Learning Through Exploration: How Children, Adults, and Older Adults Interact with a New Feature-Rich Application

Shareen Mahmud[1], **Jessalyn Alvina**[1], **Parmit K. Chilana**[2], **Andrea Bunt**[3], **Joanna McGrenere**[1]
[1]University of British Columbia, Vancouver, Canada
[2]Simon Fraser University, Burnaby, Canada
[3]University of Manitoba, Winnipeg, Canada
[1]{shareenm, jalvina, joanna}@cs.ubc.ca, [2]pchilana@cs.sfu.ca, [3]bunt@cs.umanitoba.ca

ABSTRACT
Feature-rich applications such as word processors and spreadsheets are not only being used by adults but increasingly by children and older adults as well. Learning these applications is challenging as they offer hundreds of commands throughout the interface. We investigate how newcomers from different age groups explore the user interface of a feature-rich application to determine, locate, and use relevant features. We conducted an in-lab observational study with 10 children (10-12), 10 adults (20-35) and 10 older adults (60-75) who were first-time users of Microsoft OneNote. Our results illustrate key exploration differences across age groups, including that children were careful and performed as efficiently as the adults, whereas older adults spent a longer time and repeated sequences of failed selections. Further, their exploration style was negatively influenced by their past knowledge of similar applications. We discuss design interventions to accommodate these exploration differences and to improve software onboarding for newcomers.

Author Keywords
Age-related differences; lab study; desktop/laptop GUI; newcomers' exploration strategy

CCS Concepts
•Human-centered computing $\rightarrow$ **Graphical user interfaces; Empirical studies in HCI; •Social and professional topics** $\rightarrow$ *Age;*

INTRODUCTION
Modern feature-rich applications such as word processors, spreadsheets, and 3D modeling packages offer hundreds of commands organized under various menus, toolbars, and navigation structures. These applications are powerful and highly flexible, but can be overwhelming and difficult to learn [25, 30, 40]. One common way for users to learn a new application is to explore the functionality displayed on the user interface [5, 47]. However, exploring the interface of a feature-rich application can be challenging because users must determine which features are needed to accomplish their tasks, understand how individual features work (in isolation and together), and locate relevant features in the interface [23, 45].

Increasingly, newcomers to feature-rich software include a diverse group of users. For example, children are using various productivity and learning applications in digital classrooms [41, 58]. With greater flexibility in retirement age, older adults (65+) are working longer and learning to use new applications for knowledge work [3, 22, 53]. Prior work has shown that children can be more eager to explore software than adults [9, 29] whereas older adults can be fearful to explore new applications and can have lower confidence levels [10, 36, 59]. Other work has also shed light on how different age groups approach new technologies [3, 27, 51]. However, little is known about differences and similarities in how users explore the interface of a feature-rich application when learning to use it for the first-time.

Tackling the problem of interface exploration for feature-rich software is more important than ever before: by some estimates, at least 25% of users are abandoning an application after just a single use [44]. Application onboarding can be a crucial part of a user's journey [4], but there are few insights from the HCI literature about how to design such onboarding experiences, particularly those that support the natural interface exploration styles of the different age groups and keep them motivated to learn the application.

With this issue in mind, the core research questions that we explore in this paper are: What are the age-related differences in users' exploration styles when using a feature-rich application for the first time? In particular, how do different user groups determine, locate, and use relevant features within the application? How do they deal with performance breakdowns? Characterizing these differences in exploration styles of different age groups could help designers make more inclusive design decisions.

We conducted a structured observational study with 30 newcomers to a feature-rich application, Microsoft OneNote: 10 children (10-12), 10 adults (20-35), and 10 older adults (60-75). Our goals were to identify and characterize exploration

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376414

styles of the different age groups. We captured detailed interaction data (via software logs) to quantitatively compare exploration styles. We also supplemented our quantitative analysis with brief post-session interviews to obtain participants' retrospective assessments of what made exploration easy or difficult. Based on both our quantitative and qualitative findings, we propose design implications for applications seeking to foster efficient exploration and onboarding experiences for the different age groups.

Our work contributes: (1) simultaneous investigation of the interface exploration styles of three age cohorts – children, adults, and older adults; (2) identification of the challenges that each age group faces when exploring the interface of a feature-rich application to accomplish a goal; (3) identification of the different strategies that each age group uses to deal with hurdles during interface exploration; (4) design implications to support efficient interface exploration for the different age groups; and (5) a codebook that other researchers can leverage to investigate exploratory learning in GUIs.

RELATED WORK

Users' Approaches to Learning New Applications

Studies from back in the 1980s and 1990s to the late 2000s have shown that people often have difficulties learning a new application due to different past experiences with technology [5, 47, 51]. Despite the availability of online learning resources, built-in help, and manuals, people often tend to be reluctant to read and prefer to learn via self-directed exploration [6, 32, 51]. Although learning an application through interface exploration is a preferred strategy for many, users make more errors in the early stages of learning that can cause them to feel confused and frustrated [5]. We were motivated to explore software learnability of feature-rich applications across different user groups, and this body of prior work helped us to focus on exploration as the method of learning.

Age-Related Differences in Exploratory Learning

Multiple studies in the past have looked at how adults learn to use an application via exploration [18, 35]. For example, Carol-Ina Trudel's work in the 90s focused on adults' exploration of a computer-simulated digital watch and showed that poor learners had "negative exploration strategies" where they repeated moves that had no effect, did not pay attention to feedback, and inaccurately assessed what had been learned so far [57]. Similarly, another study by Novick et al. [47] investigated the usefulness of trial & error exploration with Microsoft Publisher, and found that adults relied on interface's signifiers, which sometimes took them in the wrong direction.

When it comes to children, fewer studies have identified their exploration behaviors with a new application. In the context of using a tablet, Couse & Chen [15] found that children between the ages of 3 to 6 were able to draw using a limited version of Microsoft Word when guided by adults. They were persistent during exploration even when they encountered errors. More recent work on children's use of a feature-rich 3D modeling application has provided insights on the barriers that they face when using help resources and found that children had difficulties locating the relevant UI elements and struggled to formulate help queries [27].

Significant work has also investigated how older adults learn new applications [1, 3, 7]. Some has explicitly made design recommendations to improve learning resources and reduce cognitive load [20, 24, 52]. Leung et al. examined how older adults learned to use a smartphone and found that participants switched between trial & error and reading the manual, and also benefited from a task list [36]. Other studies have focused on helping older adults to use social networking [26, 46], healthcare [11, 56] and smart home [19, 49] applications. While the above prior work provides valuable insights into how different age groups learn new applications, the age groups were studied separately, making more detailed comparisons and contrasts difficult.

Some studies have considered more than one age cohort when learning simpler applications, such as Chin & Tat-Fu [10] who pointed out that when learning from a link recommendation system, older adults took longer to click on a link by first deciding if it is relevant or not, whereas adults were more likely to click and see it. This could be because older adults have greater computer anxiety than adults [28]. In addition, O'Brien et al., found that older adults had more difficulties attributed to insufficient prior knowledge than adults when learning to use technology in their everyday lives [48].

Furthermore, prior research has focused on intergenerational gameplay and collaboration [17, 37, 50, 62], where older adults have been shown to require assistance to understand aspects of gaming that otherwise seemed intuitive to the younger adults [62]. Hence, they benefited from being paired with younger adults which reduced their anxiety [12, 17]. The closest work to ours is a study by Chimbo et al. [9] that looked at how children and adults explored a simple gaming interface and found that adults would not make a move that they were unsure about whereas children were open to trying out different actions to get ahead in the game. Our work builds on this prior work by specifically investigating how three different age groups explore an unfamiliar feature-rich user interface, without accessing external help (e.g., through manuals, peer help, online resources). In doing so, our study findings complement prior work by isolating the age-relate differences when users are restricted to self-exploration of an interface.

Interface Guidelines to Support Age-Related Differences

To help designers build technology for children, Chiasson & Gutwin [8] presented a catalog of design principles by gathering information from past research in HCI, education, and psychology. The design guidelines aimed to support children's cognitive, physical, social and emotional development. More recently, Soni et al. [55] found conflicting guidelines in the literature such that they do not equally benefit children in different age categories. Hence, although having guidelines specifically oriented towards children are valuable, the age ranges in previous work vary considerably and more work is needed to test these guidelines across various age categories.

When it comes to designing interfaces for older adults, several studies in the HCI community [2, 31, 61] have proposed design recommendations to address older adults' physical and cognitive needs as well as their varying experiences with technology. Darejeh & Singh [16] conducted a literature review to extract design principles for older adults. They found that novice elderly users benefited from a reduced feature set, descriptive text, and system feedback. In addition, Fisk et al., [20] discuss the importance of letting older adults undo their actions, and helping them understand their interaction histories.

Lastly, Neilson, Molich, and Shneiderman [43, 54] have proposed universal user interface design guidelines in the 1990s that have continued to be revised over the years [28]. These guidelines highlight the importance of informing users about the system state and helping them recover from errors. They are not age-specific but we assume that they would be useful for different groups of users.

In our work, we discuss some of our design recommendations in the context of the guidelines above. We also introduce specific design implications that address differences in exploration of software newcomers from different age groups.

USER STUDY: STRUCTURED OBSERVATION

We conducted an in-lab observational study to understand how children, adults, and older adults explore a feature-rich application when they are learning to use it for the first time. Our goal was not to test any hypothesis, but rather to observe how the participants determine, locate and use relevant features within the interface. We also sought to understand the strategies that they use to deal with any hurdles during exploration. Our choice of feature-rich application for the study was Microsoft OneNote, a note-taking application being used in work settings, as well as by children in schools [58]. Our use of MS OneNote, which is recent compared to applications such as MS Word, also facilitated the recruitment of first-time users.

Participants

We recruited 10 children (10-12), 10 adults (20-35), and 10 older adults (60-75) with a gender balance. We selected these age ranges in order to minimize the overlap in the physical, technical, and cognitive abilities of participants. The study was advertised in a local school, university, community centers, and newspapers. All 30 participants were first-time users of OneNote, comfortable with using a computer, and free of any motor impairments. Each received $20 for participating.

Apparatus

A Microsoft Surface laptop with Windows 10 and OneNote installed was used for the study. The Tobii Eye Tracker 4C was attached to the laptop to record participants' gaze pattern.

Tasks

To observe participants' exploration styles when using OneNote for the first time, we gave them a set of tasks. They were asked to imagine that they were taking two online classes, Art and Science, for which they had to maintain a notebook. There were four main tasks: creating a new notebook, adding notes related to the Art class, adding notes related to the Science class, and sharing the notebook with a friend. The 'Create Notebook' and 'Share Notebook' tasks could ideally be completed in one step ('Minimum Selections' in Table 4) whereas

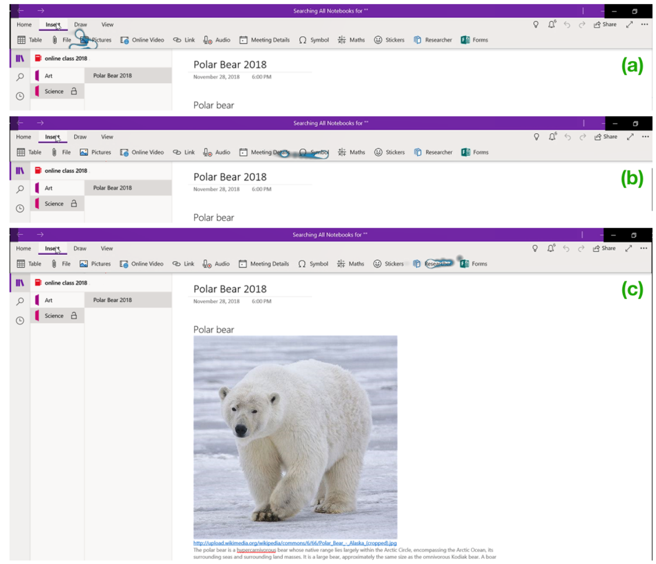

Figure 1. The interface of Microsoft OneNote. (a) and (b) show an example of C2's eye gaze (the blue bubble) traversing the features under the 'Insert' menu, moving from 'Pictures' (a) to 'Meeting Details' (b) and finally to 'Researcher' (c) while the mouse cursor stays on 'Insert'. The final result of the Science task is shown in (c).

the 'Art Task' and 'Science Task' tasks had more complexity. The Art and Science tasks had 7 and 5 sub-tasks respectively, consisting of, for example, drawing a flower, finding information about polar bears within OneNote, and locking the page with a password. Figure 1(c) shows an example of the final result for the Science task. Creating the notebook was always the first task, and sharing the notebook was always the last task. We counterbalanced the Art and Science tasks across participants. We iteratively refined the tasks and ran 6 pilot participants to ensure that they included features under different menu areas, were comprehensible for all age groups and could be completed in an hour.

Procedure

At the beginning of the study, the participants filled out an expertise questionnaire and a self-efficacy questionnaire [14]. The expertise questionnaire collected their demographic information and captured their level of computer and application expertise. The self-efficacy questionnaire asked the participants to rate their confidence level of using a new note-taking application under a variety of conditions. The conditions included having someone to help them get started, having used a similar application before, and being able to use the built-in help. After participants had answered the questionnaires, the eye tracker was set up.

Next, we introduced participants to the conceptual model of OneNote, explaining that it is similar to a physical notebook, where one can create a notebook and add several sections and pages within that notebook. Then, we gave them the list of tasks to carry out. We encouraged the participants to explore the interface in whatever way they preferred but also told them that they were not to use external help resources. Since we had different age groups taking part in the study, we motivated them by acknowledging that the tasks were intended to be slightly difficult and that they would be given a hint if they were considerably stuck. We gave the participants a hint after 3 minutes of being stuck with a sub-task and the option to move on to the next sub-task after 5 minutes of not making any progress. Each participant had one hour to complete the four main tasks and was requested to think-aloud during the session. As participants worked on the tasks, the laptop screen was recorded along with the audio and the gaze movement.

Once the tasks were completed, for comparison purposes, participants filled out the same self-efficacy questionnaire that they answered at the beginning of the study, only this time they were asked to reflect on their experience after using OneNote. The sessions concluded with a brief semi-structured interview for 10 minutes where we asked the participants to describe their exploration strategies and their overall experience of using OneNote for the first time. We also probed on any specific struggles that we observed during their exploration.

DATA ANALYSIS

Video Coding and Event Generation

We started our data analysis by coding the observational screen recordings of the participants' interaction with OneNote. The goal of the video coding was to generate and tag a set of events related to participants' exploration activities. Three authors were heavily involved in iteratively creating a codebook over a period of 3 months. The first author inspected 3 full videos to come up with the initial codes. A second author then randomly selected 2 videos and independently created their own codebook. Disagreements were discussed and settled with revisions. A third author then probed deeply on the clarity of description, completeness, and discriminability of the codes. The codebook included events such as selecting a menu or a feature, performing an action such as typing or drawing, repeating irrelevant selections, etc. We took inspiration from Trudel's [57] coding scheme of classifying exploration events when users' interacted with a digital watch and modified the event types based on the participants' interaction with OneNote. Table 1 shows our codebook with the list of events tagged. We provide a snippet of the data file from a participant's coded video in the supplemental material to the paper.

Each video had over 250 occurrences of the logged events. For most of the events in the codebook, we only analyzed their frequency of occurrence. For five of the events (including skimming and off-task actions), we also analyzed the time duration of each occurrence. We then performed statistical analysis on the data files using the REML procedure with SAS JMP, followed by Tukey HSD tests for post-hoc comparisons to understand if the number of occurrences and duration of the tagged events were different across the three age groups.

Table 1. Final list of events tagged. We counted the occurrence of all events. An event with (*) also has a duration.

Task Status	
Start Task	The user starts a sub-task
End Task	The user completes a sub-task
Give Up	The user is unsuccessful with the task despite getting hints and has not made progress for 5 minutes
Selection: The user clicks on a menu or feature	
Unique Selection	The user selects a feature for the first time
Non-Unique Selection	The user selects a feature already selected before
Successful Selection*	The user selects the right feature needed for the task
Off-Task Selection*	The user selects a feature irrelevant to the sub-task
Tooltip Selection	The user makes a selection after looking at the tooltip as captured by eye gaze
Incidental Selection*	The user selects a wrong feature for the current sub-task but that feature will be required for a later sub-task
Perform Task: After selecting a feature, the user starts any action related to executing the feature (e.g. drawing)	
Successful Action*	The user carries out actions required to succeed in the task after making a successful selection
Off-Task Action*	The user carries out actions insufficient to succeed in the task, either after making a successful selection or by making an off-task selection
Exploration Details	
Skim*	The user looks at a series of features on a sub-menu but does not select anything and moves on to another menu (captured with eye gaze). It ends when the user makes a selection or goes back to reading instructions
Missed Feature	The user does not notice the relevant feature despite being on the correct sub-menu or does not select the feature despite looking at it as indicated by the eye gaze
Unsure	The user does not understand or notice the outcome of their selection as indicated by their actions
Retry	The user selects the same feature for the same sub-task
Cycle	The user repeats the same sequence of selections for the same sub-task
Right-Click	The user opens the right-click menu
Depth	The user selects another menu
Breadth	The user selects a feature on the same sub-menu
Icon	The user selects a feature with only an icon
Text	The user selects a feature with text
Organize	The user closes tabs that are not required for the task
Eager	The user questions the interface/design
Undo	The user selects 'Undo' to revert an action
Built-In Help	The user selects the built-in help
Hint	The user gets a hint to try out a different option
Non-Exploration Events: The user's eye gaze is not on the screen	
Read Instructions	The user is reading task instructions or clarifying with the researcher
Idle	The user takes a break, e.g. to go to the washroom

Semi-Structured Interview

We used Braun and Clarke's approach to thematic analysis [13]. Two authors first analyzed the interview transcripts inductively using open coding, and then deductively based on the codebook from the video analysis. Then, all the authors discussed and iterated on the key themes during multiple meetings to contextualize the quantitative findings. Altogether, the recurring themes highlighted in children, adults, and older adults' responses were related to their overall experience of using OneNote, the specific challenges that they faced and their strategies for solving those challenges during exploration.

Questionnaire

We concluded our data analysis by looking at the responses from the three questionnaires filled out by each participant. We performed the Kruskal-Wallis H test on the expertise questionnaire along with post-hoc comparisons and a combination of the Wilcoxon signed-rank test and the Kruskal-Wallis H test on the before and after self-efficacy questionnaires.

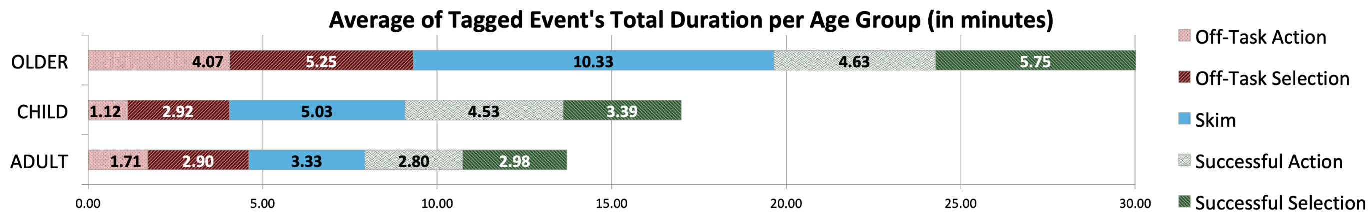

Figure 2. Average of the total duration per age group categorized by the tagged events

Participants' Expertise

Before moving on to our primary results, we describe our participants' expertise as analyzed from the demographic data.

We ran the Kruskal-Wallis H test on the expertise questionnaire data (Table 2). Participants could rate their years of experience per device from 'Less that 1 year' to '11 years and above' and their frequency of use per application from 'Never' to 'Daily.' We found a significant difference in means between the age groups when it came to their number of years of experience with desktops ($H = 14.561, p < .001$), laptops ($H = 10.292, p < .006$), and smartphones ($H = 8.810, p < .012$). A post-hoc test to check pairwise comparisons showed that children had significantly fewer years of desktop experience than adults ($p = .004$) and older adults ($p = .002$) whereas adults' and older adults' years of desktop experience was not significantly different. However, when it came to the years of laptop and smartphone experience, adults had significantly more experience than both children and older adults ($p = .011$ and $p = .024$ for laptops, $p = .048$ and $p = .019$ for smartphones) whereas children's and older adults' years of laptop and smartphone experience was not different.

Table 2. Expertise Characteristics of the Three Age Groups (Median). A category with (*) indicates significant difference between age groups.

	Children	Adults	Older Adults
Years of Experience per Device			
Desktop*	3 - 4	8 - 10	8 - 10
Laptop*	1 - 2	8 - 10	1 - 2
Tablet	3 - 4	3 - 4	3 - 4
Smartphone*	3 - 4	8 - 10	3 - 4
Frequency of Use per Application Category			
Word Processors*	Monthly	Daily	Monthly
Spreadsheets*	Never	Weekly	Rarely
Presentation*	Rarely	Monthly	Rarely
Email*	Monthly	Daily	Daily
Other Note-Taking	Rarely	Rarely	Rarely

Next, we looked at participants' frequency of use for different categories of applications over the 6 months prior to taking part in the study. We found a significant difference in their current frequency of use for word processing ($H = 10.191, p < .006$), spreadsheet ($H = 14.280, p < .001$), email ($H = 18.347, p < .000$) and presentation ($H = 6.975, p < .031$) applications. Pairwise comparisons revealed that adults used word processing, spreadsheets and presentation applications significantly more often than both children and older adults ($p = .02$ and $p = .015$ for word processors, $p = .001$ and $p = .044$ for spreadsheets, $p = .001$ and $p = .002$ for presentation) whereas children's and older adults' current usage of these applications was not significantly different. Older adults only used email applications significantly more often than children (p = .001). Also note that 7 of our older adults were still working (4 part-time, 3 full-time) and 3 were retired.

To summarize, our adults had the most experience with technology, and although our older adults had more years of desktop experience than children, their current usage of computer applications was not very different from children.

RESULTS

We present our results by first providing an overview of children's, adults', and older adults' performance of using OneNote for the first time. Next, we look at their interface exploration and unpack the differences in: (1) how the participants in different age groups determine, locate and select relevant features within the application and (2) how they deal with performance breakdowns and evaluate the usefulness of their performed actions. Finally, we touch on their overall feeling after using OneNote. In addition to the different age groups, we considered gender as an independent variable but found no significant difference in our results. Hence, to simplify, we focus on age-related differences.

Overall Performance

Children are almost as fast as adults, older adults are slower

The overall task completion times were not significantly different between children and adults, whereas older adults took a significantly longer time than the other two age groups (main effect: $F_{(2,27)} = 22.1201, p < .001$ and pairwise comparisons: $p = .001$ for older adults - adults, p = .001 for older adults - children). Children and adults completed the four main tasks with all the sub-tasks whereas 4 out of the 10 older adults gave up on at least 1 sub-task. On average, the total time to complete all four tasks (excluding the 'Read Instruction' and 'Idle' times) was 14.88 minutes for adults, 18.29 minutes for children, and 32.24 minutes for older adults.

The aggregated logged data clarifies which activities required the most time (Figure 2). It shows that children spent a bit longer than adults skimming the interface (1.5x more) and performing successful actions (1.6x more), such as drawing a flower. Surprisingly, they spent less time performing off-task actions that were not sufficient to succeed in the task, such as trying to insert a file from the computer instead of adding an online picture (1.4x less).

Older adults were the slowest overall, taking 2.2x longer than adults and 1.8x longer than children. This is partly expected as prior research has shown that, in general, older adults take more time to complete movement tasks and take more pauses [2, 31]. Although Figure 2 shows that older adults took a longer time than children and adults in many of the categories, the biggest contributors were the time they spent performing off-task actions (2.4x slower than adults and 3.6x slower than children) and skimming the interface (3.1x slower than adults and 2.1x slower than children).

To better understand the skimming and off-task action activities, we considered both the number of occurrences and the duration per event. We found a significant main effect of the occurrence of skimming and off-task action events ($F_{(2,27)} = 22.0034, p < 0.0001$ for skimming and $F_{(2,27)} = 5.0963, p < 0.05$ for off-task actions). Pairwise comparisons showed that older adults skimmed the interface and performed off-task actions significantly more often than both children ($p = .004$ for skimming, $p = .006$ for off-task actions) and adults ($p = .001$ for skimming, $p = .008$ for off-task actions). On average, the older adults skimmed 53.1 times and performed off-task actions 17 times during the whole study, while adults and children only skimmed 28.2 and 31.2 times respectively, and both performed off-task actions 9.2 times.

In the next subsection, we elaborate on how children's and (especially) the older adults' interface exploration styles resulted in their longer task completion times.

Interface Exploration Styles

Children and older adults face different challenges in locating relevant features

Adults were significantly quicker at making successful selections than older adults (main effect: $F_{(2,27)} = 5.0797, p < 0.01$ and pairwise comparisons: $p = .018$), whereas children were in between. The mean of the median duration per episode of successful selections was 4.2, 4.9, and 6.5 seconds for adults, children, and older adults respectively.

Older adults particularly struggled to locate the relevant features because they did not sufficiently investigate the features in different menus. If we consider the number of non-unique selections that were made by the three age groups, where they selected an option that they had tried before, older adults made significantly more non-unique selections than the other two age groups (main effect: $F_{(2,27)} = 3.6534, p = 0.04$ and pairwise comparisons: $p = .002$ for older adults - children, $p = .015$ for older adults - adults). The median percentages of non-unique selections were 55.5% for older adults compared to 43.9% and 47.5% for adults and children respectively. This is perhaps because older adults had the tendency to only select the sub-menu where they expected the feature to be, and dismissed other potential menus. For example, O8 struggled to find the 'Share' icon that was located at the top-right corner menu of the application. When asked, O8 said "*I thought those things won't add value. When you write on paper you write from left to right. The headings are always on the left. I assumed the important things to be on the left side.*" O6 expressed a similar thought regarding where he expected certain features to be located: "*Why are there these things [create notebook/section] at the bottom? I would expect them to be at the top, maybe under View.*"

Children faced different challenges when locating the relevant features. Even though their task completion times were similar to that of adults, they sometimes felt overwhelmed by the number of different menus and features. For example, C8 said "*It was kind of complicated. Just like they need to make it more simple in a way. It's confusing why some options are down here [a sub-menu] and some up there [top left corner menu].*" The majority of children were looking at the options at the top menu but they sometimes did not initially realize that there were more options outside 'Home'. For example, C10 said "*I didn't know that if you pressed it [Insert] it would give more options inside it. Once I looked outside Home it became obvious*" In addition, children particularly had difficulties locating features in the right-click (contextual) menu. The total number of right-click events was significantly lower for children (median 3 times during the whole study), than adults (10 times) and older adults (8.5 times), (main effect: $F_{(2,27)} = 6.4952, p = 0.0053$ and pairwise comparisons: $p = .01$ and $p = .01$) . Only 3 out of 10 children reported that they knew that they could right-click for more options. C8 said "*I actually don't use the mouse usually so I didn't know about right-clicking.*" In addition, C10 mentioned how she discovered the right-click menu accidentally '*I meant to click on it but I pressed both [right and left click] and it [right click menu] showed up.*"

Adults, being more frequent users of similar software applications, made more efficient use of the different menus resulting in their quicker successful selection times. Contrary to children and older adults, the majority of the adults found the features easy to locate. For example, A4 said "*I think they [the features] were relatively simple to discover. There were obvious features like the Draw menu. You probably saw me right-clicking several times. If it didn't have what I was looking for then I would look somewhere else.*" Similar to A4, A6 also mentioned how he would just explore another sub-menu if he could not find the feature where he expected it to be. "*I would firstly go to the menu that looks most relevant and scan from left to right. If this is not the right one, go to the next toolbar.*" Therefore, overall, adults often found the relevant features because they were comfortable exploring different menus, including the right-click menu.

Older adults struggle more than children and adults to determine the relevant sequence of selections

Older adults missed selecting a significantly higher number of correct features than children and adults, despite being on the right sub-menu (main effect: $F_{(2,27)} = 3.5099, p = 0.04$ and pairwise comparisons: $p = .004$ for older adults - children, $p = .012$ for older adults - adults). The median number of missed features was 15.5 for older adults, 9 for children and 7.5 for adults. This indicates that even though older adults were sometimes on the right track, they got confused about the features that were relevant to complete the task. For example, during the Science task, O6 had already determined the relevant sequence of selections and opened the 'Insert' menu but missed to click on the 'Researcher' feature which was the second last option from the right. In the semi-structured interview, he explained that he thought that the feature would be something like a 'magnifying glass symbol' under the 'Insert' menu. Similarly, there was O10 who had 32 instances of missed features, the highest among all the participants. When

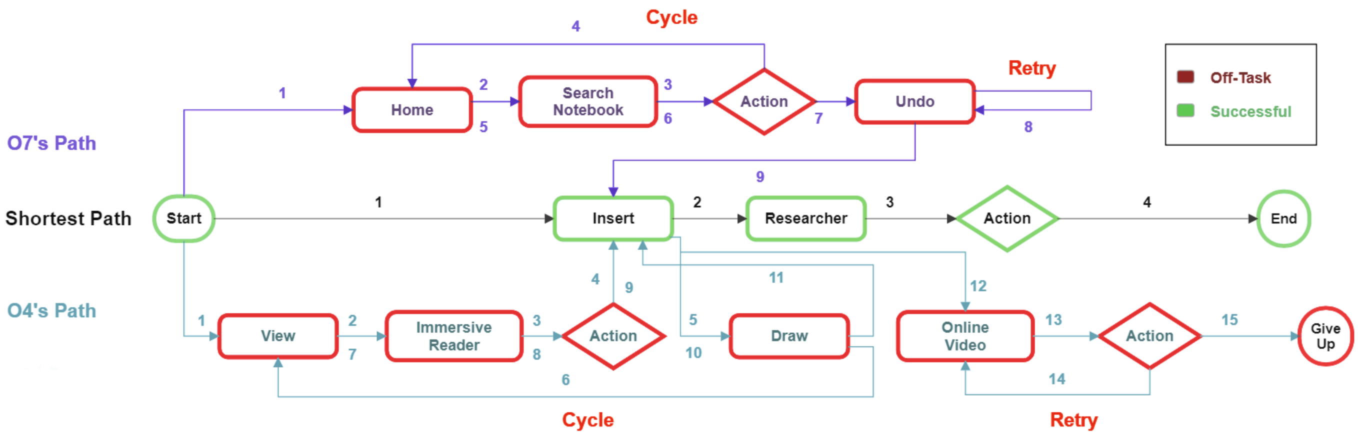

Figure 3. Example of the paths taken by O7 and O4 during a Science sub-task. Each of them deviated from the shortest path and had off-task selection cycles and retries. O7 eventually got on the right path and completed the sub-task whereas O4 gave up

asked about why she missed some of the features she said "*Maybe I am too used to looking on my left-hand side than look at my right-hand side to search for the commands. If I got used to it I think I would try to look at both sides.*"

Furthermore, some older adults tried to determine the selection sequence by searching for words similar to the task instructions on the interface. For example, O2 mentioned "*I didn't have any plan on how to approach it. I would just rely on the instruction page and then go to see if I can find a similar term. It all depends on the information available here [the task list] and here [the interface].*" This approach did not always work for them as the task list would not have keywords that could directly be mapped to the feature set.

Aside from missing the relevant features, older adults also performed off-task selections more often than adults and children (Table 3), a statistically significant difference for the Science task (main effect: $F_{(2,27)} = 4.8991, p = .002$ and pairwise comparisons: $p = .001$ for older adults - adults, $p = .002$ for older adults - children). This could be because the sub-tasks under Science were more challenging as the participants had to make the connection between the feature 'Researcher' under 'Insert' menu to 'inserting information about polar bears' and the 'Password Protect' feature under a right-click menu to 'locking the Science section with a password'.

When we further looked at the nature of the off-task selections made by the older adults in the Science task, we found that older adults had a significantly higher number of cycles, than the other two age groups (main effect: $F_{(2,27)} = 5.9271, p <$ .001 and pairwise comparison: $p = .01$ for older adults - children, $p = .04$ for older adults - adults). 7 out of the 10 older adults had at least one cycle where they repeated the same sequence of off-task selections or a retry of the same feature, both of which sometimes lead them to perform off-task actions. Figure 3 shows the deviation from the shortest path to complete a sub-task under Science by O7 and O4. O7 eventually did complete the sub-task whereas O7 gave up in the end without finding the 'Researcher' feature. The struggle to determine the relevant sequence selection sometimes caused frustration among older adults where they required a hint to proceed in the right direction.

Adults and older adults rely on their past software experience, children rely on real-life experience

The adult participants reported that they often guessed the relevant features based on their past experience with similar software, e.g. Microsoft Office, to determine the possible menus and features. For example, talking about her discovery of relevant features, A8 said "*It's pretty similar to Word. There are more different tools to explore here. I did a lot of inference from my prior knowledge to guess and check.*" This suggests positive knowledge transfer.

Although rarely occurring, we also observed a negative knowledge transfer for adults, when a participant made a wrong assumption based on their knowledge of previous software. For example, A2 had gone to the 'Home' menu to create a new notebook mentioning that he thought he should go to the menu similar to the 'File' menu of Microsoft Word to create a new file. This may potentially hinder efficient exploration if the user incorrectly narrows down the scope of exploration.

Similar to the adults, the older adults were also affected by their past experience of using similar applications but for them, negative transfer learning seemed to dominate. For example, when O9 had to add the current date to his notebook, instead of simply typing it in, he was looking for a function called 'Date' as Microsoft Word would have it. He said "*It doesn't seem logical that I can't press on some function to insert a date like in Word.*" O1 also expressed similar views regarding her confusion when she tried to relate OneNote with Microsoft Word: "*I was not too sure how to go back Home. It seems a little bit different from what I mostly work on. I mostly work on Word and whenever I am not too sure I will just go back to Home, this [OneNote] doesn't seem to work like that. It has to depend on my luck.*"

In contrast, children seemed to rely more on their real-life experiences as they had less computer experience, which had a positive effect on their exploration. For example, explaining how he found the 'Researcher' option, C2 mentioned "*When the teacher gives you homework to do and you don't know much about it, she says you do research. So I thought Researcher was something that could research for you.*" Similar to C2, C1 also expressed how he mapped the features to real-

Table 3. Median Number of Selections by Age Group Per Task. Older adults performed more off-task selections, especially during the 'Science Task'

Age Group	Create Notebook Task			Art Task			Science Task			Share Notebook Task		
	Minimum Selections	Successful Selections	Off-Task Selections	Minimum Selections	Successful Selections	Off-Task Selections	Minimum Selections	Successful Selections	Off-Task Selections	Minimum Selections	Successful Selections	Off-Task Selections
Children	1	1.1	2.7	14	18	12.7	8	9.4	9.4	1	1	5.5
Adults		1.1	4.5		19.8	11.7		8.5	11.9		1.1	4.7
Older Adults		1.1	5		21.3	11.4		10.9	21.6		1.1	5.8

life examples "*Like, if you look at the Ruler option, I think about the ruler straight away. You can measure. Math means you can do plus, minus, division.*"

Facing Breakdowns

Children are quick at recovering from breakdowns, older adults take time

Children had the shortest average total duration of off-task actions, only 1.1 minutes, compared to 1.7 and 4.1 minutes for adults and older adults (Figure 2). This suggests that children were able to understand the system feedback and address the outcome of their actions. Even when they made an off-task selection, they were quick to detect it and move on to another option. For example, in the Art task, a lot of participants did not realize right away that they had to activate the drawing mode by clicking a 'Hand' icon first before drawing. Talking about her experience of being able to quickly recover from an off-task action, such as drag the pen tool icon to draw which was not a supported interaction, C3 said "*I was dragging it [pen] as I was really trying to draw. Then I looked again and saw the hand and hovered over that to see what it did.*"

Similar to children, adults were also relatively good at understanding the feedback from the system to determine whether things worked as they expected or not. Some adults also went through the breakdown of not clicking on the 'Hand' icon and only selecting a pen. However, they were also able to quickly realize that they had to find a workaround. A2 mentioned,"... *because it wasn't working, this is like the obvious [Hand] icon, like the finger, just that icon is very readable.*"

Older adults, on the other hand, spent a significantly longer time carrying out various off-task actions than children and adults (main effect: $F_{(2,27)} = 9.1356, p < .01$ and pairwise comparisons: $p = .04$, $p = .04$) and were also significantly more unsure of their actions (main effect: $F_{(2,27)} = 9.5846, p < .01$ and pairwise comparisons: $p = .002$, $p = .003$ for older adults - children and older adults - adults). Reflecting on his experience with the drawing sub-task, O3 mentioned how he did not know he had to select both the pen tool and the 'Hand' icon: "*When you asked me to draw, I thought it was challenging, I couldn't think of the fact that I had to choose a color [pen]. If you weren't here [to give a hint] I would either have to go to the web or call somebody to ask.*" In addition, many of our older adults had difficulties understanding system feedback. For example, some selections trigger interface changes such as a pop-up menu or a mode-switch (e.g. clicking on the 'Hand' icon only activates the drawing mode), rather than a change to the data. Older adults often did not realize this difference, and tended to use the 'Undo' button hoping to undo a selection's impact on the interface. For example, in Figure 3, O7 is seen selecting 'Undo' hoping to undo the effect of the 'Search Notebook' selection that had impacted the interface by opening up a pop-up menu but she did not understand that she was removing her content instead. O2, O4, O6, O9 and O10 also faced this breakdown where they were unable to understand the outcome of selecting the 'Undo' button and accidentally removed their data.

Children rely on reading when facing a breakdown, older adults try out random options

Children were careful in making selections and did not want to do something wrong. Their fear of trying out wrong options could possibly be due to past experiences, such as C10 who mentioned "*I have used the Word document before, I accidentally highlighted the text. Then deleted it. I tried to solve it myself first and then I was at school, so I asked my friends and then the teacher.*" When facing a breakdown such as not being able to locate the correct feature or being stuck with an off-task action, 7 out of 10 children said that relied on the text labels to make an educated guess about the correct feature. Referring to how text supported him during breakdowns, C8 said "*Well, I would just kind of go into 'Insert', 'Draw', 'View' and read the options. Then work off from there.*" Figure 1 (a) and (b) further show an example of C2, who opened the 'Insert' menu and scanned the text labels systematically, as indicated by her eye-gaze, before moving the cursor to making a selection. Besides reading the text label, 2 out of the 10 children mentioned that they found the tooltip to be useful such as C10 who stated "*When I was in Home, I didn't know what they meant so I would go to the option and then stay on it. Then I could read what it was.*"

Adults and older adults, on the other hand, resorted to trial & error when they got stuck. However, for older adults, this approach was less effective as they would often click on random features just to see if they would work once their initial sequence of actions was unsuccessful. Overall, 8 out of 10 older adults mentioned trial & error. For example, talking about her approach to deal with breakdowns O8 said "*Because I often couldn't find a logical sequence, I was clicking on everything one after the other to see if something fit.*" Similar to O8, O5 expressed "*It was more hit and miss. It wasn't that straightforward. I was trying out things many times to see if it works.*" This further explains why older adults had a significantly higher number of off-task selections and cycles as previously discussed.

Although adults also adopted the trial & error approach when they faced a breakdown, they seemed to have better deduction strategies instead of selecting options in a random manner. For example, A6 said "*I think my approach is read, try it, click around. I would firstly go to the menu that looks most relevant. Mostly just try it. If this is not the right one, go to the next toolbar. It won't be a big deal to click on a wrong button.*" A3, A4, A5 and A9 also expressed similar ideas. A9 described her approach as "*Reading, seeing, trying. I wouldn't just choose random ones. I went with the first one that made sense if that did not work, I will look for another one.*" Therefore, it suggests that even when adults were doing trial & error, they were doing it in a targeted manner.

Overall Feeling

Children and adults seem fairly content with self-exploration, older adults feel disappointment

To understand how the three age groups felt about using a new application and if it affected their confidence levels, we analyzed their self-efficacy questionnaires before and after the tasks were performed. Although the within-group analysis showed no significant change in the confidence levels within each group before and after they had used OneNote, our between-group analysis revealed another story.

Older adults started the tasks with significantly lower confidence levels than adults ($p = .035$) and children ($p = .039$), and even after interacting with OneNote, their confidence levels stayed significantly lower ($p = .008$ for older adults - adults and $p = .042$ for older adults - children). This indicates that their experience of using OneNote did not improve their initial low confidence levels. If anything, it caused frustrations, as O7 mentioned "*It's a headache. You shouldn't have to touch this touch that. It's too much of clicking here and clicking there. If I want to do something quickly I don't want to click so much. I just wouldn't use it.*" Similarly, O1, O2, O8, O9, and O10 also reported that they felt unhappy with their experience. O3 further mentioned that he had expected the interface to be more intuitive "*It wasn't as intuitive and as easy as I thought it would be. I guess with any new application I need practice to get used to it.*"

In contrast, children and adults maintained a high confidence level both before and after they had used OneNote. It seems that adults had expected OneNote to be more complicated, like some of the other applications they had used in the past. For example, A5 mentioned "*It's simple and I guess the design, it's not as complicated as I thought. I expected something like, you know, Microsoft Excel.*" Overall, 9 out of 10 adult participants expressed similar views of finding OneNote easy to use, with the exception of A2 who found it difficult: "*I found it a bit hard to use. It's not intuitive at all. Like there are some conventions where you would expect the options to be and I don't see them there.*"

Similar to adults, children also expressed satisfaction with their experience of using OneNote, although they mentioned that the interface could be further simplified. Such as C1 who said "*I am happy but what I dislike is there is all this [sub menu options]. Then you have these [main menu options]. You don't know where what is.*" In addition, C10 mentioned how he understood the interface better as he progressed in the tasks: "*For this one [Science], it was a bit difficult. But when I did Art I got used to it.*"

DISCUSSION AND IMPLICATIONS FOR DESIGN

We reflect on our key findings and discuss their implications to foster efficient exploration for children, adults and older adults.

We were surprised to see that children performed almost as well as the adults when using Microsoft OneNote, for the first time. Prior work on children's use of problem-solving software had indicated that children tend to feel lost in different parts of an application and tend to try out many different actions to get ahead [9, 27]. Yet, we saw that even when children were exploring the application, they seemed careful to avoid making mistakes. They often read the text labels and sometimes the tooltips for guidance. Even when they had a breakdown, i.e. carrying out off-task actions, they were quick to detect it and recover. In contrast, older adults fell into the active user paradox [5] where they resorted to less systematic trial & error strategies once their initial sequence of actions had failed. This confirms prior work on older adults being more negatively affected by errors [10, 36, 59], which then impacts their initial exploration strategy and causes them to struggle.

While it is not surprising that older adults had slower task completion times – there are natural declines in cognitive and motor abilities due to aging, as well as documented fears of exploring a new application [10] – our study revealed additional factors that contributed to their longer task completion times. Older adults had multiple selection cycles and retries during exploration where they deviated from the shortest path and selected the same sequence of irrelevant features, that often led them to perform unnecessary actions and consequently slowed them down (Figure 3). Adults and children, on the other hand, had significantly fewer such cycles in their interaction data, instead moving on to trying out different sets of options. This could be an indication of short-term memory loss where older adults forget the sequence of actions that they have carried out or are simply unsure of whether they have performed the actions correctly [61].

Implication for Design: *Detect selection cycles and offer support.* The system could detect cycles and retries in users' selections and offer support. A simple possibility would be to suggest the use of built-in help as some of our older adults did not notice that there was a built-in help option. In addition, based on the user's interaction, the system could invoke a heatmap showing the recent cycles and retries of selected options, similar to Patina [39], and encourage the users to better understand the use of those features by using the tooltip. Moreover, seeing a visual representation of their repeated selections could motivate users to reflect on their previous actions (as advocated in prior work on older adults [20]) and prevent them from repeating the wrong path. In providing this increased, personalized support, the hope would be that older adults would experience greater task success and be less inclined to abandon the application [34].

Another surprising finding was that the impact of knowledge transfer on exploration style seemed to be different for children, adults, and older adults. Adults were the most experienced with recent technologies, which generally seemed to help them perform well. Although our older adults reported comparable past experience with similar applications, they had not used these regularly in the past six months. This might have caused a mismatch in their mental models for how they expected the application to behave and the way that it is designed today. Hence, although they spent a long time skimming the interface, they were making more non-unique selections. This is consistent with prior work on understanding the effect of older adults' prior knowledge on interactions with technology in general [33, 48], where our observational study provides additional insights on how having past knowledge of similar applications more specifically does not necessarily make self exploration efficient. Children, on the other hand, had used fewer computer applications than our older adults but were still as confident as our adults. They were most familiar with using tablets that offer relatively simpler applications than desktop computers. Their lack of experience with personal computers and pointing devices (e.g. a mouse) may hinder their discovery of interactions such as the right-click menu, which has the metaphor of right-clicking a mouse, something more foreign to them.

Implication for Design: *Support feature discovery and skimming by appropriately revealing signifiers for hidden menus, including drop-downs and the right-click menu.* The interface could detect the user's eye-gaze on the screen together with skimming behaviour and then subtly provide signifiers. For example, the interface could highlight a 'Reveal' button similar to ExposeHK [38], which the user could click on to discover hidden menus. These signifiers could be ignored or acted upon, without unnecessarily cluttering the interface for all users. In addition, this might inspire users to explore the menu regions they had not previously paid attention to.

Lastly, while prior work has stressed the importance of 'Undo' mechanisms for older adults [20], we were surprised to find that older adults struggled to understand its scope although it is universally used today. They often tried to use 'Undo' hoping to undo a selection that had impacted the interface, whereas the current 'Undo' mechanism only acts on selections that involve operations on data. Hence, some of the older adults ended up selecting 'Undo' incorrectly and accidentally removing important content, which caused confusion.

Implication for Design: *Provide feedback of user actions for both changes to the data/content and to the interface, and enable users to undo either type of action.* Although recent work has widely explored 'Selective Undo', where users can undo specific operations instead of backtracking in a linear manner [42, 60], we recommend expanding the 'Undo' scope by distinguishing between undoing an operation that affects the data/content or merely undoing a selection. For example, the system could offer a feature such as 'My Past Actions' that could show the user a list of the features that they had selected along with whether a feature caused a change to their data/content or just to the interface. The user could then hover on the list item to undo the effect of certain selections.

We envision these design recommendations to be particularly useful for both children and older adults. At the same time, they would not get in the way of adults who do not need the extra assistance, as the support would be triggered based on the user's interactions. The adults in our study did not seem to struggle very much. We do not interpret this as inconsistent with prior work, but rather complementary – OneNote is not likely as feature-rich as applications such as Photoshop or AutoCAD, with which adults have been shown to have difficulty [21, 32]. The nature of our tasks might also not have been as complex for the adults as those studied before.

Altogether, our findings have implications for the future design of personalized application learning support. One crucial point for such learning support is right at the outset of use, when onboarding a newcomer. Applications today commonly include a generic "getting started" type tutorial, but none, to our knowledge, are adapted based on age or exploration style. This may partially explain, in our view, their limited success. One possibility would be to have educational versions of software, such as OneNote, provide onboarding that specifically covers interactions like the right-click menus, given that our children were largely unfamiliar with them. Onboarding tutorials for computers in senior centers could alternatively educate on the scope of 'Undo' and the value of reading tooltips.

Beyond our study findings, our work also contributes a detailed codebook, which substantially extends the codes provided in prior work [57]. Besides other researchers using our codebook for investigating exploratory behaviour in GUIs in the future, it could be leveraged for more automatic coding of data streams. An example would be using the codes to label features for machine learning purposes, to train a model to automatically detect effective exploration behaviours across age groups.

LIMITATIONS

Although our study provides insights on the interface exploration styles of the three age cohorts, it is limited by the fact that we investigated their performance with only one type of application- Microsoft OneNote. Future work could expand the choice of application with various degrees of complexity, beyond productivity, and investigate the effectiveness of the design implications with a broader sample size. While we chose to limit their help-seeking approach to specifically observe their interface exploration style and gave participants tasks to ensure consistent experience, future work could also consider supporting other means of help-seeking along with a task-free approach.

CONCLUSION

When learning a feature-rich application for the first time, users often explore different menus and features to accomplish their desired tasks. Today, these applications are being used by children, adults and older adults alike. Our study contributes insights into the interface exploration styles of the three age groups, the challenges that they face and the strategies that they use to deal with breakdowns. We found, among other things, that children explore the interface carefully but struggle to locate contextual menus (because of lack of mouse exposure), whereas older adults have difficulties determining relevant sequence of features and repeat failed selections. Our work is an important step towards understanding the diversity in users' approaches to learning through exploration which has implications for improving their application onboarding experiences through design.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

[1] Anne Aula. 2005. User Study on Older Adults' Use of the Web and Search Engines. *Univers. Access Inf. Soc.* 4, 1 (Sept. 2005), 67–81. DOI: http://dx.doi.org/10.1007/s10209-004-0097-7

[2] Maxim Bakaev. 2008. Fitts' Law for Older Adults: Considering a Factor of Age. In *Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems (IHC '08)*. Sociedade Brasileira de Computação, Porto Alegre, Brazil, Brazil, 260–263. http://dl.acm.org/citation.cfm?id=1497470.1497502

[3] Yvonne Barnard, Mike D. Bradley, Frances Hodgson, and Ashley D. Lloyd. 2013. Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. *Computers in Human Behavior* 29, 4 (2013), 1715 – 1724. DOI: http://dx.doi.org/https://doi.org/10.1016/j.chb.2013.02.006

[4] Jacobson Ben. 2018. Why Better Onboarding Means Less Churn: Proven Tactics You Can Try Today. (Jul 2018). https://postfunnel.com/better-onboarding-means-less-churn-proven-tactics/

[5] John M. Carroll and Caroline Carrithers. 1984. Training Wheels in a User Interface. *Commun. ACM* 27, 8 (Aug. 1984), 800–806. DOI: http://dx.doi.org/10.1145/358198.358218

[6] John M. Carroll and Mary Beth Rosson. 1987. The Paradox of the Active User. In *Interfacing Thought: Cognitive Aspects of Human-computer Interaction*, John M. Carroll (Ed.). MIT Press, Cambridge, MA, USA, 80–111. http://dl.acm.org/citation.cfm?id=28446.28451

[7] Micaela Y. Chan, Sara Haber, Linda M. Drew, and Denise C. Park. 2014. Training Older Adults to Use Tablet Computers: Does It Enhance Cognitive Function? *The Gerontologist* 56, 3 (06 2014), 475–484.

[8] Sonia Chiasson and Carl Gutwin. 2005. *Design Principles for Children's Software*. Technical Report HCI-TR-05-02.

[9] Bester Chimbo, J. H. Gelderblom, and Mary Ruth de Villiers. 2011. A comparative analysis of the meaning of 'learnability' for child and adult users.

[10] Jessie Chin and Wai-Tat Fu. 2012. Age Differences in Exploratory Learning from a Health Information Website. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 3031–3040. DOI: http://dx.doi.org/10.1145/2207676.2208715

[11] Wonchan Choi. 2013. What Makes Online Health Information Credible for Older Adults?: An Exploratory Study. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 2671–2676. DOI: http://dx.doi.org/10.1145/2468356.2479491

[12] Puay-Hoe Chua, Younbo Jung, May O. Lwin, and Yin-Leng Theng. 2013. Let's Play Together: Effects of Video-game Play on Intergenerational Perceptions Among Youth and Elderly Participants. *Comput. Hum. Behav.* 29, 6 (Nov. 2013), 2303–2311. DOI: http://dx.doi.org/10.1016/j.chb.2013.04.037

[13] Victoria Clarke and Virginia Braun. 2014. *Thematic Analysis*. Springer Netherlands, Dordrecht. 6626–6628 pages. DOI: http://dx.doi.org/10.1007/978-94-007-0753-5_3470

[14] Deborah R. Compeau and Christopher A. Higgins. 1995. Computer Self-Efficacy: Development of a Measure and Initial Test. *MIS Quarterly* 19, 2 (1995), 189–211. http://www.jstor.org/stable/249688

[15] Leslie J. Couse and Dora W. Chen. 2010. A Tablet Computer for Young Children? Exploring its Viability for Early Childhood Education. *Journal of Research on Technology in Education* 43, 1 (2010), 75–96. DOI: http://dx.doi.org/10.1080/15391523.2010.10782562

[16] Ali Darejeh and Dalbir Singh Valbir Singh. 2013. A review on user interface design principles to increase software usability for users with less computer literacy. *Journal of Computer Science* 9, 11 (2013), 1443–1450. DOI:http://dx.doi.org/10.3844/jcssp.2013.1443.1450

[17] Teresa de la Hera, Eugene Loos, Monique Simons, and Joleen Blom. 2017. Benefits and Factors Influencing the Design of Intergenerational Digital Games: A Systematic Literature Review. *Societies* 7 (07 2017). DOI:http://dx.doi.org/10.3390/soc7030018

[18] Sjaak de Mul and Herre van Oostendorp. 1996. Learning user interfaces by exploration. *Acta Psychologica* 91, 3 (1996), 325 – 344. DOI:http://dx.doi.org/https://doi.org/10.1016/0001-6918(95)00060-7 Usage of Modern Technology by Experts and Non-professionals.

[19] George Demiris, Debra Parker Oliver, Geraldine Dickey, Marjorie Skubic, and Marilyn Rantz. 2008. Findings from a Participatory Evaluation of a Smart Home Application for Older Adults. *Technol. Health Care* 16, 2 (April 2008), 111–118. http://dl.acm.org/citation.cfm?id=1378666.1378670

[20] Arthur D Fisk, Wendy A Rogers, Neil Charness, Sara J Czaja, and Joseph Sharit. 2009. *Designing for older adults: Principles and creative human factors approaches*. CRC press.

[21] C. Ailie Fraser, Mira Dontcheva, Holger Winnemöller, Sheryl Ehrlich, and Scott Klemmer. 2016. DiscoverySpace: Suggesting Actions in Complex Software. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 1221–1232. DOI: http://dx.doi.org/10.1145/2901790.2901849

[22] Susan L. Gatto and Sunghee H. Tak. 2008. Computer, Internet, and E-mail Use Among Older Adults: Benefits and Barriers. *Educational Gerontology* 34, 9 (2008), 800–811. DOI: http://dx.doi.org/10.1080/03601270802243697

[23] Tovi Grossman, George Fitzmaurice, and Ramtin Attar. 2009. A Survey of Software Learnability: Metrics, Methodologies and Guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 649–658. DOI: http://dx.doi.org/10.1145/1518701.1518803

[24] Jamye M Hickman, Wendy A Rogers, and Arthur D Fisk. 2007. Training older adults to use new technology. *The journals of gerontology. Series B, Psychological sciences and social sciences* 62 Spec No 1 (Jun 2007), 77–84.

[25] Idris His and Colin Potts. 2000. Studying the Evolution and Enhancement of Software Features. In *Proceedings of the International Conference on Software Maintenance (ICSM'00) (ICSM '00)*. IEEE Computer Society, Washington, DC, USA, 143–. http://dl.acm.org/citation.cfm?id=850948.853409

[26] Alexis Hope, Ted Schwaba, and Anne Marie Piper. 2014. Understanding Digital and Material Social Communications for Older Adults. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3903–3912. DOI: http://dx.doi.org/10.1145/2556288.2557133

[27] Nathaniel Hudson, Benjamin Lafreniere, Parmit K. Chilana, and Tovi Grossman. 2018. Investigating How Online Help and Learning Resources Support Children's Use of 3D Design Software. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 257, 14 pages. DOI: http://dx.doi.org/10.1145/3173574.3173831

[28] Jeff Johnson. 2010. *Designing with the Mind in Mind: Simple Guide to Understanding User Interface Design Rules*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[29] Matleena Kähkönen and Saila Ovaska. 2006. Initial Observations on Children and Online Instructions. In *Proceedings of the 2006 Conference on Interaction Design and Children (IDC '06)*. ACM, New York, NY, USA, 93–96. DOI: http://dx.doi.org/10.1145/1139073.1139098

[30] Leah Kaufman and Brad Weed. 1998. Too Much of a Good Thing?: Identifying and Resolving Bloat in the User Interface. *SIGCHI Bull.* 30, 4 (Oct. 1998), 46–47. DOI:http://dx.doi.org/10.1145/310307.310370

[31] Simeon Keates and Shari Trewin. 2005. Effect of Age and Parkinson's Disease on Cursor Positioning Using a Mouse. In *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '05)*. ACM, New York, NY, USA, 68–75. DOI: http://dx.doi.org/10.1145/1090785.1090800

[32] Kimia Kiani, George Cui, Andrea Bunt, Joanna McGrenere, and Parmit K. Chilana. 2019. Beyond "One-Size-Fits-All": Understanding the Diversity in How Software Newcomers Discover and Make Use of Help Resources. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 340, 14 pages. DOI:http://dx.doi.org/10.1145/3290605.3300570

[33] Bran Knowles and Vicki L. Hanson. 2018a. Older Adults' Deployment of 'Distrust'. *ACM Trans. Comput.-Hum. Interact.* 25, 4, Article 21 (Aug. 2018), 25 pages. DOI:http://dx.doi.org/10.1145/3196490

[34] Bran Knowles and Vicki L. Hanson. 2018b. The Wisdom of Older Technology (Non)Users. *Commun. ACM* 61, 3 (Feb. 2018), 72–77. DOI: http://dx.doi.org/10.1145/3179995

[35] Benjamin Lafreniere, Andrea Bunt, and Michael Terry. 2014. Task-centric Interfaces for Feature-rich Software. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design (OzCHI '14)*. ACM, New York, NY, USA, 49–58. DOI: http://dx.doi.org/10.1145/2686612.2686620

[36] Rock Leung, Charlotte Tang, Shathel Haddad, Joanna Mcgrenere, Peter Graf, and Vilia Ingriany. 2012. How Older Adults Learn to Use Mobile Devices: Survey and Field Investigations. *ACM Trans. Access. Comput.* 4, 3, Article 11 (Dec. 2012), 33 pages. DOI: http://dx.doi.org/10.1145/2399193.2399195

[37] Abdullah Al Mahmud, Omar Mubin, Suleman Shahid, and Jean-Bernard Martens. 2010. Designing social games for children and older adults: Two related case studies. (Sep 2010). https://www.sciencedirect.com/science/article/pii/S1875952110000066

[38] Sylvain Malacria, Gilles Bailly, Joel Harrison, Andy Cockburn, and Carl Gutwin. 2013. Promoting Hotkey Use Through Rehearsal with ExposeHK. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 573–582. DOI: http://dx.doi.org/10.1145/2470654.2470735

[39] Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2013. Patina: Dynamic Heatmaps for Visualizing Application Usage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 3227–3236. DOI:http://dx.doi.org/10.1145/2470654.2466442

[40] Joanna McGrenere and Gale Moore. 2000. Are We All in the Same "Bloat"?. In *Proceedings of the Graphics Interface 2000 Conference, May 15-17, 2000, Montr'eal, Qu'ebec, Canada*. 187–196. http://graphicsinterface.org/wp-content/uploads/gi2000-25.pdf

[41] Katherine McKnight, Kimberly O'Malley, Roxanne Ruzic, Maria Kelly Horsley, John J. Franey, and Katherine Bassett. 2016. Teaching in a Digital Age: How Educators Use Technology to Improve Student Learning. *Journal of Research on Technology in Education* 48, 3 (2016), 194–211. DOI: http://dx.doi.org/10.1080/15391523.2016.1175856

[42] Brad A. Myers, Ashley Lai, Tam Minh Le, YoungSeok Yoon, Andrew Faulring, and Joel Brandt. 2015. Selective Undo Support for Painting Applications. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 4227–4236. DOI: http://dx.doi.org/10.1145/2702123.2702543

[43] Jakob Nielsen and Rolf Molich. 1994. User Interface Design Guidelines: 10 Rules of Thumb. (Apr 1994). https://www.interaction-design.org/literature/article/user-interface-design-guidelines-10-rules-of-thumb

[44] Jackson Noel. 2019. Why user onboarding is the most important part of the customer journey. (Nov 2019). https://www.appcues.com/blog/user-onboarding-customer-journey

[45] Donald A. Norman and Stephen W. Draper. 1986. *User Centered System Design; New Perspectives on Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

[46] Chris Norval, John L. Arnott, and Vicki L. Hanson. 2014. What's on Your Mind?: Investigating Recommendations for Inclusive Social Networking and Older Adults. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3923–3932. DOI: http://dx.doi.org/10.1145/2556288.2556992

[47] David G. Novick, Oscar D. Andrade, and Nathaniel Bean. 2009. The Micro-structure of Use of Help. In *Proceedings of the 27th ACM International Conference on Design of Communication (SIGDOC '09)*. ACM, New York, NY, USA, 97–104. DOI: http://dx.doi.org/10.1145/1621995.1622014

[48] Marita A. O'brien, Wendy A. Rogers, and Arthur D. Fisk. 2012. Understanding Age and Technology Experience Differences in Use of Prior Knowledge for Everyday Technology Interactions. *ACM Trans. Access. Comput.* 4, 2, Article 9 (March 2012), 27 pages. DOI: http://dx.doi.org/10.1145/2141943.2141947

[49] François Portet, Michel Vacher, Caroline Golanski, Camille Roux, and Brigitte Meillon. 2013. Design and Evaluation of a Smart Home Voice Interface for the Elderly: Acceptability and Objection Aspects. *Personal Ubiquitous Comput.* 17, 1 (Jan. 2013), 127–144. DOI: http://dx.doi.org/10.1007/s00779-011-0470-5

[50] Mark Rice, Wah Pheow Tan, Jeremy Ong, Lih Jie Yau, Marcus Wan, and Jamie Ng. 2013. The Dynamics of Younger and Older Adult's Paired Behavior when Playing an Interactive Silhouette Game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1081–1090. DOI: http://dx.doi.org/10.1145/2470654.2466138

[51] John Rieman. 1996. A Field Study of Exploratory Learning Strategies. *ACM Trans. Comput.-Hum. Interact.* 3, 3 (Sept. 1996), 189–218. DOI: http://dx.doi.org/10.1145/234526.234527

[52] Wendy A. Rogers, Arthur D. Fisk, Sherry E. Mead, Neff Walker, and Elizabeth Fraser Cabrera. 1996. Training Older Adults to Use Automatic Teller Machines. *Human Factors* 38, 3 (1996), 425–433. DOI: http://dx.doi.org/10.1518/001872096778701935 PMID: 8865767.

[53] Ranu Sewdas, Astrid de Wind, Lennart G.L. van der Zwaan, Wieke E van der Borg, Romy Steenbeek, Allard J van der Beek, and Cécile R. L. Boot. 2016. Why older workers work beyond the retirement age: a qualitative study. In *BMC Public Health*.

[54] Ben Shneiderman. 1997. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (3rd ed.). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

[55] Nikita Soni, Aishat Aloba, Kristen S. Morga, Pamela J. Wisniewski, and Lisa Anthony. 2019. A Framework of Touchscreen Interaction Design Recommendations for Children (TIDRC): Characterizing the Gap Between Research Evidence and Design Practice. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children (IDC '19)*. ACM, New York, NY, USA, 419–431. DOI: http://dx.doi.org/10.1145/3311927.3323149

[56] Craig D. Stewart, Vicki L. Hanson, and Thomas J. Nind. 2014. Assisting Older Adults in Assessing the Reliability of Health-related Websites. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 2611–2616. DOI:http://dx.doi.org/10.1145/2559206.2581243

[57] Carol-Ina Trudel and Stephen J. Payne. 1995. Reflection and goal management in exploratory learning. *International Journal of Human-Computer Studies* 42, 3 (1995), 307 – 339. DOI:http://dx.doi.org/https://doi.org/10.1006/ijhc.1995.1015

[58] Laton Vermette, Joanna McGrenere, Colin Birge, Adam Kelly, and Parmit K. Chilana. 2019. Freedom to Personalize My Digital Classroom: Understanding Teachers' Practices and Motivations. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 318, 14 pages. DOI: http://dx.doi.org/10.1145/3290605.3300548

[59] Nicole Wagner, Khaled Hassanein, and Milena Head. 2010. Review: Computer Use by Older Adults: A Multi-disciplinary Review. *Comput. Hum. Behav.* 26, 5 (Sept. 2010), 870–882. DOI: http://dx.doi.org/10.1016/j.chb.2010.03.029

[60] Young Seok Yoon and Brad A. Myers. 2015. Supporting Selective Undo in a Code Editor. In *Proceedings of the 37th International Conference on Software Engineering - Volume 1 (ICSE '15)*. IEEE Press, Piscataway, NJ, USA, 223–233. http://dl.acm.org/citation.cfm?id=2818754.2818784

[61] Mary Zajicek. 2001. Interface Design for Older Adults. In *Proceedings of the 2001 EC/NSF Workshop on Universal Accessibility of Ubiquitous Computing: Providing for the Elderly (WUAUC'01)*. ACM, New York, NY, USA, 60–65. DOI: http://dx.doi.org/10.1145/564526.564543

[62] Fan Zhang. 2018. *Intergenerational Play Between Young People and Old Family Members: Patterns, Benefits, and Challenges*. 581–593. DOI: http://dx.doi.org/10.1007/978-3-319-92034-4_44