The Effect of Folder Structure on Personal File Navigation

Ofer Bergman
Department of Information Science, Bar-Ilan University, Ramat Gan, 52900 Israel. E-mail: oferbergman@gmail.com

Steve Whittaker
IBM Research, IBM Almaden Research Center, San Jose, CA. E-mail: sjwhitta@us.ibm.com

Mark Sanderson
Department of Information Studies, Sheffield University, Regent Court, 211 Portobello St, Sheffield, S1 4DP, UK. E-mail: m.sanderson@sheffield.ac.uk

Rafi Nachmias
Department of Education, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel. E-mail: nachmias@post.tau.ac.il

Anand Ramamoorthy
Department of Experimental Psychology, Universiteit Ghent 9000, Ghent, Belgium. E-mail: Anand.Ramamoorthy@Ugent.be

Folder navigation is the main way that personal computer users retrieve their own files. People dedicate considerable time to creating systematic structures to facilitate such retrieval. Despite the prevalence of both manual organization and navigation, there is very little systematic data about how people actually carry out navigation, or about the relation between organization structure and retrieval parameters. The aims of our research were therefore to study users' folder structure, personal file navigation, and the relations between them. We asked 296 participants to retrieve 1,131 of their active files and analyzed each of the 5,035 navigation steps in these retrievals. Folder structures were found to be *shallow* (files were retrieved from mean depth of 2.86 folders), with *small folders* (a mean of 11.82 files per folder) containing many *subfolders* ($M = 10.64$). Navigation was largely successful and efficient with participants successfully accessing 94% of their files and taking 14.76 seconds to do this on average. Retrieval time and success depended on folder size and depth. We therefore found the users' decision to avoid both deep structure and large folders to be adaptive. Finally, we used a predictive model to formulate the effect of folder depth and folder size on retrieval time, and suggested an optimization point in this trade-off.

Received March 29, 2010; revised July 6, 2010; accepted July 6, 2010

 • Published online 17 August 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.21415

Personal file navigation (*navigation* herein) is a two-phase process. First, users manually traverse their organizational hierarchy until they reach the folder in which the target file is stored. Second, they locate the file within that folder (Bergman, Beyth-Marom, Nachmias, Gradovitch, & Whittaker, 2008).

Most information retrieval research has focused on public data sources such as databases, libraries, and the Web, developing various theories and methods for organizing and retrieving such public information. Yet all of us expend considerable effort organizing and accessing our personal information, using predominantly manual methods to prepare for subsequent retrieval. Surprisingly little is known about this process, in terms of how successful people are at organizing and retrieving their personal data.

This study therefore attempts to empirically investigate various questions relating to navigational retrieval, personal folder organization, and the relationship between them. We present large-scale quantitative data about (a) participants' folder structure and organizational strategies, (b) navigation success and efficiency, and (c) the effects of folder structure on retrieval success and efficiency. In contrast to previous research that focused on file structure alone, our study also quantitatively investigated file navigation retrieval in a natural setting, and examined the effect of structure on folder navigation.

There has been some prior research on how people organize their personal information. Early studies looked at the organization of personal paper archives (Malone, 1983; Whittaker & Hirschberg, 2001) finding two prevalent strategies: filing and piling. Because of the characteristics of filing cabinets and folders, early studies found only few instances of complex subfoldering of paper archives (Cole, 1981). More recent work has documented organizational strategies across different types of digital data, detailing how people organize e-mails (Whittaker & Sidner, 1996), Web data (Abrams, Baecker, & Chignell, 1998; Tauscher & Greenberg, 1997), photos (Kirk, Sellen, Rother, & Wood, 2006), documents (Gonçalves & Jorge, 2003; Hardof-Jaffe, Hershkovitz, Abu-Kishk, Bergman, & Nachmias, 2009a, 2009b Henderson & Srinivasan, 2009; Jones, Phuwanartnurak, Gill, & Bruce, 2005) or common strategies across all these data types (Bergman, Beyth-Marom, & Nachmias, 2006, 2008; Boardman & Sasse, 2004). Although such studies have looked at how we manually *organize* personal information, less attention has been paid to how people exploit these structures to *access* that information.

Some recent studies document the problems people experience with organizing personal information. People find it hard to organize e-mails, making folders that are either too big or too small (Fisher, Brush, Gleave, & Smith, 2006; Whittaker & Sidner, 1996). For example, Whittaker and Sidner (1996) found that almost 40% of e-mail folders contain two or fewer items and Henderson and Srinivasan (2009) showed that 8% of file folders created are empty, showing that people create structures that they fail to actively exploit for organization. In contrast, with digital photos, people create large folder structures that contain heterogeneous pictures from many different events, making it hard to find older digital photos (Whittaker, Bergman, & Clough, 2009). Other studies show that Web bookmark folders are often not useful in supporting retrieval of Web documents (Abrams et al., 1998; Aula, Jhaveri, & Kaki, 2005; Tauscher & Greenberg, 1997). And when users are asked to explain their organization in PIM (personal information management) "desktop tours," they usually express dissatisfaction and modify their organization as they give the tour (Boardman & Sasse, 2004; Whittaker & Sidner, 1996).

One response to these organizational problems has been to propose a move to *desktop search.* Much novel desktop search technology has been developed over the last few years, e.g., Google Desktop, Microsoft Windows Search, and Macintosh Spotlight. According to its advocates, desktop search promises to minimize users' organizational problems because it reduces the need to manually organize personal information, which is automatically indexed by the search engine. Search has other potential advantages: It allows flexible and efficient ways to query one's personal information (Cutrell, Dumais, & Teevan, 2006; Russell & Lawrence, 2007). Despite its promise, however, various studies still show a strong preference for navigation over search when both are available for accessing personal information (Barreau & Nardi, 1995; Boardman & Sasse, 2004; Kirk et al., 2006; Teevan, Alvarado, Ackerman, & Karger, 2004). Moreover, the use of improved search engines has been shown to have little effect on this preference (Bergman et al., 2008). Bergman et al. (2008) showed that regardless of search engine quality, there was a strong preference for navigation. Search was predominantly used as a last resort only when users could not remember the location of a file. There was also little evidence that using improved desktop search leads people to change their filing habits to become less reliant on hierarchical file organization.

It therefore seems that (at least for the foreseeable future) manual file organization and navigation will be critical PIM behaviors. This study therefore attempts to explore and quantify various research questions relating to three topics: folder structure, navigation performance, and the effect of structure on retrieval.

Folder Structure

There are important trade-offs to be made in organizing files and folders. Folder hierarchies may lie between two extremes: (a) broad and shallow, or (b) deep and narrow. Broad shallow hierarchies allow faster access to folders, but increase the time needed to scan within each folder. In contrast, deep narrow hierarchies allow faster scanning of each folder, but users have to access more folders overall. Previous work is inconclusive about which of these strategies people most commonly use.

In an early study, Barreau (1995) studied 7 participants using DOS, OS/2, Windows, and Macintosh operating systems. Only three of her participants used folders at all, the other four grouped their files simply by placing them on separate floppy disks. More recent studies have generated contradictory findings about the structure of personal file systems. Gonçalves and Jorge (2003) studied the folder structure of 11 computer scientists using Windows (8), Linux (2), and Solaris OS (1). Their results show extremely deep, narrow hierarchies. The average directory depth was found to be 8.45, with an average branching factor (which is an estimate of the mean number of subfolders per folder) of 1.84. In contrast, a larger scale study by Henderson and Srinivasan (2009) looked at the folder structure of 73 university employees using Windows OS. The structures they found were much shallower, being only 3.4 folders deep on average (similar results were obtained by Boardman & Sasse, 2004). Folders tended to be broader with an average of 4.1 subfolders per folder, for non-leaf folders. Both studies found relatively small numbers of files per folder: 13 for Gonçalves and Jorge (2003) and 11.1 for Henderson and Srinivasan (2009).

DOI: 10.1002/asi

However, one significant limitation of the above studies is that they examine the user's *entire folder archive*, which may contain thousands of inactive files in archival structures that have not been touched for years. For example, Gonçalves and Jorge (2003) document that *over half* the files in the users' system had not been modified for over a year. Instead, our study focused on *active* parts of the structure from which the user had recently retrieved files. Other work has documented a strong tendency to access *recent* personal information (Bergman et al., 2008, Dumais et al., 2003, Tang et al., 2008), and we wanted to focus on these more typical access situations.

Our study investigated the following research questions regarding folder structure:

1.1 Depth: At what depth in the folder hierarchy are active files stored? Are they stored in deep structures as found in Gonçalves and Jorge (2003) or shallow ones as in Henderson and Srinivasan (2009)?
1.2 Size: How big are file folders?
1.3 Internal Structure: How many subfolders and files are in each folder?
1.4 Relations between Structure and Depth: Does folder depth affect folder size, number of subfolders, and percentage of subfolders?
1.5 Subfoldering Distribution: What percentage of each folder is taken up with subfolders? How is this subfolder percentage distributed across all folder items? What explains this distribution?

Navigation Success

Prior research has consistently shown that navigation is the main way in which users retrieve their files (Bergman et al., 2008; Boardman & Sasse, 2004; Kirk et al., 2006; Teevan et al., 2004). However, no prior research has quantified *how* people actually navigate to their folders in their natural setting. Our study examined *retrieval success rate* and the *time* users took to navigate to their files. We can interpret these results in terms of users' memory of their file locations.

Our research questions for navigation success were as follows:

2.1 Success Rate: How often are participants successful in retrieving their files?
2.2 Factors Affecting Success: We collected information about retrieval strategies. We examined the *number of retrieval steps*, i.e., the number of times a user opened a new folder, as well as *step duration*—the time taken to scan each folder. We asked the following questions: What is the distribution of retrieval outcome? How does retrieval outcome relate to retrieval time, number of steps per retrieval and step duration? And what do these results imply for users' memory for file location?

The Effects of Structure on Retrieval

Although prior work has documented different organizational strategies, it hasn't examined *the effect of these strategies on retrieval*. It seems, however, that there are trade-offs in how users choose to organize their information. Broad shallow hierarchies reduce the number of folders to be scanned, but increase the time to scan the contents of each folder. In contrast, narrow, deep hierarchies reduce scan time per folder, but mean that more folders have to be accessed overall.

Although the effect of structure on retrieval has not been examined for personal files, it has been studied extensively for menu navigation (Jacko & Salvendy, 1996; Kiger, 1984; Miller, 1981; Snowberry, Parkinson, & Sisson, 1984) and for Web page navigation (Furnas, 1997; Kim, Li, Moy, & Ni, 2001; Larson & Czerwinski, 1998; Shneiderman & Plaisant, 2010; Zaphiris & Mtei, 1997; Zhang, Zhu, & Greenwood, 2004). Overall, breadth is better than depth in terms of both error rate and retrieval time, i.e., choosing broad shallow hierarchies leads to more effective retrieval. For example, Miller (1981) tested four artificial menu structures with 64 bottom level nodes: 2^6 (six levels of depth with two items of breadth), 4^3 (three levels of depth each with four items of breadth), 8^2 (two levels of depth with eight items of breath) and 64^1 (64 top level items). Of the four structures, the 8^2 supported fastest retrieval and lowest error rate. These results suggest that some hierarchical organization reduces the visual overcrowding found in the 64^1 structure; however, deep structures should also be avoided. Indeed, later studies (which did not test the "no hierarchy" option) found that retrieval time is positively correlated with depth for both menus and Web pages (Furnas, 1997; Jacko & Salvendy, 1996; Kiger, 1984; Kim et al., 2001; Zaphiris & Mtei, 1997). For Web design, a widely quoted heuristic for navigation design is the "three clicks rule," which states that the user should be able to get from the homepage to any other page on the site within three mouse clicks, arguing for shallow organizational structure (Zhang, Zhu, & Greenwood, 2004).

Our research questions for the effect of structure on retrieval were as follows:

3.1 Folder Depth and Retrieval Time: Does folder depth affect retrieval time?
3.2 Folder Size and Retrieval Time: Does folder size affect step duration and retrieval time?
3.3 Folder Size, Folder Depth and Success: Do structural elements (folder size and depth) affect retrieval success?
3.4 Predictive Modeling: How do folder depth and size predict retrieval time?

Method

Previous work examined organizational strategies in relatively small numbers of participants. In contrast, in our study, to increase external validity, we collected data from large numbers of users sampled in a naturalistic setting. The requirement for lightweight, nonintrusive data collection led us to a procedure in which we recruited users and videotaped their screens as they accessed files from their own computers. We did not install software on people's machines

DOI: 10.1002/asi

to record organization and retrieval behaviors. Installation is error prone, and pilot interviews showed that users were concerned about its intrusiveness and potential implications for their privacy.

Other studies have tried to profile people's entire document collections (Gonçalves & Jorge, 2003; Henderson & Srinivasan, 2009; Tang et al., 2007). However, this runs the risk of cataloguing large numbers of documents that may not have been accessed for very long periods. Instead, we wanted to look at typical access behaviors. Other research shows that users tend to most frequently access recent information items regardless of whether these are files, Web pages, or e-mails (Bergman et al., 2008; Dumais et al., 2003; Tang et al., 2008). We therefore videoed participants navigating to files in their Recent Documents list, i.e., personal files that they had recently spontaneously retrieved and opened from their own computers, as part of their everyday computer use. There were a number of other important benefits to this approach. Focusing on recent files meant users were trying to access files that we were confident were present on users' disks and that were definitely retrievable by the user. It also allowed us to identify active files without having to manipulate or access participants' file systems, avoiding encroaching on their privacy.

Participants

Participants were 296 everyday computer users: 163 men, 133 women. The large majority of participants were students and employees at Sheffield University. The participants were directly approached by the researchers in the university and students' hall of residence (nonrandom selection). We knocked on their doors and asked them to spare a few minutes for the study. Participants' ages ranged from 16 to 64 years ($M = 26.44$, $SD = 9.63$). The majority of participants were Windows OS users (246: 181 XP, 62 Vista, 3 Windows 2000), 43 used a Mac, and 7 used a Linux operating system.

Procedure

Participants used their own computers for the retrieval task. The tester printed out the participants' Recent Documents list, asking them to navigate to each file (the target) in that list in order. Participants were asked to click on the target file once but not open it. We did this to preserve users' privacy as these files might contain sensitive information. Participants were asked to close all folders before each navigation task took place, so that all retrievals started from the desktop. Participants were asked to skip a file in the list when they had already navigated to that target folder during a previous access task. We did this to prevent access to these items being primed because that folder had already been accessed. We asked our participants to access only files saved on their computer and to avoid retrieving files on external drives (such as a memory stick) and e-mail attachments that had not been saved as files on their hard drive. The procedure took approximately 10 minutes.

Retrievals

Our study includes 1,131 valid retrievals. Of the initial overall set of 1,158 recorded retrievals, we excluded 2% that were deemed invalid for the following reasons: 15 retrievals were interrupted by external events such as phone calls or instant messenger alerts. In a further six retrievals, participants did not follow the above procedure (e.g., they moved the mouse-pointer over the Recent Documents list to look up the file's path instead of using the printout); 3 participants used a library computer so the Recent Documents list did not contain any of their personal files; for 2, the video recording was not clear enough to be analyzed, and 1 participant had deleted all files on the list prior to the experiment.

The target files of these retrievals were in various formats: 469 text files (e.g., doc files), 160 pictures (e.g., jpg files), 126 pdf files, 64 Excel files, 49 MP3 files, 40 PowerPoint files, 28 video clips (e.g., avi files), 16 SPSS files, 14 html files, 48 files in unidentified format, and 117 files in other, less common formats.

Retrieval Time Measurements

Recordings of user interactions were made using a high-definition digital video camera (with 1080 horizontal scan lines). This was sufficient resolution to allow the user interaction to be timed accurately, with text on screen being readable by our analysts almost all the time.

We measured retrieval time by analyzing the videos frame-by-frame. In a pilot, it was found that in the camera's default setting, frames were not of equal duration, making timing calculations very complex. This problem was resolved by adjusting the camera so that frames were recorded at a fixed rate of 25 frames per second, making each frame 40 milliseconds (0.04 seconds) long.

Retrieval time. Retrieval time was measured from the first mouse movement made by a participant in the navigation, until the moment when they either clicked on the target file (in successful retrievals) or announced that they could not find it (in failure retrievals).

Step duration. We use the term *step* for each folder opened in the navigation process. In our study, we measured 5,035 steps. Step duration was measured from the time a folder was opened until the time the user either (a) clicked on the next folder, (b) reverted to a parent folder (if the relevant item was not found), (c) clicked on the target file, or (d) said, "I give up." We excluded the time taken from clicking on a folder to that folder's opening, as pilots showed that this time was inconsistent across different computers depending on their configuration and performance. Because of this correction, the total time for aggregated steps is slightly shorter than the overall retrieval time.

Research Limitations

As users, we are very oriented to the semantics of our files. We organize files and give files and folders names based

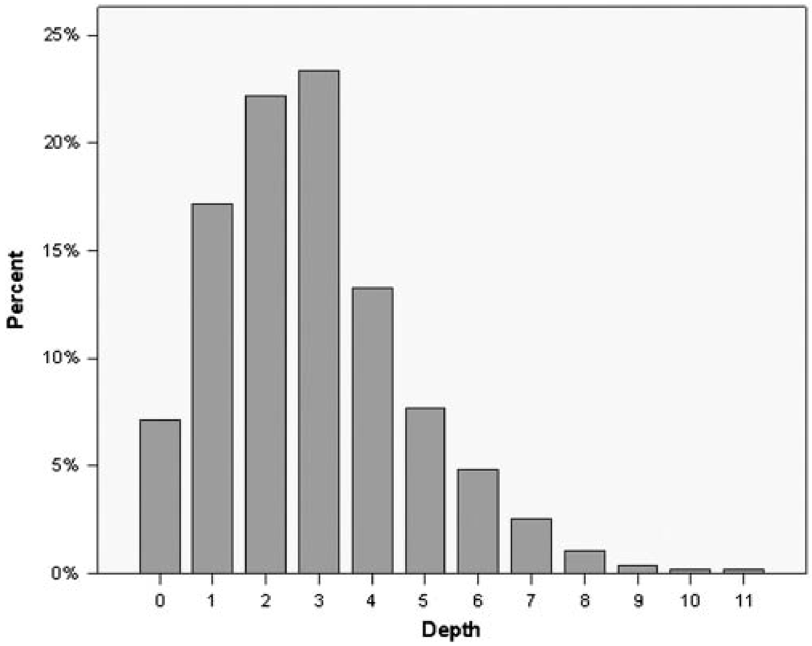

FIG. 1. Frequency distribution of depth of target file.

on their intrinsic meaning. Semantics undoubtedly affects navigation success and retrieval time. However, our research focuses on *structural* rather than semantic elements and their effect on retrieval. Each person's semantic organization is highly individual (Boardman & Sasse, 2004), making it hard to compare the effects of semantics across individuals. Evaluating these effects was beyond the scope of this research and should be addressed in future work.

Results

Folder Structure

The video recordings provide information about the users' folder organization strategies. We were able to collect information about the organization of the folders that participants accessed as they navigated to the target file. In this section, we describe properties of the hierarchical structure, such as folder depth, size, and breadth (number of subfolders).

1.1 At what depth in the folder hierarchy are active files stored? Folder depth is the number of steps in the folder path that the participants traversed to get directly to the folder containing the target file. The folder depth of the desktop is 0, and the folder depth of the root folder (e.g., My Documents) is 1. Figure 1 presents a frequency distribution of the depth of the target file for 1,054 successful retrievals. We obviously could not determine the depth when users failed to find the target. We excluded eight additional retrievals because the recordings were not clear enough. We treated shortcuts in different ways depending on whether we were analyzing folder structure or retrieval. In the current section we are interested in folder structure, when participants used a shortcut for access, we identified the depth of the file rather than of the shortcut, as we were interested in the location and context of where the file was logically stored. In subsection 3.1, we describe how we treat shortcuts in retrieval context. Full numerical results are given in the Appendix, Table A1, second column.

The mean folder depth of the target was 2.86 ($SD = 1.85$). The median folder depth was 3. Furthermore, the majority of retrieved files (82%) were stored at depths of 4 or less (see Figure 1). This is in clear contrast to previous studies that report overall depths of 8.45 for entire archives (Gonçalves & Jorge, 2003). There was also considerable use of the desktop: in 115 retrievals (11% of all retrievals) participants used a desktop folder shortcut and in 75 (7%) retrievals they used files placed on the desktop.

A possible explanation for the shallow hierarchical position of active files is that people rely on default locations (such as My Documents and My Pictures). However, only 136 retrievals (12%) were made from such default storage locations (e.g., files retrieved directly from My Documents folder, as opposed to subfolders inside it). The default location folders used in these retrievals contained an average of 19.42 files on the average ($SD = 37.28$). This clearly indicates that these folders are not large enough to serve as the users' only file repository. Lack of reliance on defaults implies that the majority of participants made efforts to construct their own organizational hierarchies rather than relying on placement by the application.

These findings inform us about hierarchical depth of the folder containing the target file; in the next subsections (1.2–1.5), we report on folders at each individual step in the retrieval process. Whenever our results include depth we report only *Direct Navigation* retrievals: retrievals in which the user went directly to the target file, without making mistakes by accessing irrelevant folders. In this case,

DOI: 10.1002/asi

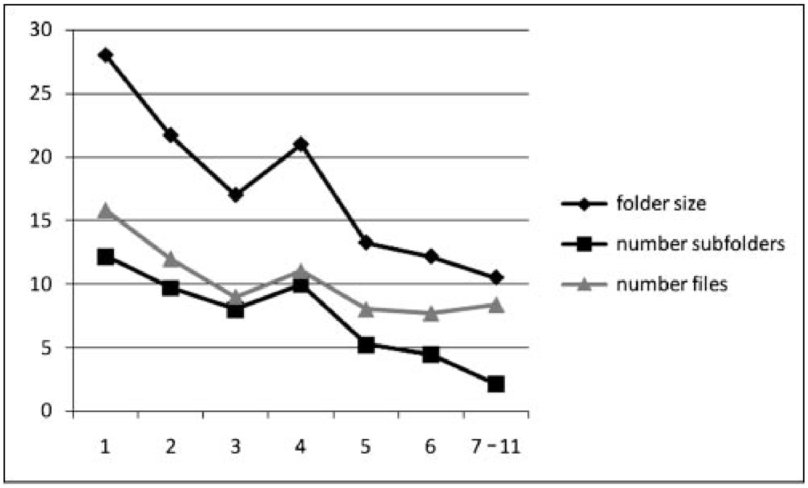

FIG. 2. Folder properties at different depths.

the hierarchical depth of each folder along the path was consistent with the step number (each step increases the depth, except the last one from which the file is retrieved). We omitted results regarding folder depth 0, as in the first step, participants did not navigate using a folder, but used either a menu (e.g., Start → My Documents) or the desktop instead.

1.2 How big were file folders? On average, folders that participants used in their navigation contained 22.46 information items (i.e., files and subfolders; $SD = 32.30$). The median folder size was much smaller: 15 information items. This difference between the average and median was due to a long tail of very big folders, some of which contained a large number of machine-generated files (e.g., picture folders populated by camera software or music folders managed by music software).

1.3 How many subfolders and files were in each folder? On average, participants' folders contained 10.64 subfolders ($SD = 23.54$) and 11.82 files ($SD = 27.47$). When calculating the average percentage of subfolders in relation to all information items (files and subfolders), we find that about half of the information items in the folders were subfolders ($M = 54\%$, $SD = 36\%$). This is again striking: instead of organizing information into a small number of folders containing huge numbers of files, the large number of subfolders suggests that users spend time and effort to create structure in their file system, in anticipation of future retrievals.

1.4 Does folder depth affect folder size, number of subfolders and percentage of subfolders? The average folder size at different depth levels is represented by the top diamond line in Figure 2 (for numerical values including standard deviations, see Table A2 in the Appendix). As is evident from the graph, there is a negative correlation between folder depth and folder size, with folders becoming smaller at greater depths (Pearson $r(2{,}248) = -0.13$, $p < 0.01$). A possible explanation for this is that deeper folders are added later than shallower ones, so participants have less time to populate them with files and subfolders. Alternatively, participants keep active files on higher levels to promote accessibility.

Figure 2 also shows the mean number of files (center triangle line) and subfolders (bottom square line). Both graphs seem to decay with depth at approximately the same rate. Although there is a small negative correlation between folder depth and percentage of subfolders, $r(2{,}642) = -0.06$, $p < 0.01$), each folder depth has an average of about 50% files and 50% subfolders (except for the more infrequent 7–11-level deep folders) as confirmed in Table A2.

This constant average percentage of subfolders disconfirms the common intuition that higher folder levels serve as structural aids and are populated mostly by folders, whereas deeper folder levels contain mostly files.

1.5 What is the distribution of the percentage of subfolders in all folder items? And what explains this distribution? A histogram of subfolder distribution of all information items in folders is presented in Figure 3.

Figure 3 clearly shows a bimodal distribution of subfolder percentages. Moreover, 32% of the folders contain either only files (331 folders—12 % of all folders measured) or only subfolders (521–20% of the folders). What explains this bimodal distribution? Why do some of the folders contain exclusively or mainly files, while other folders contain exclusively or mainly subfolders? The answer to this question is not in the folder structure: we found (in the previous section) that folder depth has little effect on subfolder percentage. The explanation relates to the difference between *Target Folders* (folders containing the target files) and those which are navigated through on the way to the target. Figure 4 divides Figure 3 into two histograms: *Target Folders* and *Navigation Folders* (folders that precede the target in the navigation path).

Figure 4 shows that *Target Folders* contained mostly files and are responsible for the "all files" peak in the bimodal distribution of Figure 3, while *Navigation Folders* contained mostly subfolders and are responsible for the "all subfolders" peak in the bimodal distribution. An independent

DOI: 10.1002/asi

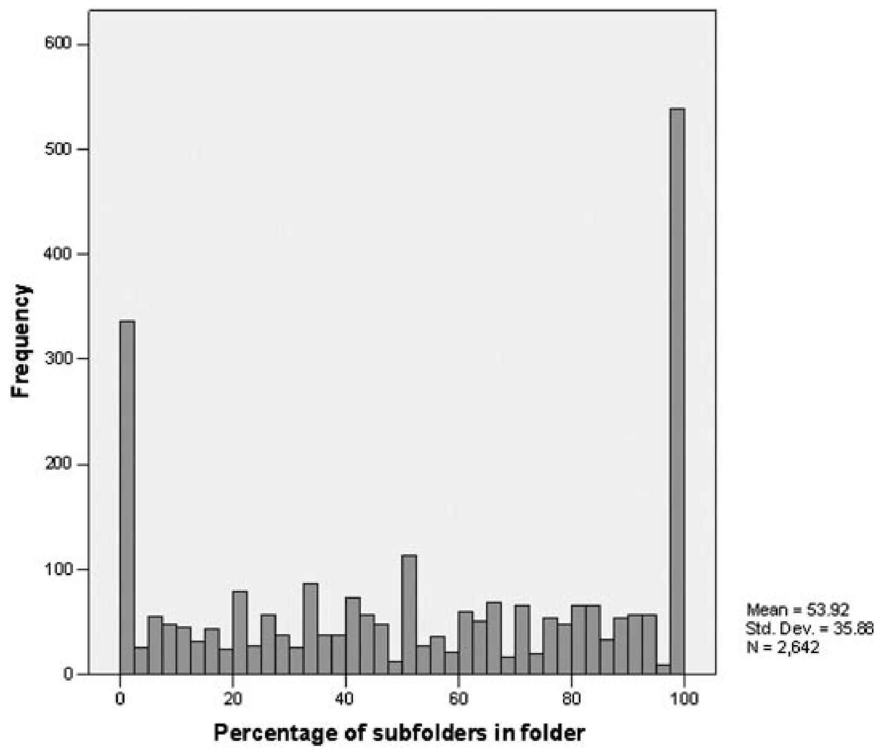

FIG. 3. Percentage of subfolders in folders.

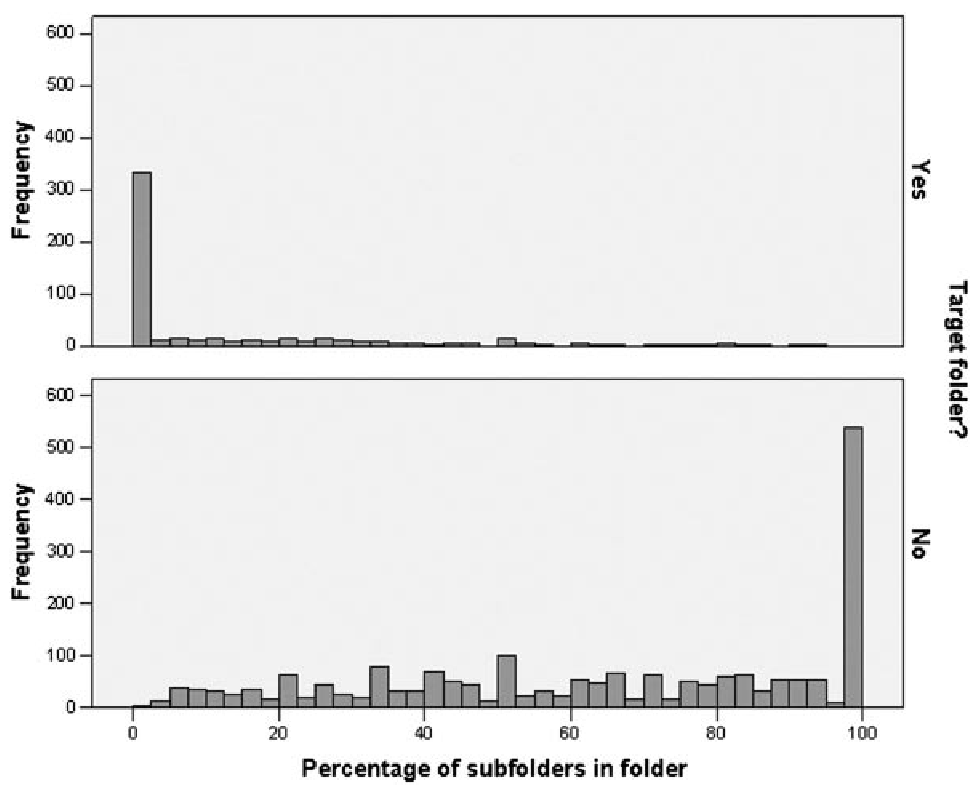

FIG. 4. The subfolders histogram divided into target and navigation folders.

sample t test shows that the subfolder percentage of *Target Folders* ($M = 13\%$, $SD = 22\%$) was significantly smaller than the subfolder percentage of the *Navigation Folders* ($M = 65\%$, $SD = 31\%$), $t(2, 641) = 37.52$, $p < 0.01$. The effect (a difference of 52% between averages) is large (however, notice that *Target Folders* still contained an average of 13% of subfolders and the *Navigation Folders* contained an average of 35% files).

DOI: 10.1002/asi

TABLE 1. Retrieval outcome and retrieval factors.

	Number of retrievals (Overall %)	Mean retrieval time (*SD*) (Seconds)	Mean steps per retrieval (*SD*) (Seconds)	Mean step duration (*SD*) (Seconds)
Direct success (DS)	893 (79%)	12.29 (9.29)	3.82 (1.84)	2.72 (3.23)
Eventual success (ES)	169 (15%)	27.82 (16.49)	7.2 (3.2)	3.48 (4.33)
Failure (F)	69 (6%)	45.05 (31.92)	5.77 (4.75)	7.28 (12.77)
Total	1,131 (100%)	16.61 (15.9)	4.44 (2.67)	3.27 (5.12)
Statistical results		DS < ES < F	DS < F < ES	DS < ES < F

Navigation Success

In this section, we study retrieval success. We also examine types of retrieval outcomes (i.e., whether or not retrieval was successful, as well as *types* of success). We relate outcome to various parameters (retrieval time, number of steps per retrieval, and step duration). We then use this data to shed some light on users' memory for file location.

2.1 How often did participants succeed in retrieving their files? Participants found 94% of their requested files (1,062 out of 1,131 files). The average time to navigate to these files was 14.76 seconds ($SD = 12.16$). They took an average of 4.44 steps, and each step took 3.27 seconds. This shows that for active files, participants are generally able to find their files quickly and accurately. We know that users tend to access recently used files (Bergman et al., 2008; Dumais et al., 2003, Tang et al., 2008), so success in navigating to active files may partially explain other findings that navigation is the preferred method for accessing files (Barreau & Nardi, 1995; Bergman et al., 2008; Capra & Pérez-Quiñones, 2005; Kirk et al., 2006; Teevan et al., 2004).

2.2 What is the distribution of retrieval outcome? How does retrieval outcome relate to retrieval time, number of steps per retrieval and step duration? And what do these results imply about users' memory for file location? Not all access attempts were immediately successful. We identified three different retrieval outcomes. In the majority of occasions (79%), participants navigated through the folder hierarchy directly to the target file's location without diversions or missteps. We refer to these as *direct successes*. On another 15% of occasions, they were eventually able to find the file, but en route they opened at least one incorrect folder and had to retrace their steps. We called these *eventual successes*. On the remaining 6% of occasions, participants attempted to find the file, but were unable to do so. We called these *failures*. Table 1 presents the distribution of retrieval outcome as well as retrieval time, number of steps per retrieval, and step duration for each type of retrieval. The statistics in the last row are t tests with a Bonferroni correction.

Retrieval outcome distribution. Although participants were directly successful in 79% of all navigations, navigating straight to the target file without error, in 21% cases (238 retrievals), they had difficulty remembering the location of the files. However, for 169 of these (71%), they were eventually able to find the file by navigation. This, too, may explain other findings of strong preferences for navigation over search (Bergman et al., 2008). In general, participants tend to remember the exact location of their active files, but even when they don't remember the exact location, they know that if they persist, navigation will usually be successful.

Relation between outcome and retrieval time. As we expected, retrieval time for *Direct Success* navigations was shorter than for *Eventual Success,* $t(1{,}060) = 17.21$, $p < 0.01$, which in turn is shorter than for *Failure,* $t(236) = 5.46$, $p < 0.01$. It should be noted that on 24 of the 69 *Failure* retrievals, participants said in advance that they did not remember the location of the file and did not attempt to navigate to it. As there was no navigation in these cases, we could not report on their retrieval time and omitted them from this calculation. Table 1 shows that the effect of retrieval outcome on retrieval time was large: Retrieval time almost doubles when we compare direct and eventual successes. It almost doubles again when people cannot find the file.

Relationship between retrieval outcome and number of steps per retrieval. As we expected, the number of steps per retrieval was greater in the *Eventual Success* than the *Direct Success* case, $t(1, 060) = 18.99$, $p < 0.01$. To our surprise however, the number of steps in the *Eventual Successes* ($M = 7.2$ steps) was greater than for *Failures* ($M = 5.77$ steps) $t(236) = 2.69$, $p < 0.01$. This result is counterintuitive because one would expect participants who cannot remember, to exhaustively search for the target, opening many folders before giving up. A possible explanation for the reduced number of *Failure* steps is that in the *Eventual Success* cases, participants had a correct intuition that they would eventually find the file, and consequently tried harder to find the file than for *Failure* retrievals where they gave up more easily.

Relation between retrieval outcome and step duration. Our results show that *Direct Success* step durations (i.e., the time taken to scan each folder) were shorter than *Eventual Success* step durations, $t(4, 627) = 6.47$, $p < 0.01$, which in turn were shorter than *Failure* step durations, $t(1, 611) = 8.91$, $p < 0.01$. In particular, there was

DOI: 10.1002/asi

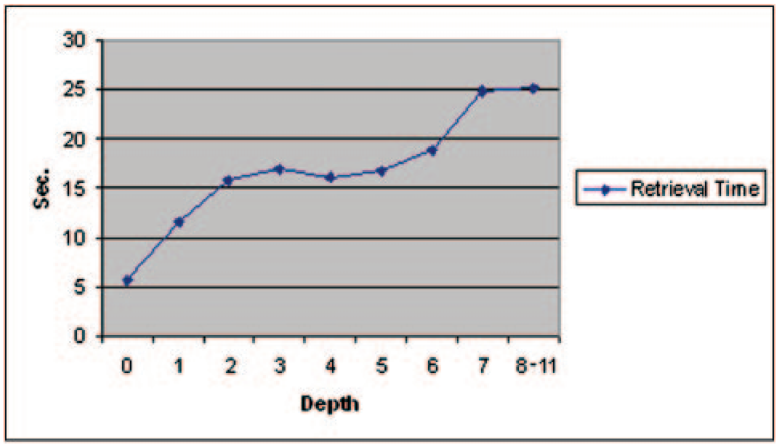

FIG. 5. Overall retrieval time at different depths.

a substantial difference in step duration between *Eventual Successes* (3.48 seconds) and *Failures* (7.28 seconds).

Access outcome and memory. The access outcome is a reflection of the users' memory for the target file location. For *Direct Successes*, participants navigated directly to the file, indicating that they remembered exactly where it was. For *Eventual Successes,* participants made at least one mistake during navigation, but eventually found the target, indicating they didn't remember exactly where the file was located. Finally, for *Failures,* participants did not find the file, indicating that they clearly didn't remember where it was. Our results thus reveal relations between memory and retrieval. In the *Direct Success* case, users are able to remember the file location and hence directly navigate to the file, through highly efficient steps, in which users are quickly able to select the target folder at each step. In contrast, in the *Failure* case, users seem unable to remember much about the file location and in consequence, when they open a folder they spend large amounts of time scanning files and subfolders to look for clues about where the target might be stored. There is a large difference in variance in step duration for *Failure* vs. *Direct Success* retrievals ($SD = 12.77$ seconds for *Failure* compared with 3.23 seconds for *Direct Success* navigations). This variance difference may arise in the following way: In the *Failure* case where different aspects of navigation may have very different time courses; participants may quickly navigate to a folder where they guess the file is located (leading to a short step duration). They then scan it exhaustively, but when they can't find the target, attempt to think of an alternative location before possibly giving up (leading to a long step duration). In contrast in the *Direct Success* case, participants know exactly where to go at each phase of the navigation leading to short, uniform steps of low variance. Finally, *Eventual Successes* are slightly longer per step than *Direct Successes*, but involve more steps overall (on average, participants made 3.1 mistaken steps for the *Eventual Success* retrievals $SD = 2.34$ steps). This suggests that on such occasions, users don't remember the exact location of the file, and look for it in more than one location before finding it.

Effects of Structure on Retrieval

In this section, we analyze the effect of folder depth and size on the speed and success of retrievals. Finally, we use a regression model to predict retrieval time by folder size and depth.

3.1 Does retrieval depth affect retrieval time? In contrast to the Folder Structure section, in our analysis of retrieval depth, folder shortcuts were analyzed as having a depth of 1 (as the target file is retrieved in two steps). Figure 5 indicates there is a positive correlation and a linear relation between retrieval depth and retrieval time, $r(1,054) = 0.29^{**}$, $p < 0.01$ (see also the third and fourth columns of Table A1 in the Appendix). The deeper the file, the more time it takes to find it: from an average of 5.6 seconds for desktop files to 25.22 seconds for files located 8–11 levels deep (the fact that the graph flattens at levels 4 and 5 can be explained by random observational errors caused by the dramatic drop in frequency of retrievals at these depth levels). As we expected, deeper hierarchies require more navigation steps and each step is an action that requires time.

3.2 Does folder size affect step duration and retrieval time? There is a positive correlation between the overall number of information items in each folder and the step duration, $r(3,971) = 0.24$, $p < 0.01$. The correlation for *Direct Success* retrievals between folder size and step duration is even higher, $r(2,688) = 0.31$, $p < 0.01$, presumably because this data does not contain missteps where step duration is influenced by users' attempts to remember the file location (see Figure 6). There was also a positive correlation for entire retrievals between the average number of information items in each retrieval path and overall retrieval time, $r(848) = 0.14$, $p < 0.01$. The more information items in a folder, the longer it takes for the participant to locate the correct file or folder.

DOI: 10.1002/asi

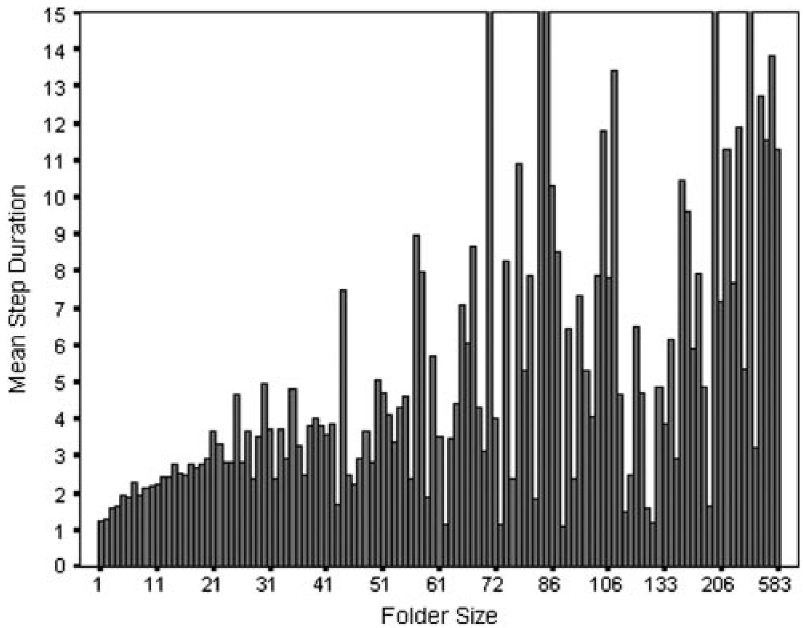

FIG. 6. Mean step durations at different folder sizes.

This again is consistent with cognitive work on visual search (Neisser, 1964; Treisman & Gelade, 1980).

3.3 Do structural elements (folder size and depth) affect retrieval success?

The effect of folder size on retrieval outcome. We tested the relations between folder size and retrieval outcome using *t* tests with a Bonferroni correction. We included in the analysis all folders that users touched in the course of their overall navigation. The mean folder size for *Direct Success* retrievals ($M = 22.72$, $SD = 32.30$) was smaller than for *Eventual Success* retrievals ($M = 28.38$, $SD = 83.92$), $t(3{,}650) = 2.94$, $p < 0.01$. The mean folder size for *Failure* retrievals ($M = 27.45$, $SD = 36.09$) was significantly larger than for *Direct Success* retrievals, $t(3{,}005) = 2.44$, $p < 0.02$, but not significantly different from the mean folder size for *Eventual Success* retrievals, $t(1{,}281) = 0.19$, $p > 0.05$. The folder size effect on retrieval outcome could therefore be explained either directly (it is easy to overlook the next information item in the navigation path in larger folders), or indirectly via memory (the bigger and more cluttered the folders are, the harder it is to remember where the file is located).

The effect of folder depth on retrieval success. The depth of *Direct Success* retrievals ($M = 2.81$, $SD = 1.81$) was lower than the depth of *Eventual Success* retrievals ($M = 3.10$, $SD = 2.05$), and this approached significance, $t(1, 052) = 1.81$, $p = 0.07$. This suggests that there are benefits for shallower organizational schemes. We cannot report on the depth of *Failure* retrievals, as these files were not found so we do not know the depth of their location.

TABLE 2. Regression model for retrieval time.

Factor	Coefficient	*SE*	*t*	*p*
Constant	4.956	0.71	6.945	<0.01
Depth	2.236	0.2	11.34	<0.01
Folder size	0.106	0.01	14.2	<0.01

3.4 How do folder depth and size predict retrieval time? To model the effect on overall retrieval time of target file folder depth and average folder size in the navigation path we used linear regression analysis. We first excluded folder shortcut data. The regression model is presented in Table 2. This model is significant ($R^2 = 0.22$, $p < 0.01$).

The predictive model presented in Table 2 is therefore:

$$\text{Retrieval Time} = 4.956 + 2.236 \times \text{Depth} + 0.106 \times \text{Size}$$

We will discuss the implications of this model at length in the discussion.

Discussion

Folder Structure

In this section, we discuss folder depth, overall size (including number of subfolders), the effect of depth on size, and the difference between target and navigation folders.

Folder depth. Active files were retrieved from an average depth of 2.86 folders. This suggests a shallow folder structure. This result is consistent with that of Henderson & Srinivasan

DOI: 10.1002/asi

(2009) who found an average folder depth of 3.4 and of (Boardman & Sasse, 2004) with an average folder depth of 3.3, but not with those of Gonçalves and Jorge (2003) who found an extremely high average folder depth of 8.45. A possible explanation for the contrast between the studies is that in the Gonçalves and Jorge (2003) study, the user population—a small number of computer scientists—may have storage behavior that is different from that of the majority of users.

On the other hand, our findings contrast with claims that participants are reluctant to organize their information, instead of saving it in rudimentarily organized structures (Cutrell et al., 2006; Dourish et al., 2000; Raskin, 2000; Russell & Lawrence, 2007). Only 12% of retrievals were made from default folders provided by the operating system such as My Documents, or other application-defined locations. In the other 88% of cases, files were retrieved from user-created folders. Moreover, these default location folders contained an average of 19.42 files, suggesting that they are only rarely used to store files. These results confirm previous studies that indicate that users are willing to invest time and effort in organizing their personal file collections (Bergman, 2006; Boardman, 2004).

Folder size. Our research found an average of 22.46 information items per folder ($SD = 32.30$). These numbers are bigger than those found in previous studies (Gonçalves & Jorge, 2003; Henderson & Srinivasan, 2009). A closer look at the results shows that the average number of files found in our research—11.82—is consistent with findings of Gonçalves and Jorge (2003)—13 files—and of Henderson and Srinivasan (2009)—11.1 files per folder. The difference is due to a difference in the number of subfolders, which we discuss next.

Breadth (no. of subfolders). Our research found an average of 10.64 subfolders per folder, compared to 4.1 subfolders found by Henderson and Srinivasan (2009) and a branching factor of 1.84 found in Gonçalves and Jorge (2003). This difference can be partly explained by differences in what was measured: Henderson and Srinivasan (2009) measured the average number of subfolders in the entire folder structure; we measured the average number of subfolders at each step of the retrievals. As each retrieval starts with top level folders (which tend to have a higher number of subfolders), the contribution of the top level folders in our calculation is greater than when computing the average number of subfolders for the entire folder structure. However, this does not explain all the differences between the results. When looking at the average number of subfolders at each depth (presented in the Appendix in Table A2, column 4), we see that our participants had slightly more subfolders than those of Henderson and Srinivasan (2009), and the number of subfolders decreased only gradually with folder depth. Our study therefore portrays a picture of a wider hierarchical tree than the ones reported in previous research.

Depth effect on size. Deeper folders tended to be smaller in size, presumably because they were newer and had less time to be populated. As reported in Henderson and Srinivasan (2009), we found that deeper folders contained fewer subfolders and fewer files. Interestingly, the relative numbers of files and subfolders in each folder remained steady regardless of folder depth: About half of a folder was populated with files and the other half with subfolders. These results contrast with the intuitive assumption that higher folder levels are populated mostly by folders and deeper folder levels mostly by files.

Target folders and navigational folders. Our results show a difference between (a) *Target Folders,* which contained mainly files (an average of 87% files and 13% subfolders); and (b) *Navigation Folders* used in the preceding steps of navigation, which contained significantly fewer files and more subfolders (35% files and 65% subfolders). These results (which explain the bimodal subfolder percentage distribution) indicate that users tend to make a clear distinction between two kinds of folders: some are used mostly as file repositories; others are "corridors" to navigate to these repositories. How can we explain why *Navigation Folders* still contain files? There are two independent explanations. First, subfolders are created gradually, in a bottom-up manner, as users observe that many of their files relate to the same topic (Jones, Phuwanartnurak, Gill, & Bruce, 2005). However, after these new subfolders are created, users may neglect to relocate older files into the relevant subfolders, both because this requires extra work and because these files are obsolete and therefore less likely to be retrieved. However, failing to remove older files is not adaptive because they compete for the users' attention and increase retrieval time (see the results of question 3.2). A second explanation is that users deliberately insert such target files in a higher hierarchical level because they assume that they are likely to be retrieved often. This is an adaptive behavior, because we found that files at higher levels of the hierarchy are retrieved faster and retrievals tend to be more successful (see the results of questions 3.1 and 3.3). Further research should explore these competing explanations for users' populating *Navigation Folders* with files.

Navigation Success

Our results show that participants were able to find 94% of the target files. Moreover, they seemed relatively efficient at accessing active files, taking, on average, 14.76 seconds. In the majority of cases, participants remembered where their files were: In 79% of the retrievals, participants navigated directly to the target file, in a further 15%, they eventually succeeded in finding the file. Only in 6% of the retrievals did they fail to retrieve the files. Because files were taken from the Recent Documents list, participants were probably familiar with their location. However, this pattern of accessing recent files reflects users' common naturalistic behaviors (Bergman et al., 2008; Dumais et al., 2003; Tang et al., 2008). Our results therefore indicate that users are generally able to navigate to active files quickly and accurately.

These results are consistent with our prior work on navigation and search. Our previous study (Bergman et al., 2008) analyzed the use of four different search engines. Overall, participants estimated that they remembered the exact location of their files in 74–90% of the retrievals. This is consistent with the 79% *Direct Success* retrievals found in this study. In that study, they also stated that they used a search engine for 4–13% of the retrievals—when they couldn't find their files by using navigation. This is consistent with the 6% *Failure* retrievals found in the current study. However, these estimations in Bergman et al. (2008) are based on memory and it is well known that people tend to remember evocative events (such as failing to find a file) much better than routine events (such as finding it). Future research could tackle this problem by using methods that do not rely on memory such as direct observation, logs, and diary studies.

More importantly, future research should compare hierarchical storage and navigation retrieval with alternative solutions. Articles written over two decades suggested three such directions for alternative solutions: (a) *Multiple Classification* allowing users to assign the information item to more than one category (e.g., tagging) (Lansdale, 1988; Malone, 1983); (b) *Automatic Classification*, which spares the user from having to manually classify the information (e.g., applying a predominant default classification parameter such as time) (Malone, 1983); and (c) *Search*, using any attribute that the user happens to remember about it, thus avoiding classification altogether (Lansdale, 1988). During these two decades, many new applications consistent with these directions have been developed, both experimentally and commercially. However, to date, there is no evidence that any of them is better than the existing hierarchical method. Our current results suggest that navigation is effective for active documents, providing an explanation for why users have not embraced search. Future research should systematically compare new alternative solutions with hierarchical navigation, with regard to parameters such as retrieval time, error rate and users' preferences. Stating that the hierarchical method is passé is simply not enough.

Showing that users are effective in accessing active documents supports previous work showing a preference for navigation over search (Barreau & Nardi, 1995; Bergman et al., 2008; Capra & Pérez-Quiñones, 2005; Kirk et al., 2006; Teevan et al., 2004). There may be profound reasons for this. Navigation in the physical environment has been the traditional way of finding items throughout millions of years of evolution (e.g., hunter—gatherers looking for food where they had previously stored it, or a dog digging for a bone where it hid it). As humans, we have well-developed cognitive and neurological structures that support navigation in physical locations and these may be used for computer folder navigation as well. This could be determined by future neuroscientific studies testing whether similar parts of the brain (such as the hippocampus) are activated in physical navigation and file navigation, determining whether the same mechanisms are involved.

Another possible reason for the success of navigation is the familiarity that users have with the structure of their own personal information. Personal information can be simply defined as "Stuff I've Seen" (Dumais et al., 2003), in which case users are likely to try to find it in the same location, using the same route as the previous times they saw it, with each navigation making the path more familiar. Files may be particularly familiar to users because users store and organize files in folders that they create according to their own subjective needs (Bergman, Beyth-Marom, & Nachmias, 2003; Jones et al., 2005). This is unlike previously seen Web pages where users rarely organize information (Jones et al., 2003). Users are naturally more likely to remember the classification and location they personally created, than an organization imposed by others. Possible cognitive explanations for file navigation preference can be found in Bergman et al. (2008).

The Effect of Structure on Retrieval

Our results show that both folder depth of the target file and the average size of the folders along the navigation path increase retrieval time. This is consistent with research on menu navigation (Jacko & Salvendy, 1996; Kiger, 1984; Miller, 1981) and Web navigation (Furnas, 1997; Kim et al., 2001; Larson & Czerwinski, 1998; Zaphiris & Mtei, 1997). The effect of depth can be explained by the fact that every step along the navigation takes its time for visual scanning, and cognitive and motor activity. The size effect is simply an instance of a well-known cognitive phenomenon: The time it takes to find a target visual stimulus is positively correlated with the number of other visual stimuli that distract the scanning (Neisser, 1964; Treisman & Gelade, 1980).

There is an obvious trade-off between depth and size. At one extreme, users can minimize the cost of retrieving deep in the hierarchy by storing all their items in a single folder; at the other extreme, users can create very deep hierarchies, reducing the size of their folders. Prior Web and menu navigation literature indicates that choosing either extreme of the trade-off increases retrieval time and the number of errors. But where is the "sweet spot" that minimizes retrieval time in this trade-off? We can use the predictive model derived from the regression presented in the results for question 3.5 to suggest such an optimization point in that trade-off. The predictive model is:

$$\text{Retrieval Time} = 4.956 + 2.236 \times \text{Depth} + 0.106 \times \text{Size}$$

According to the model, each additional folder step increases retrieval time by 2.236 seconds and each new information item in a folder increases retrieval time by 0.106 seconds. Therefore, the trade-off between depth and size is $2.236/0.106 = 21.09$. Each step down the hierarchy equals about 21 information items in terms of its effect on retrieval time. Therefore, as a heuristic, we can recommend that users try to avoid storing more than 21 information items per folder

DOI: 10.1002/asi

and create an additional level of subfolders instead. We call this the *up to 21 heuristic*. Interestingly, users seem to intuitively comply with this rule. Our study shows that mean folder size was found to be 22 information items and that 67.3% of the folders contained up to 21 items.

The file collection can grow in three different dimensions: in folder size, folder depth, and folder breadth (number of subfolders per folder). In the following paragraphs, we compare these three growth strategies.

Increasing folder size strategy. Miller's (1981) research has shown that creating a flat menu containing all 64 options slows down retrieval time and increases the number of mistakes over a two-level 8^2 menu. Storing thousands of files in a single folder[1] and finding them using navigation is simply not a realistic option. Indeed, our participants clearly did not choose to create huge folders as their median folder size was 15 items and the majority of their large personal file folders seemed to have been automatically created (e.g., camera, MP3 player) software. Our data showed small folders to be an adaptive behavior as we found a positive correlation between folder size and retrieval time. By keeping folders relatively small, participants avoided having many visual distracters that increase the time taken to find the target (Neisser, 1964; Treisman & Gelade, 1980). Our data also indicate that *Direct Success* retrievals had significantly smaller folders than *Failure* retrievals, indicating that larger folders increase error rate.

Increasing folder depth strategy. Research in menu and Web navigation has consistently shown that creating deep hierarchies increases retrieval time and error rates (Furnas, 1997; Jacko & Salvendy, 1996; Kiger, 1984; Kim et al., 2001; Miller, 1981; Zaphiris & Mtei, 1997). Our research showed a significant positive correlation between hierarchical depth of the target file and retrieval time, all arguing against creating deep folder structures. Interestingly, we found that users did not choose the deep hierarchy strategy, retrieving files from an average depth of 2.86 folders (i.e., between one and two levels below their main repository).

Increasing breadth (number of subfolders). Research in menu and Web navigation has shown that increasing breadth is preferred to increasing depth (Furnas, 1997; Jacko & Salvendy, 1996; Kiger, 1984; Kim et al., 2001; Miller, 1981; Zaphiris & Mtei, 1997). Our participants clearly chose the breadth option with an average of 10.64 subfolders per folder. Moreover, about half of the information items in folders were subfolders, regardless of the folder's hierarchical depth. It can be argued that increasing the number of subfolders increases folder size. This is true to some extent, but this increase is small compared to having all the files in these subfolders (and their subfolders, etc.) located in the original folder.

[1]Henderson and Srinivasan's (2009) participants' collections contained 5,850 files on average.

Conclusions

Many millions of computer users navigate to their personal files multiple times a day. Somewhat surprisingly, there has been very little research into this topic and as far as we are aware, ours is the first study to quantitatively investigate file navigation retrieval in a natural setting, and to examine the effect of structure on folder navigation. Because file navigation is so pervasive, improving navigation time by only a few milliseconds could save large enterprises several working months each day. Below are our conclusions regarding folder structure, navigation success, and the effect of structure on retrieval.

Folder Structure

Participants tended to create structure and use subfolders. They did not restrict their organization to default storage locations, such as My Documents. However, they also did not tend to create deep hierarchies and, typically, retrieved files from two levels below their main repository folder. They also did not create structures where higher levels were "organizational," containing mainly subfolders, and lower levels were used for storage, containing mainly files. Instead, files and folders occurred in approximately the same proportions on all levels. The overall picture is of a shallow, wide hierarchy containing relatively small folders that are a mix of files and subfolders.

Navigation Success

Our study showed a high success rate and reasonable retrieval time for folder-based navigations. This may partly explain previous research that showed navigation preference over search. Further research should use cognitive psychology and neuropsychological research methods to determine the reasons for this preference. Research should also compare the hierarchal method with alternative ones (i.e., multiple classification, automatic classification, and search) which have been claiming to outdate it for the last two decades.

The Effect of Structure on Retrieval

Our research indicates that increasing the breadth of folders is preferred to increasing their size or depth. Our participants clearly chose the breadth storage strategy, intuitively complying with the up to 21 heuristic rule. This allowed them to retrieve the majority of their files within 3–4 clicks (Zhang et al., 2004), which may explain their ability to find 94% of their target files in 14.76 seconds on the average. Future research should further investigate the relationship between folder structure and navigation retrieval using either large-scale studies, or controlled laboratory studies using eye tracking and the logging of participants' actions, possibly also taking semantics into consideration.

There are direct design implications to our results. We showed that increased folder size decreases retrieval efficiency because there are more items to scan within a folder. One reason why users accumulate large folders is because they tend to keep files of low subjective importance that they are unlikely to use (Boardman & Sasse, 2004). This may be because current system designs allow only two options regarding unimportant files: to delete the file (making it unavailable if needed) or keep it (and have it clutter the folder and compete for the users' attention). In earlier work the user-subjective approach (Bergman et al., 2003; Bergman, Beyth-Marom, & Nachmias, 2008) suggested the *demotion* principle. The demotion principle proposes that PIM systems should allow users to demote unimportant information items (making them less visually salient) so as to reduce distraction. Unlike deletion and archiving, demotion keeps items in their original context. We implemented this principle in a system called *GrayArea* (Bergman, Tucker, Beyth-Marom, Cutrell, & Whittaker, 2009) that allows users to demote files of low subjective importance by dragging them to a gray area at the bottom of the folder. A system evaluation showed that use of *GrayArea* reduced visual clutter in folders. According to the results of the current study we expect it to reduce retrieval time. We also proposed other user-subjective designs (such as *Old'nGray* that automatically grays out old versions of files to distinguish them from the latest version) to address this accumulation of items of low subjective importance.

Theoretically and empirically we need to develop better models of organization and its relation to retrieval. Our current study did not consider folder semantics, but this is an important determinant of both structure and retrieval that deserves more research attention. In addition, we need to determine whether our findings extend to different data types, e.g., e-mail or Web bookmarks. Are shallow, broad hierarchies also optimal for e-mail retrieval for example? Another question is whether e-mail folder navigation is short and successful similar to file folder navigation. This question is important because several mail systems attempt to replace folders with tags. Another important retrieval parameter is collection size and we need to better understand how this affects organization. In addition we did not look here at organization of, and navigation to, older nonactive files. Of course, we might expect success and efficiency for older files to be reduced compared with active file retrieval, but how quickly does memory for location degrade?

In conclusion, we need much more theoretical and technical work into manual organization and retrieval, prevalent activities that have strong implications for everyday productivity, but which remain critically underresearched.

Acknowledgments

We thank the participants, Prof. Ben Shneiderman, Isabella Gamsu, Dr. Kate Gee, Dr. Paul Clough, Dr. Xin Fan, and Prof. Dan Maoz for helping us with this study. This research was partially funded by the European Union Marie Curie Grant, PERG-GA-2009-248997.

References

Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks: Personal Web space construction and organization. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 41–48). New York and Reading, MA: ACM Press and Addison-Wesley Publishing Co.

Aula, A., Jhaveri, N., & Kaki, M. (2005). Information search and re-access strategies of experienced Web users. In Proceedings of the 14th International Conference on World Wide Web (pp. 583–592). New York: ACM Press.

Barreau, D.K. (1995). Context as a factor in personal information management systems. Journal of the American Society for Information Science, 46(5), 327–339.

Barreau, D.K., & Nardi, B.A. (1995). Finding and reminding: File organization from the desktop. SIGCHI Bulletin, 27(3), 39–43.

Bergman, O. (2006). The use of subjective attributes in personal information management systems. Unpublished doctoral dissertation (in Hebrew), Tel Aviv University, Israel.

Bergman, O., Beyth-Marom, R., & Nachmias, R. (2003). The user-subjective approach to personal information management systems. Journal of the American Society for Information Science and Technology, 54(9), 872–878.

Bergman, O., Beyth-Marom, R., & Nachmias, R. (2006). The project fragmentation problem in personal information management. In Proceedings of the Conference on Human Factors in Computing Systems (CHI 2006) (pp. 271–274). New York: ACM Press.

Bergman, O., Beyth-Marom, R., & Nachmias, R. (2008). The user-subjective approach to personal information management systems design—evidence and implementations. Journal of the American Society for Information Science and Technology, 59(2), 235–246.

Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N., & Whittaker, S. (2008). Improved search engines and navigation preference in personal information management. ACM Transactions on Information Systems, 26(4), 1–24.

Bergman, O., Tucker, S., Beyth-Marom, R., Cutrell, E., & Whittaker, S. (2009). It's not that important: Demoting personal information of low subjective importance using GrayArea. In the Proceedings of the 27th International Conference on Human Factors in Computing Systems (pp. 269–278). New York: ACM Press.

Boardman, R. (2004). Improving tool support for personal information management. Unpublished doctoral disseration, Imperial College, London.

Boardman, R., & Sasse, M.A. (2004). "Stuff goes into the computer and doesn't come out": A cross-tool study of personal information management. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 583–590). New York: ACM Press.

Capra, R.G., & Pérez-Quiñones, M.A. (2005). Using Web search engines to find and refind information. Computer, 38(10), 36–42.

Cole, I. (1981). Human aspects of office filing: Implications for the electronic office. In Proceedings of the Human Factors Society 26th Annual Meeting (Vol. 26, pp. 59–63). Santa Monica, CA: Human Factors Society.

Cutrell, E., Dumais, S.T., & Teevan, J. (2006). Searching to eliminate personal information management. Communications of the ACM, 49(1), 58–64.

Dourish, P., Edwards, W.K., LaMarca, A., Lamping, J., Petersen, K., Salisbury, M., et al. (2000). Extending document management systems with user-specific active properties. ACM Transactions on Information Systems, 18(2), 140–170.

Dumais, S.T., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R., & Robbins, D.C. (2003). Stuff I've seen: A system for personal information retrieval and re-use. In Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 72–79). New York: ACM Press.

Fisher, D., Brush, A.J., Gleave, E., & Smith, M.A. (2006). Revisiting Whittaker & Sidner's "email overload" ten years later. In Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work. New York: ACM Press.

DOI: 10.1002/asi

Furnas, G.W. (1997). Effective view navigation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM Press.

Gonçalves, D., & Jorge, J.A. (2003, June). An empirical study of personal document spaces. Paper presented at the DSV-IS'03 Design, Specification, and Verification of Interactive Systems Conference, Funchal, Portugal.

Hardof-Jaffe, S., Hershkovitz, A., Abu-Kishk, H., Bergman, O., & Nachmias, R. (2009a, July). In T. Barnes, M. Desmarais, C. Romero & S. Ventura (Eds.), How do students organize personal information spaces? Paper presented at the Second International Conference on Educational Data Mining, Cordoba, Spain.

Hardof-Jaffe, S., Hershkovitz, A., Abu-Kishk, H., Bergman, O., & Nachmias, R. (2009b). Personal information space organization strategies in learning context. Special Issue of Journal of Digital Information on the Social and Psychological Aspects of Personal Information Management, 10(5).

Henderson, S., & Srinivasan, A. (2009). An empirical analysis of personal digital document structures. In Proceedings of the Symposium on Human Interface 2009 on Universal Access in Human-Computer Interaction. Part I: Held as Part of HCI International 2009. Lecture Notes in Computer Science, 5617, 394–403.

Jacko, J.A., & Salvendy, G. (1996). Hierarchical menu design: Breadth, depth, and task complexity. Perceptual and Motor Skills, 82(2), 187–1201.

Jones, W., Bruce, H., & Dumais, S. (2003). In how do people get back to information on the Web? How can they do it better? In M. Rauterberg et al. (Eds.), Proceedings of the Ninth IFIP TC13 International Conference on Human–Computer Interaction (INTERACT 2003) (pp. 793–796). Amsterdam: IOS Press.

Jones, W., Phuwanartnurak, A.J., Gill, R., & Bruce, H. (2005). Don't take my folders away! Organizing personal information to get things done. In CHI '05 Extended Abstracts on Human Factors in Computing Systems (pp. 1505–1508). New York: ACM Press.

Kiger, J.I. (1984). The depth/breadth trade-off in the design of menu-driven user interfaces. International Journal of Man-Machine Studies, 20(2), 201–213.

Kim, J., Li, Y., Moy, N., & Ni, M. (2001). The menu design and navigational efficiency of the E-Maryland Portal. Retrieved November 5, 2009, from http://otal.umd.edu/SHORE2001/eMaryland/index.html

Kirk, D., Sellen, A., Rother, C., & Wood, K. (2006). Understanding photowork. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 761–770). New York: ACM Press.

Lansdale, M.W. (1988). The psychology of personal information management. Applied Ergonomics, 19(1), 55–66.

Malone, T.W. (1983). How do people organize their desks? Implications for the design of office information systems. ACM Transactions on Office Information Systems, 1, 99–112.

Larson, K., & Czerwinski, M. (1998). Web page design: Implications of memory, structure and scent for information retrieval. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 25–32). New York and Reading, MA: ACM Press and Addison-Wesley Publishing Co.

Miller, D.P. (1981). Depth/breadth tradeoff in hierarchical computer menus. In Proceedings of the 25th Annual Human Factors Society Meeting (pp. 296–300). Santa Monica, CA: Human Factors Society.

Neisser, U. (1964). Visual search. Scientific American, 210(6), 94–102.

Raskin, J. (2000). The humane interface: New directions for designing interactive systems. New York and Reading, MA: ACM Press and Addison-Wesley Publishing Co.

Russell, D., & Lawrence, S. (2007). Search everything. In W. Jones & J. Teevan (Eds.), Personal information management (pp. 153–166). Seattle, WA: University of Washington Press.

Shneiderman, B., & Plaisant, C. (2010). Designing the user Interface: Strategies for effective human-computer interaction (5th ed.). Reading, MA: Addison-Wesley.

Snowberry, K., Parkinson, S.R., & Sisson, N. (1984). Computer display menus. Ergonomics, 26(7), 699–712.

Tang, J.C., Drews, C., Smith, M., Wu, F., Sue, A., & Lau, T. (2007). Exploring patterns of social commonality among file directories at work. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM Press.

Tang, J.C., Wilcox, E., Cerruti, J.A., Badenes, H., Nusser, S., & Schoudt, J. (2008). Tag-it, snag-it, or bag-it: Combining tags, threads, and folders in e-mail. In CHI '08 Extended Abstracts on Human Factors in Computing Systems. New York: ACM Press.

Tauscher, L., & Greenberg, S. (1997). How people revisit Web pages: Empirical findings and implications for the design of history systems. International Journal of Human-Computer Studies, 47(1), 97–137.

Teevan, J., Alvarado, C., Ackerman, M.S., & Karger, D.R. (2004). The perfect search engine is not enough: A study of orienteering behavior in directed search. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 415–422). New York: ACM Press.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12, 97–136.

Whittaker, S., Bergman, O., & Clough, P. (2009). Easy on that trigger dad: A long term family pictures retrieval study. Personal and Ubiquitous Computing, 13(5), 17–30.

Whittaker, S., & Hirschberg, J. (2001). The character, value, and management of personal paper archives. ACM Transactions on Computer-Human Interaction, 8(2), 150–170.

Whittaker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Common Ground (pp. 276–283). New York: ACM Press.

Zaphiris, P., & Mtei, L. (1997). Depth vs breadth in the arrangement of Web links. Retrieved November 5, 2009, from http://otal.umd.edu/SHORE97/

Zhang, Y., Zhu, H., & Greenwood, S. (2004). Website complexity metrics for measuring navigability. In Proceedings of the Fourth International Conference on Quality Software (pp. 172–179). Piscataway, NJ: IEEE.

Appendix

This appendix contains numerical representations for the graphs in the results section, standard deviations where only averages are presented, and additional data.

TABLE A1. Target file hierarchal depth distribution and related retrieval time.

Hierarchal depth	Frequency of folder depth[1] (%)	Frequency of retrieval depth[2] (%)	Retrieval time[3] – *M*(*SD*)
0	75 (7)	75 (7)	5.6 (7.86)
1	181 (17)	286 (27)	11.65 (10.09)
2	234 (22)	202 (19)	15.78 (14.74)
3	246 (23)	211 (20)	16.87 (12.34)
4	140 (13)	117 (11)	16.2 (9.47)
5	81 (8)	74 (7)	16.78 (8.83)
6	51 (5)	43 (4)	18.92 (11.49)
7	27 (3)	27 (3)	24.76 (12.31)
8–11	19 (2)	19 (2)	25.22 (12.63)
Total	1,054 (100)	1,054 (100)	14.8 (12.09)

[1]Folder shortcuts are counted as the depth of the target file.
[2]Folder shortcuts are counted as 1st level depth.
[3]Of retrievals listed in column 3.

TABLE A2. Step folder depth distribution and related results for *Direct Success* retrievals.

Folder depth	Frequency (%)	Folder size – *M*(*SD*)	Subfolders – *M*(*SD*)	Subfolders % – *M*(*SD*)	*Target folders* – *N* (% of all folders)	Step duration – *M*(*SD*)
1	822 (33)	28.03 (44.18)	12.18 (20.39)	57 (29)	68 (8)	3.19 (3.92)
2	673 (27)	21.71 (27.96)	9.7 (11.35)	56 (36)	115 (17)	3.15 (3.77)
3	467 (18)	17 (18.58)	7.99 (12.71)	51 (39)	96 (21)	2.68 (3.15)
4	274 (11)	21 (45.21)	9.96 (37.09)	50 (42)	66 (24)	2.42 (2.76)
5	154 (6)	13.25 (13.81)	5.22 (9.39)	46 (43)	36 (23)	2.70 (3.06)
6	78 (3)	12.16 (9.9)	4.45 (5.38)	48 (38)	16 (20)	2.58 (2.60)
7–11	57 (2)	10.5 (11.78)	2.09 (2.74)	28 (43)	19 (33)	2.74 (2.89)
Total	2,525 (100)	22.46 (32.3)	10.64 (23.54)	54 (36)	416 (16)	2.94 (3.53)

DOI: 10.1002/asi