenberg's test buttress many of my opinions identified above. For example:

i. Nearly all test subjects viewing an app-description page understood that it was possible to incur additional charges within an app that offered in-app-purchasing opportunities;

ii. All or nearly all test subjects understood the language used in the "PLEASE NOTE" and Key Details notices relating to in-app purchasing and Parental Controls, and nearly all subjects understood that they could set Parental Controls to limit in-app purchases; and

[46] Ibid. at 00265969.

[47] Ibid. at 00265976.

[48] Expert Report of Craig Rosenberg (December 7, 2015).

iii. Nearly all test subjects understood the language used in the dialog box introduced in May 2013 for first-time in-app purchases relating to in-app purchasing and Parental Controls, and nearly all subjects understood the dialog box presented the option of requiring a password for future in-app purchases.[49]

IV. Conclusion

88. Heuristic evaluations of user interfaces are subjective and, particularly when limited to a single evaluator, are prone to overestimate the number and severity of potential problems. It is my opinion that the heuristic evaluation by Jennifer King is an incomplete analysis and draws unreliable conclusions. Contrary to Ms. King, I conclude that Amazon's initial design and refinement of its user interface relating to in-app purchasing were reasonable and consistent with the practice of the industry. They did not fall so far below contemporary, accepted business practices that they would have unfairly confused parents about the presence of and opportunity to make in-app purchases, about the presence of and opportunity to enable Parental Controls to restrict in-app purchases, or the availability of and means to request a refund for any accidental or unwanted in-app purchases.

Dated December 7, 2015

_____

Andrew L. Sears, Ph.D.

[49] The test results revealed similarly positive results for the version of the dialog box introduced in June 2014.