

Folder vs. Tag Preference in Personal Information Management

Ofer Bergman¹

Bar-Ilan University, Israel

Dept. of Information Science, Bar-Ilan University, 52900, Israel

Tel: 972-52-3583842, Fax: 972-3- 7384027

Email: oferbergman@gmail.com

Noa Gradovitch

Bar-Ilan University, Israel

Dept. of Information Science, Bar-Ilan University, 52900, Israel

Tel: 972-3-7323232, Fax: 972-3- 7384027

Email: noa.gradovitch@live.biu.ac.il

Judit Bar-Ilan

Bar-Ilan University, Israel

Dept. of Information Science, Bar-Ilan University, 52900, Israel

Tel: 972-3-5318351, Fax: 972-3- 7384027

Email: barilaj@mail.biu.ac.il

Ruth Beyth-Marom

Dept. of Education and Psychology, The Open University of Israel

108 Ravutski St. POB 808, Raanana 43107, Israel

Tel: 972-9-7781465, Fax: 972-9-7780632

E-mail: ruthbm@openu.ac.il

This is a preprint of an article accepted for publication in *Journal of the American Society for Information Science and Technology* copyright © 2012 (American Society for Information Science and Technology).

¹ Corresponding author.

Abstract

Users' preference for folders vs. tags was studied in two working environments where both options were available to them. In the *Gmail* study we informed 75 participants about both folder-labeling and tag-labeling, observed their storage behavior after a month, and asked them to estimate the proportions of different retrieval options in their behavior. In the *Windows 7* study, we informed 23 participants about tags and asked them to tag all their files for two weeks, followed by a period of 5 weeks of free choice between the two methods. Their storage and retrieval habits were tested prior to the learning session and after 7 weeks, using a designated *classification recording software* and a retrieval habits questionnaire. A controlled retrieval task and an in-depth interview were conducted. Results of both studies show a strong preference for folders over tags for both storage and retrieval. In the minority of cases where tags were used for storage, participants typically used a single tag per information item. Moreover, when multiple classification was used for storage, it was only marginal used for retrieval. The controlled retrieval task showed lower success rates and slower retrieval speeds for tag use. Possible reasons for participants' preferences are discussed.

Personal Information Management (PIM) is a basic human-computer behavior in which the user stores his/her information items (e.g. files, email, and Web favorites) in order to retrieve them later. Traditionally, PIM systems provided folders for information storage and retrieval; however, as a consequence of the popularity of Web 2.0, tags also defused into PIM systems.

It is widely claimed that tags have two fundamental advantages over folders; tags enable multiple classification and eliminate the need for hierarchies. (a) *Multiple Classification*: In the folders method, an information item can only be stored in a single folder; however, the user may have a number of possible classifications related to that item (Dourish et al., 2000). For example, pictures from a conference in Copenhagen can be stored under "Pictures," "Trips," "Conferences," or "Copenhagen." As time passes, users may forget the choice they initially made, making retrieval difficult. In contrast, the tagging method enables users to apply any number of tags to their information item, and

use any of these tags to retrieve it. (b) *No Hierarchical Location*: Folders may hide information items from view because they do not show items contained in subfolders (Malone, 1983). The tagging method consciously refutes hierarchies and locations. Instead, all information items are stored in a single repository and are retrieved via non-hierarchical means, such as tag search, tag selection or tag clouds.

Which option is preferable when both are available and users are familiar with the use of both? To the best of our knowledge, this is the first research to test this question. To do so, we conducted two studies: In one, the *Gmail* study, we first introduced 75 participants to the folder-labeling and tag-labeling options of the Gmail interface, waited a month, and then observed their mailbox to analyze their actual behavior. In the second study, the *Windows 7* study, we asked 23 participants to tag all files they used for a period of two weeks, then returned after five weeks to observe, using a designated software, the amount of tagging performed on new files, and conducted in-depth interviews regarding the users' behavior. In the *Windows 7* study, we also compared the retrieval efficiency of the two methods using a controlled test.

Theoretical Background

Folder Hierarchies

Hierarchical storage was first introduced to end-users in the Multics operating system in the mid 1960s. Users were allocated a personal directory, in which they could create their own subdirectories, sub-subdirectories, etc., and store their files in any of these "locations." This directory structure was later applied in the Unix and the Linux operating systems. The location metaphor became even clearer with the creation of digital folders first introduced in Xerox Star in 1981. A folder is a visual metaphor for a location: users can see the information items "inside" the folders, and they can manipulate items and folders in various straightforward ways, e.g. drag and drop information items from one folder to another, etc. This folder hierarchy metaphor was later applied by Apple in the Mac operating systems and then by Microsoft in the Windows operating systems. Thus, location-based storage has been used without significant modifications, continuously and almost exclusively, for several decades.

Through most of its long history, the hierarchical method has met with criticism for two main reasons, single classification and the hierarchical-locations metaphor, described in the following sections.

Single classification. Folder hierarchies are designed for single classification; however, in the user's mind, an information item could fit several different categories (Dourish et al., 2000). The user then needs to decide which category fits best, a process that can be cognitively challenging (Dumais & Landauer, 1983; Lansdale, 1988; Malone, 1983). Such categorization also needs to anticipate future usage (Bruce, 2005), which may be difficult because usage changes over time (Kidd, 1994). This causes a retrieval problem, as clearly articulated by Lansdale (1988): "*Placing a document into a filing system under one category places the information out of reach if retrieval is required for some other reason*" (p. 57). True, folder hierarchies do allow for short-cuts or aliases from one folder to another; however, this seems to go against the "this is there" location metaphor (Civan, Jones, Klasnja, & Bruce, 2008) and consequently are rarely implemented by users in practice (Bergman, 2006; Dourish et al., 2000). Criticism of the single classification of the hierarchical method is well established in PIM literature (Bloehdorn & Völkel, 2006; Dourish et al., 2000; Heckner, Heilemann, & Wolff, 2009; Hsieh, Chen, Lin, & Sun, 2008; Lansdale, 1988; Marsden & Cairns, 2003; Quan, Bakshi, Huynh, & Karger, 2003).

The hierarchical-locations metaphor.

Retrieval difficulty. Malone (1983) noted that filing (in his experiment, filing documents into physical folders) has a disadvantage – it loses its reminding function (becoming aware of the information item without thinking about it). As one of his participants said, "*You don't want to put it [a pile on the desk] away [to a folder] because that way you'll never come across it again*" (p. 107). The notion that folders can hide files is repeatedly mentioned in PIM literature (Civan et al., 2008; Jones, Phuwanartnurak, Gill, & Bruce, 2005; Lansdale, 1988). Quan et al. (2003) complain that users must remember the ordered sequence of topics and subtopics that were used to organize the information when attempting to retrieve it, even though the topics of interest during retrieval might be different from those during organization. Concerning emails

Elsweiler, Baillie, & Ruthven (2011) claimed that users find retrieval more difficult when emails are organized in folders than when left in the inbox.

Retrieval duration. In principle, one can imagine a transparent hierarchy (where you can see files contained in a folder's subfolders, sub-subfolders, etc.); however, since the location metaphor is used (and in a physical location, it is not enough to go to the closet to get a pair of jeans, one needs to know on which shelf the jeans were stored), the user needs to navigate the full path in order to reach the information item (Bloehdorn & Völkel, 2006), causing a possible long retrieval time. This claim gets some support in the Whittaker, Matthews, Cerruti, Badenes, & Tang (2011) study on emails. They found that emails organized in folders were not retrieved faster than the ones left in the Inbox.

Despite the above mentioned disadvantages the Hierarchical-Locations Metaphor and its application in the folder system gets support when comparing users' preferences for folder navigation and for search: there is consistent evidence that the former is preferred over the later (Barreau & Nardi, 1995; Bergman, Beyth-Marom, Nachmias, Gradovitch, & Whittaker, 2008; Boardman & Sasse, 2004; Capra & Pérez-Quñones, 2005; Kirk, Sellen, Rother, & Wood, 2006; Teevan, Alvarado, Ackerman, & Karger, 2004).

Tags as an Alternative to Folders

The above criticism motivated a search for a different model. Tags were suggested as a substitute for folders. Tags are a kind of metadata that describe the information item through a keyword or a term. Unlike folders, tags are non-hierarchical and users can assign as many tags as they want to an information item. Additional support for tags was given by Hsieh et al, (2008) who claimed that tags are more compatible than folders with our own cognitive system: "*In contrast to the hierarchy model, Collins and Loftus have proposed a spreading activation model of semantic processing. The non-hierarchical structure of semantic processing is similar to that of flat-structure tag networks.... Since most data stored in computers are declarative concepts, user-defined tags can be viewed as cognitive nodes in a concept network*" (pp. 1-2). In recent years, there has been extensive development of tag-related PIM prototypes, including: *Phlat* (Cutrell, Robbins, Dumais, & Sarin, 2006), *TagFS* (Bloehdorn & Völkel, 2006), *Gnowsis* (Sauermann et al., 2006), *ConTag* (Adrian,

Sauermann, & Roth-Berghofer, 2007), *TapGlance* (Robbins, 2008), *Zotero* (Ma & Wiedenbeck, 2009), *TAGtivity* (Oleksik et al., 2009), *BlueMail* (Tang et al., 2008; Whittaker et al., 2011) and *TagStore* (Voit, Andrews, & Slany, 2012).

The Web 2.0. The Web 2.0 revolution, in which users share their content on the Internet, had a crucial effect on the dissemination of tag usage. Applications such as Flickr and YouTube allow users to upload their pictures and video movies to the Internet together with the relevant tags describing them. Using these tags, the content could then be searched for by other users. In a paper entitled, “Why do tagging systems work?,” Furnas et al. (2006) attributes the success of tagging to the social aspect of Web 2.0:

Tagging systems have become increasingly popular after an element of social interaction was introduced. Social Tagging Systems connect the individual bookmarking activities of users into a network of tags and resources shared among multiple users. Social tagging systems, then, allow users to share their tags for particular resources. The tag sharing allows multiple added benefits, in discovery as well as retrieval... Social tagging systems may thus offer a way to overcome the “Vocabulary Problem” ... showing different users use different terms to describe the same things. (p. 37).

Note however, that these socially-related advantages of tags are irrelevant to PIM (Pak, Pautz, & Iden, 2007) as personal information management is, by definition, solitary rather than social information behavior.

Tags in Current PIM Systems

In recent years, tags defused from content sharing systems to PIM systems, and are currently offered by systems for managing Web Favorites, emails and files.

Web favorites. Del.icio.us is Web-based bookmarking software that uses tags instead of folders. It is both a PIM tool because users can bookmark Web pages for their own personal use, and a Web 2.0 tool because it allows users to share their Web favorites with other users across the Internet.

Emails. Google's Gmail, first introduced in 2004, supplied users with what they termed "labels." At first, Gmail allowed users to use labels only as tags (hereafter tag-labels): users could (and still can) add as many tag-labels as they wanted to an email. This method did not move the emails from the single repository where they were contained – the Inbox (see Figure 1a). Beginning in 2009, Gmail allowed users to use

labels also as folders (hereafter folder-labels). Participants can drag the email from the Inbox into a label (see Figure 1b). This is single classification because the email receives only one tag. (It disappears from the Inbox and loses its "Inbox" tag and moreover, if dragged into a second label, it loses its original one). The user needs to "open" the label and go through the messages it "contains" in order to retrieve the information needed. Therefore, Gmail users currently have two ways of using labels: tag-labeling, in which labels are dragged to an email, and folder-labeling, in which emails are dragged into labels. Regardless of the way that they are categorized, there are two ways of retrieving email: by using a location; i.e. opening a label in a folder-navigation-like way or using non-location based search (either by using tags or other search words).

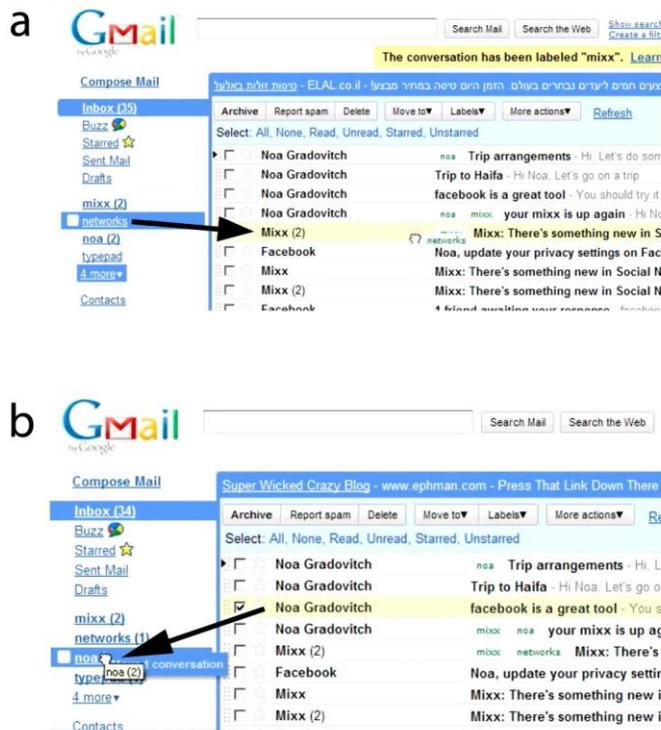


Figure 1. Tag labeling (a) and folder labeling (b) in Gmail.

Files. Microsoft introduced tags into its operating system in Windows Vista along with the traditional file folders hierarchy and they were retained in Windows 7. The files

can be stored either in a single repository (e.g. My Documents) or in designated subfolders with as many tags as the users choose to assign to the file. The users have three different options for using tags for retrieval: by using search (with one or more tags as keywords), combined with folder navigation (sorting the folder by tags), and using the "arrange by" option (arranging the files by tags instead of by folders).

Comparison between Folder and Tag Usage

Several laboratory studies compared between folder and tag usage in PIM. Five studies compared folder use to their own prototypes, testing for storage time, retrieval time, retrieval mistakes and number of mouse clicks: Quan et al. (2003), Ma & Wiedenbeck (2009), Voit et al. (2012), Hsieh et al. (2008) and Sajedi, Afzali, & Zabardast (2012). Three other studies tested in addition to efficiency also for cognitive load and degree of frustration (Civan et al., 2008; Gao, 2001; Pak et al., 2007). Summary of the results of the eight articles is presented in Table 1.

TABLE 1. Summary of papers testing folder vs. tag for efficiency, cognitive load and frustration level.

| | Folders are better | No Substantial / Significant Difference | Tags are Better |
|---------------------------------|-----------------------------------|--|---|
| Storage Time | Voit et al., Pak et al., (exp2) | Sajedi et al., Ma & Wiedenbeck | Quan et al., Pak et al., (exp1) |
| Mouse clicks/ physical diff. | Civan et al., | Pak et al., (exp2) | |
| Cognitive load Frustration | Gao Pak et al., (exp1), Gao | Pak et al., (exp1,2) | Civan et al., |
| Retrieval Time | | Voit et al., Ma & Wiedenbeck , Civan et al., Gao | Hsieh et al., Sajedi et al., Quan et al., |
| Mistakes Mouse clicks | Gao Hsieh et al., | Gao | Sajedi et al Voit et al., Ma & Wiedenbeck , Pak et al., (exp2) |
| Cognitive load | Gao (secondary task) | Gao (subjective evaluation) | |

Table 1 presents mixed results regarding almost all comparisons between folders and tags, giving no clear indication that tag usage is superior to folder usage. All eight studies were using controlled tasks in which information was given to the users by the researchers. While such methodology has important advantages it does not use participants' own personal information. Moreover, in none of the studies participants could choose whether to work with folders or tags and express their preferences via real choice behavior. Preferences were inferred from preference ratings. Finally, none of the above authors studied participants choices related to single or multiple classification.

The only study we are aware of that monitored users' choices for tags vs. folders (but did not monitor users behavior regarding multiple vs. single classification) in a natural setting, was done by two members of the Google research team who were involved in introducing folder-like Gmail labels (Rodden & Leggett, 2010). This change was done because "*While the flexibility and power of labels was appreciated by the users who discovered them, labels were often not discoverable or understandable for another set of users, who would have been happy with the simpler solution of folders*" (p. 4588). Indeed their results showed that, with the new possibility of folder-labeling new users' chances of creating labels doubled and the percentage of folder-label storage exceeded that of tag-label storage. They concluded that "*With this design, the team expects that users will be able to discover and use the more advanced aspects of labeling (such as multiple inclusion) at the point where they identify a need, but it will be important to verify that*" (p. 4595). Our research had tested just that, however unlike (Rodden & Leggett, 2010) we acknowledged our participants of both options so that results could not be explained by users' ignorance.

The main purpose of the present research is to test users' preferences for folders vs. tags in PIM environments (Gmail and Windows 7) that support both options. This was done in several ways: (a) Directly, by observing folder and tag related behaviors; (b) By testing single vs. multiple classification behaviors, and (c) By studying hierarchical vs. non-hierarchical behavior. In addition, we wanted to test the retrieval efficiency of the two methods.

Research Questions

Folder vs. Tag Preference

Folder vs. tag preference is important to test because it can inform future systems' design.

1. Which method is used more frequently when *storing* information items – folders or tags? As mentioned, this question was tested in a natural setting by (Rodden & Leggett, 2010) however their participants may not have been aware of both options, which could serve as an alternative explanation for their results. In our experiments we were careful to explain both options to our participants.

2. Which method is used more frequently when *retrieving* information items – folders or tags? To the best of our knowledge this has never been tested before. We tested this by using self-estimation questionnaires. The validity of these questionnaires is discussed in the Method section.

Single vs. Multiple Classification Preference When Using Tags

The single classification of the hierarchic method was widely criticized and the multiple classification widely advocated (see Theoretical Background section), however to the best of our knowledge users' preference of single vs. multiple classification was never tested in a naturalistic settings. It was implicitly assumed that because there are different possible classifications for each information item people would use them for storage and retrieval whenever possible; however this assumption was never empirically tested. Notice that tagging does not necessarily lead to multiple classification because users can store and retrieve information items using a single tag for each information item.

3. Which classification option is used more frequently when *storing* information items – single classification or multiple classification? This was measured by the percentage of items that received a single tag compared to the percentage of items that received two tags or more from the 100% of the participant's tagged items.

4. Which classification option is used more frequently when *retrieving* information items – single classification or multiple classification? Multiple classification can be used for

retrieval in two ways: a. By using more than one tag in a single retrieval. This conjunction action drastically reduces the results list and could lead to a faster retrieval.
b. By using a different single tag in different retrievals, allowing the users to utilize each of the items' tags that they happened to remember at the time.

Hierarchical vs. Flat Storage

5. Which structure is preferred for information items storage – hierarchical or flat? We have surveyed some of the literature that criticizes the hierarchal method (see The ‘Hierarchical-Locations Metaphor’ section in the Theoretical Background) for example for hiding away files (Civan et al., 2008; Jones et al., 2005; Lansdale, 1988; Malone, 1983) . This raises the question of whether users prefer a hierarchal storage or a flat one such as suggested by the tag system.

Location-based vs. Non-location Retrieval

Regardless of the way the information items are stored, the user can choose between retrieving information items using location-based navigation or by using non-location based retrieval methods, such as search.

6. How do people prefer to retrieve their information items – by using location-based navigation or non-location retrieval methods? Our section ‘The Hierarchical-Locations Metaphor’ focuses on critic regarding navigation for example that users must remember the ordered sequence of topics and subtopics that were used to organize the information when attempting to retrieve it, even though the topics of interest during retrieval might be different from those during organization (Quan et al., 2003). Therefore, tag based systems deliberately omitted the possibility of location based navigation, replacing it with non location retrieval methods such as tag search. Although location vs. non location preference was studied in the past (Barreau & Nardi, 1995; Bergman, Beyth-Marom, Nachmias et al., 2008; Boardman & Sasse, 2004; Capra & Pérez-Quñones, 2005; Kirk et al., 2006; Teevan et al., 2004) showing clear preference for navigation over other research methods, in these studies, however, tags were not introduced to participants. We wanted to know whether introducing tags would alter users’ preference.

Tagging Retrieval Efficiency

7. Does the use of tags increase retrieval efficiency? Table 1 indicates mixed results regarding retrieval efficiency. Regarding retrieval time some of the studies did not find a significant or substantial differences (Civan et al., 2008; Gao, 2011; Ma & Wiedenbeck, 2009; Voit et al., 2012), while others found tag-based retrieval to be faster than not using tags (Hsieh et al., 2008; Quan et al., 2003; Sajedi et al., 2012). Regarding mistakes Gao (2011) found more mistakes using tags while Sajedi et al. (2012) found more mistakes when using non tag retrieval. In contrast to these previous studies ours relates to the participants' own personal files.

Possible Age and Computer Literacy Effect

8. Do age and computer literacy affect storage and retrieval behavior?

We used this question to test the external validity (generalizability) and possible interpretations of our findings. If, for example, younger participants tended to use tags or multiple classifications more than older participants, this might raise the possibility that habit plays a crucial role in the adoption of new technologies/applications.

Research Method

To test users' preferences for tags or folders, we ran two studies in two different environments, to increase the external validity of our results. While requiring more research effort, converging results allow for greater generalizability. Furthermore, to increase the convergent validity of our measurements we used a multi-operational approach: we collected various kinds of data (quantitative and qualitative) using surveys, experimental manipulation and, natural observations. We conducted two studies, the *Gmail* study, to test email management preferences, and the *Window 7* study to test preferences regarding file management. In both studies, we exposed our participants to folders and tags. In the Gmail study, we examined their mailbox for folder and tag labeling behavior after a month during which they could use whichever method they chose. This study tested all research questions except question 7. In the *Windows 7* study, we asked our participants to use tags instead of folders for two weeks, conducted a controlled retrieval task, and then returned five weeks later to test users' folder/tag preferences after a period during which they managed their folders as they chose. This

study tested all research questions except question 8. Unlike the majority of previous literature testing for folder/tag preference, both studies were conducted in naturalistic settings, i.e. participants managed their own information items on their own computers as part of their daily routine. This highly contributes to the ecological validity of the results. The studies were approved by the IRB at Bar-Ilan University, and all participants signed an informed consent form. The stages and procedures of the two studies are detailed in the following sections. After a short description of the two studies, we summarize their main features in Table 2.

Gmail Study

Participants. Participants were 75 Israeli students who used Gmail with a Gmail interface (i.e. not using an external email client) as their primary mailbox. Twenty six of the participants were men and 49 women. Their ages ranged from 20 to 58 ($M = 30.02$, $SD = 6.96$). Participants who used labels reported to have used Gmail for a period of 1-60 months prior to the study ($M = 20.15$, $SD = 18.99$). Thirty five of the participants (47%) had used Gmail prior to the introduction of the folder-labeling option when only tag-labeling was available. Participants self-reported a medium-high degree of Gmail expertise ($M = 3.72$, $SD = 0.92$ on a 1-5 Likert scale, see Appendix A, question 11).

Procedure. The study included three stages. (1) First Meeting: The first meeting was dedicated to a *learning session* in which we met with each of the participants and showed them a 15 minute video. The video demonstrated the various storage and retrieval possibilities that Gmail offers for both folder-labeling (dragging the email from the Inbox into a label) and tag-labeling (add tag labels to emails in the Inbox).² Special emphasis was placed on explaining multiple classification. The video was informational and we were careful not to express any opinions regarding the two options. (2) Activity Period: The following month was one of *free choice activity* – participants were free to manage their email in any way they choose. (3) Email Transaction: At this stage, participants were instructed to send us, via email, *print screens of their mailboxes* showing their last 100 emails, in the All Mail mode. This mode includes tag-labeled emails, folder-labeled

² The video (in Hebrew) can be viewed at

http://telem.openu.ac.il/cgi-bin/telem/tools/open_video/showVideo.pl?p=public&v=gmail_noa_160310.aspx

emails, unclassified emails and other data not analyzed (outgoing emails, deleted email, drafts and chats). These print screens supplied us with data about actual folder/tag labeling behavior. Participants also answered a questionnaire regarding their retrieval habits during the activity period as well as other information. The *questionnaire* (see Appendix A) was sent to them via email.

Windows 7 Study

The *Windows 7* study differed from the *Gmail* study in four respects: (a) The environment (files vs. emails); (b) An opportunity to practice for two weeks using tags only, as we felt that teaching participants how to tag their files was not enough because they had been storing their files in folders for years and we wanted to minimize the habit effect on preference as much as possible; (c) We conducted controlled retrieval tasks as well as (d) in-depth interviews with the participants. By using such interviews (possible only with the relatively small size of the sample), we included some qualitative methods in the research design. In this study, we used a specially developed *classification recording software* with which we scanned the participants' storage area and for each file, recorded its storage location and the number of tags attached to it.

Participants. Participants were 23 Israeli students who used the Windows 7 operating system on their laptops. Twelve participants were men and 11 were women. Their ages ranged from 21 to 38 ($M = 26.82$, $SD = 4.49$). They had used Windows 7 for 2-18 months prior to the study ($M = 12.39$, $SD = 8.06$). All participants had used computers for over 8 years and self-reported themselves as highly computer literate ($M = 4.43$, $SD = 0.72$ on a 1-5 Likert scale).

Procedure. The study had 5 stages. (1) First Meeting: Prior to the learning session, we gathered some data: we scanned the participants' storage area with our *classification recording software* to learn about their initial folder and tag classification habits and asked them to answer a *retrieval questionnaire* regarding their habits (see Appendix B). In the *learning session*, we showed participants how to use tags for file storage and retrieval. (2) First Activity Period: A *forced tagging activity* of two weeks, in which participants were instructed to tag all files they encountered and to retrieve files by using these tags. (3) Second Meeting: We used our *classification recording software* for

a manipulation check which indicated that they had complied with our instructions: On average, they tagged 71% of the 505 files which they accessed ($SD = 26\%$) and used multiple classification (added more than one tag) to 55% of their tagged files ($SD = 40\%$). *The second meeting* also included a controlled *retrieval task* in which we asked participants to retrieve files in two conditions: (a) using tags and (b) not using tags. We compared retrieval time and success between these conditions. More specifically we used a different specially designed software to randomly select a tagged file; the software then paired the file with another file used retrieved by the user at approximately the same time. We asked participants to retrieve the first file by using tags and its paired file by using any method they chose other than tags. Conditions were counter-balanced and the sessions were video screen-captured and analyzed. In order to avoid biases (e.g. an attempt to please the tester), we told our participants that the experiment was over but we would like to meet them again for a different experiment. (4) Second Activity Period: *Free choice activity* in which participants could manage their files as they chose for approximately five weeks. They were unaware that we intended to record their choices. (5) Third Meeting: we (a) ran our *classification recording software* again to test folder and tag storage on files that were accessed in the past three weeks (giving a gap of two weeks between the *forced tagging* activity to time actually tested in the free choice activity to avoid a momentum effect); (b) asked participants to answer our *retrieval questionnaire* regarding their behavior during the free choice activity; and (c) conducted an *in-depth interview* in which participants were asked to reflect upon their actual behavior. Typically, the interviewer informed them about specific results extracted from our *classification recording software* or *retrieval questionnaire* and asked for their reaction. The answers to these questions were qualitatively analyzed into distinctive categories. A between-judges test using these categories showed complete agreement.

TABLE 2. Stages and procedures in the two studies.

| Stage | Gmail Study | Window 7 Study |
|--|--|---|
| (1) First Meeting | <p>Learning session: folder-labels and tag-labels</p> | <p>Data gathering: - check folder and tag storage via <i>classification recording software</i> -retrieval habits questionnaire</p> <p>Learning session: tags</p> |
| (2) First Activity Period | <p>Free choice activity (1 month): -"arrange your Emails as you choose"</p> | <p>Forced tagging activity (2 weeks): -"use only tags to arrange your files"</p> |
| (3) Email Transaction / Second Meeting | <p>Data gathering (via email): -print screen of the Mailbox - retrieval habits questionnaire</p> | <p>Data gathering (in meeting): -use <i>classification recording software</i> for a manipulation check -a controlled retrieval task</p> |
| (4) Second Activity Period | | <p>Free choice activity (5 weeks): -"arrange your files as you choose"</p> |
| (5) Third Meeting | | <p>Data gathering: -check folder and tags storage via <i>classification recording software</i> -retrieval habits questionnaire -in-depth interview</p> |

Results

Folder vs. Tag Storage Preference

RQ1: Which method is used more frequently when storing information items – folders or tags?

Gmail study. We investigated this question by analyzing participants' email printouts after the *free choice activity* period. Of our 75 participants, 60 used labels (in either way). Our results section will only report on these 60 participants. On average, they created 32.31 labeled messages ($SD = 26.29$), which represented 45% of their messages ($SD = 26\%$). Examining the labeled messages of these 60 participants, there was an average of 33% ($SD = 41\%$) tag-labeled messages (labeled messages in the Inbox) and 64% ($SD = 42\%$) folder-labeled messages (messages that were moved to a label). The difference between the two was significant, $t(59)=2.97$, $p<0.01$.

There is an additional way to label – by first adding labels to the message and then archiving it. These messages could have multiple labels but are not in the Inbox. We did not mention this possibility in our informational movie. However 10 of our participants had used this technique as we found that some of their messages have two labels or more but no Inbox label. Overall this behavior seems marginal: on average only 3% of the messages had two labels or more but no Inbox label ($SD = 9.43\%$). It is also possible that some of the folder-labeled messages were created by adding a single tag and then archiving the message, in which case the participants may possibly work under a tagging mental model. However since these messages have a single location and a single classification they are in effect not different than other folder labeled messages.

In the questionnaire, participants were first asked to state their preference (either folder-labeling or tag-labeling), and then to explain their choice (see Appendix A, question 3). Table 3 classifies their explanations for folder-labeling or tag-labeling behavior.

TABLE 3. Reasons for using folder-labels or tag-labels usage when storing email ($N= 51$).

| Reason | Folder-labelers | Tag-labelers |
|------------------------------|-----------------|--------------|
| a. Decrease inbox size | 13 | - |
| b. Order | 9 | - |
| c. Easier retrieval | 9 | 6 |
| d. Habit | 3 | 7 |
| e. See all messages in Inbox | - | 10 |
| f. See message in two places | - | 4 |

Table 3 shows that the most frequently mentioned reason for folder-labeling behavior was to decrease inbox size (row a). The two next most frequent reasons were to force order on the Mailbox (row b) and to allow for easier retrieval (row c). Only three participants ascribed their folder-labeling to habit (row d). As for tag-labeling the most frequently mentioned reason (row e) was to see all the messages in the same location (i.e. the Inbox). Other reasons given for tag-labeling were easier retrieval (row c) and habit (row d). The habit may have been formed because tag-labeling was the only storage option for several months before the folder-label method was introduced in Gmail. Only 4 participants mentioned multiple classification (row f) as a reason for their tagging behavior.

Windows 7 study. On average, participants created 57.13 files ($SD = 128.66$) during the second activity period in which they could manage their files as they chose (between the second and the third meetings). On average, 2 files (4%) were stored in the Windows default location – My Documents ($SD = 11%$). This percentage did not change from the one measured prior to the *forced tagging activity* ($M = 4%$, $SD = 21%$). Participants chose a specific folder for the remaining 96% of the files, clearly indicating folder storage behavior.

Of our 23 participants, only 6 (26%) created tags during the *free choice activity* period. One was an avid tagger who tag-labeled all 14 of her files, two participants tag-

labeled only one (3%) and 8 (13%) of their files, and the other three participants tag-labeled 180 (30%), 3 (23%) and 10 (24%) of their files.

In the in-depth interviews, we asked the participants why they stopped or continued using tags after the forced tagging period. Their answers were classified into categories detailed in Table 4:

TABLE 4. Reasons for using/not using tags ($N = 23$).

| Tag use (N participants) | Reasons (N responses) | Quotation |
|-----------------------------|-------------------------------------|--|
| Did not use (17) | a. Tags are difficult to use (7) | p5: "In the two weeks I used it, I found it very difficult to use and impractical." |
| | b. Folders allow fast retrieval (7) | p22: "My files are ordered enough for me to find things fast." |
| | c. Folders maintain order (5) | p10: "I don't use tags, I keep a well ordered folder hierarchy. So that I have fast access to my files. That way everything is neatly in place." |
| | d. Tagging is time consuming (5) | p21: "Both in storage and retrieval it delays [me], takes too much time." |
| | e. Habit (4) | p12: "No particular reason, [it's] just a habit" |
| | f. No benefit (1) | p4: "I don't understand what's the benefit of this whole tagging thing." |
| Used tags (6) | g. Liked the method (3) | p1: "I like this method more than it appears [from the % of tagged files]. I didn't have much time for it in the last three weeks." |
| | h. Order maintenance(2) | p14: "It's more ordered this way." |

Table 4 shows that some of the reasons for not using tags have to do with folder preference for storage (c) or retrieval (b), while other reasons criticize tagging as difficult to use (a) and time consuming (d), or as having no apparent benefit (f). Only four participants ascribed not using tags to their old folder habits. Participants who used tags

had expressed appreciation for the tagging method (g) and thought that it kept their information items well in order (h).

Combined overview. Both studies showed a clear preference for folder over tag storage. There was, however a difference between the two studies with regard to tagging behavior: while in the *Windows 7* study, tagging was marginal, in the *Gmail* study, it was quite substantial. However, the dominant reason given by the participants for tagging in the Gmail study was not multiple classification but the desire to keep categorized messages in the Inbox with the rest of the messages.

Folder vs. Tag Retrieval Preference

RQ2: Which method is used more frequently when retrieving information items – folders or tags?

Gmail study. After the free choice stage, we asked our participants to estimate the percentages in which they had used each of the available retrieval options out of the 100% of retrieval occurrences (see Appendix A, question 4)³. Results are presented in Table 5.

TABLE 5. Mean and standard deviation of estimated percentages for different retrieval options in the *Gmail* study.

| Retrieval option | M (SD) |
|---------------------------------|-----------|
| a. scanning the inbox | 53% (33%) |
| b. opening a label | 16% (20%) |
| c. search without using labels | 27% (29%) |
| d. search using a single label | 3% (8%) |
| e. search using multiple labels | 0.1% (1%) |
| f. other options | 0.9% (8%) |
| Total | 100% |

³ Subjective retrieval estimates based on answers to a questionnaire were compared to actual behavior in a previous study (Bergman, Beyth-Marom, Nachmias et al., 2008). Results indicated high validity of estimates.

As Table 5 indicates, participants estimated that they used folder-label retrieval (row b) in 16% of their retrievals and tag-label retrieval (rows d and e) in only 3% of their retrievals on average. The difference between these results is significant, $t(72) = 2.61, p < 0.05$.

Windows 7 study. In order to answer this question we used a retrieval questionnaire presented in Appendix B.

Results of the questionnaire administered at the first meeting showed no tag-related retrievals prior to the experiment. Results of the questionnaire administered at the third meeting relating to the *free choice activity* are presented in Table 6.

TABLE 6. Mean and standard deviation of estimated percentages for different retrieval options in the *Windows 7* study.

| Option | M(SD) | Tag usage | M(SD) |
|------------------|-----------|---------------|-----------|
| Navigation | 61% (7%) | No tags | 93% (13%) |
| | | Single tag | 7% (13%) |
| Desktop shortcut | 20% (28%) | - | - |
| Search | 10% (18%) | No tags | 92% (15%) |
| | | Single tag | 8% (15%) |
| | | Multiple tags | 0% |
| Recent doc. list | 7% (15%) | - | - |
| Arrange by | 1.5% (4%) | No tags | 97% (2%) |
| | | Single tag | 3% (2%) |
| Others | 0.5% (3%) | - | - |
| Total | 100% | | |

Table 6 indicates that on average participants estimated that they used folder-based navigation for 61% of their retrievals ($SD = 7\%$). This is not different from the estimated percentage of navigation they had given before the *forced tagging* activity ($M = 61\%, SD = 5\%$). The average estimated percentage of tag-based retrievals for all

participants was much lower – 5% ($SD = 15\%$)⁴. For the 6 participants who stored files using tags during the free choice stage, the average percentage of retrieval by tags was 19% ($SD = 25\%$) compared to 37% by navigation ($SD = 31\%$). Thus, even those participants who did use tags for storage used folders more than tags in their retrieval behavior.

Reasons for folder retrieval preference over tag retrieval included difficulty of tag usage ($N=6$, p5: "*it's not easy to use*") and ease of folder use ($N=3$; p18: "*we all know how easy it is to find things in a tree [structure]*"), habit ($N=4$; p13: "*It's all about habits, I'm just used to working without it*"). Other participants found tagging unnecessary ($N=4$; p15: "*I know where my files are [located], [so] I don't need tags*"). Another interesting reason (p1) is that "*if I use tags I lose the entire picture [of the folder hierarchy]*." Three participants expressed positive opinions regarding tag retrieval during the interview (p6: "*If something gets lost it will be easier for me to find it*"). One participant said she uses tag retrieval but only as a "*last resort*" (p14).

Combined overview. Both studies showed little use of tagging when retrieving information items, even for participants who used tags to store their items. Folder navigation was the dominant retrieval option.

Single vs. Multiple Classification Storage Preference

RQ3: Which classification option is used more frequently when storing information items – single classification or multiple classification?

Gmail study. We tested multiple/single classification preferences only for the 26 participants who had been using tag-labels (i.e. that their mailbox contained labeled messages with no Inbox label). Email printout results across participants show on average that they added a single label to 92% ($SD = 23\%$) of their labeled messages ($M = 22.3$, $SD = 20.61$ messages, in absolute numbers) and added two or more labels to only 8% of their messages ($M = 1.07$, $SD = 1.49$ messages). The difference between results was again significant, $t(25)=10.67$, $p<0.001$. This indicates that even when tag-labels were used and multiple classification was easy to deploy (simply by dragging more than one

⁴ We calculated the estimated tag use percentage in the following way: (% tags in navigation * % navigation) + (% tags in search * % search) + (% tags in arrange by * % arrange by).

label to a message) multiple classification was the exception, while single classification was the norm.

Windows 7 study. Table 7 details the amount of multiple classification for each of these 6 participants that tagged their files.

TABLE 7. The amount of multiple-classification among participants who used tags ($N = 6$).

| Participant | No. of files with two tags or more | % of files with two tags or more of the tagged files | % of files with two tags or more of all files |
|-------------|------------------------------------|--|---|
| p1 | 0 | 0 of 180 = 0% | 0 of 593 = 0% |
| p2 | 0 | 0 of 3 = 0% | 0 of 13 = 0% |
| p6 | 7 | 7 of 14 = 50% | 7 of 14 = 50% |
| p7 | 1 | 1 of 1 = 100% | 1 of 37 = 3% |
| p9 | 1 | 1 of 1 = 100% | 1 of 8 = 13% |
| p14 | 10 | 10 of 10 = 100% | 10 of 41 = 24% |

Table 7 shows that p1 and p2 did not use multiple classification at all (notice that p1 creates an exceptionally large number of files and uses single classification tagging for almost one-third of them); p7 and p9 used multiple classification for the single file each of them created; and two participants (p6 and p14) tagged a substantial number of the their files using multiple classification.

In the interview, the two participants who used multiple classification for a substantial number of their files said that they do so to enrich the description of their files (p14: *“I wanted to have as many details on it as possible”*). The two participants who each used multiple classification for only one file said that they did so because they thought it would ease retrieval (p7: *“Because [each information item] can have different attributes and so I can make a conjunction and reach a much higher level of precision upon retrieval”*). Participants who used only single classification explained that it is simpler than multiple classification ($N=3$; p1: *“Multiple tags are somewhat cumbersome. One short name for a tag is something that can be remembered and eases retrieval”*).

Others thought single classification to be a better method for reasons that had to do with storage ($N=3$; p2: “*I didn’t see a reason to use more than one tag. I prefer a single tag it’s more ordered and focused*”) and retrieval ($N=3$; p2: “*Multiple classification has simply not an advantage as far as I’m concerned. Single classification allows for easy retrieval*”). Interestingly, ten participants who did not use multiple classification (or tagging at all) remained silent and seemed somewhat embarrassed when not responding to this question during the interviews.

Combined overview. Results from the *Gmail* study shows only 8% of multiple classification among the tag-labeling participants. In the *Windows 7* study only 6 participants tagged their files (two of them tagged a single file each) so no conclusion can be drawn, however the fact that so few participants tagged their files does not indicate an enthusiasm regarding multiple classification.

Single vs. Multiple Classification Retrieval Preference

RQ4: Which classification option is used more frequently when retrieving information items – single classification or multiple classification?

Gmail study. Table 5 indicates that on average only 0.1% of the messages were retrieved via search using two or more labels. The Gmail interface allows users who stored their messages with multiple labels to “open” each of these labels (in a folder-like way) and retrieve the same message using different labels. In order to test the use of this option, we first asked our participants: “Have you ever added two labels or more to a message?” Twenty-five participants answered this question positively. Those participants were then asked to specify which of six predefined options they usually use to retrieve their multiple-labeled messages (see Appendix A, question 6). Results are presented in Table 8.

As shown in Table 8, of the 25 participants, 12 (48%) indicated that they used single classification (options b and e) when retrieving a message with multiple labels. Only 4 participants (16%) indicated that they used multiple classification (options c and f) when retrieving a message with multiple labels. These results indicate that even in the rare cases when participants did use multiple classification storage, they tend to use only one label when retrieving messages.

TABLE 8. Number of participants who used each option to retrieve their multiple-classified messages ($N = 25$).

| Retrieval option | Participants N (%) |
|---|--------------------|
| a. scanning the inbox | 5 (20%) |
| b. opening a single label | 11 (44%) |
| c. opening multiple labels | 2 (8%) |
| d. search option without labels | 4 (16%) |
| e. search option always with the same label | 1 (4%) |
| f. search option with different labels on different occasions | 2 (8%) |

Windows 7 study. When asked about the use of multiple classification for file retrieval in the file retrieval questionnaire (see Appendix B, question 6), none of the participants reported such behavior, including the four participants who used multiple classification during the storage phase. One of them said that the use of a single tag is enough for retrieval: *“As soon as I think of one tag it’s enough. I find what I need.”*

Combined overview. Results from both studies show that there is little or no use of multiple classification for retrieval, even in the minority of cases where it is used for storage.

Hierarchical vs. Flat Storage

RQ5: Which structure is preferred for information item storage – hierarchical or flat?

Gmail study. At the time the experiment was conducted, the Gmail design was flat and did not support hierarchical label structure. We therefore presented a hypothetical hierarchical label structure in the questionnaire and asked our participants which system would they prefer if they had the choice between the current system and the hypothetical hierarchic one (see Appendix A, question 7).

Of the 73 participants who answered this question, 58 (79%) indicated that they would prefer the hierarchical option, and only 5 (7%) preferred the current flat method. The remaining 10 participants (14%) were indifferent. These results suggest that

theoretically Gmail users prefer a hierarchical label structure. No doubt, future research will have to test whether these preferences express themselves in practice: when exposed (in the future) to a hierarchical label structure, will users use it.

Windows 7 study. In Windows 7, participants do not necessarily need to use a hierarchic nested folder structure. Instead, they can use flat locations and store their files on the desktop, in My Documents, or in any folder which the users defined as their root directory (e.g. C:/David). On average, only 14% of the files that participants stored during the free choice stage were stored in flat locations ($SD = 26\%$) compared to 86% stored deeper in the folder hierarchy. The percentage of files stored in flat locations increased from 10% prior to the *forced tagging* activity ($SD = 21\%$) to 14% in the test phase ($SD = 14\%$); however, this cannot be due to the use of tags because for the six participants who used tags, the average percentage of files stored in flat folders decreased (from 25% prior to the *forced tagging* to 20% in the test phase).

When asked in the interview why they used the deep hierarchic structure, participants said that it helped them maintain order ($N=11$; p1: "*Order is important for me*"), eases retrieval ($N=7$; p9: "*It's important because otherwise I wouldn't know where to find them*"). Other reasons were habit ($N=1$; p12: "*I can't change*") and ease of use ($N=1$; p10: "*it's much easier*").

Combined overview. Both studies showed participants' preference for hierarchical over flat storage. The *Gmail* study participants could not use hierarchies but their answer to the hypothetical question indicates that they would like to, and the *Windows 7* study participants stored only a minority of their files in flat locations (such as the desktop and My Documents), and the majority of files deeper in the folder hierarchy.

Location-based vs. Non-location Retrieval

RQ6: How do people prefer to retrieve their information items – by using location-based navigation or non-location retrieval methods?

Gmail study. Table 5 clearly indicates that users prefer location based retrieval through navigation to non-location based search. Participants estimated that they used search for 30% of their messages on average (options c, d, and e) and navigated to 69% of their messages (options a and b). The difference was statistically significant, $t(74)=5.4$,

$p < 0.001$. These results indicate that Gmail users prefer to retrieve their messages using location-based navigation rather than by a non-location based search.

Windows 7 study. In addition to the location-based retrieval method (folder navigation), Windows 7 provides four non-location based retrieval methods: search, recent documents, shortcuts, and “arrange by” (using any attribute other than folders, e.g. date). We asked about the use of these options in the retrieval questionnaire (Appendix B, question 1). On average, participants estimated that they used navigation for 61% of their retrievals ($SD = 33\%$). All other non-location retrievals added up to 39% (short-cuts, 21%; search, 10%; recent documents, 7%; “arrange by,” 1%). These were the exact percentages which the users estimated prior to the *forced tagging* (only then the standard deviation was 25% and the non-location retrievals percentages divided slightly differently). These results suggest that participants prefer location over non-location retrieval and that the *forced tagging* activity had no effect on this preference.

When asked why they prefer navigation to search, participants mentioned ease of use ($N=9$; p15: “*that's the easiest way for me*”), habit ($N=8$; p6: “*It's a habit – that's the way I learned it, so this is how I do it*”), order ($N=2$; p10: “*it's more ordered: you get into a folder, a subfolder and another folder beneath it*”) and consistency (p26: “*because that's the way I save it*”).

Combined overview. Both studies show that most users retrieve their personal information items through navigation. Interestingly, in both emails and files, navigation accounts for about two thirds of the retrievals, while non-location retrieval accounts for the remaining third.

Tagging Retrieval Efficiency

RQ7: Does the use of tags increase retrieval efficiency?

We tested this research question only in the *Windows 7* study. During the second meeting (immediately after the *forced tagging* period), we conducted a controlled retrieval task. We used a designated software to randomly select 6 tagged files for each participant. The software then paired each of these files with another file created/modified at approximately the same time. The retrieval task had two conditions: tag-retrieval and non-tag retrieval. In the tag-retrieval, we asked our participants to

retrieve files of the first selection by using tags in any way they choose (search by tag, arrange by tag or navigation and then sorting by tag). In the non-tag retrieval task, we asked them to retrieve the paired files in any way they choose, other than using tags. The order of conditions was counter-balanced and the computer screens were video recorded and analyzed. 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. Retrieval time was measured from the time the tester announced the name of the file to be retrieved, 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). Altogether, participants retrieved 119 pairs of files (5-6 pairs for each participant). When asked to use tags, participants mostly used search (74%); 20% used "arrange by" and 6% used tag sorted navigation. When conducting non-tag retrieval, participants mostly used navigation (64%), but also used search (25%) and short cuts (11%).

Results show that the average retrieval time when using tags was 21.21 seconds ($SD = 20.9$), compared to 16.44 seconds for non-tag retrieval ($SD = 12.59$). A paired t test showed that non-tag retrieval was significantly faster than tag retrieval, $t(118)=2.22$, $p<0.05$. When using tags, participants failed to retrieve 6 of their files, compared to only one failure when not using tags. A Wilcoxon related-samples non-parametric test showed results on the verge of significance, $W=1.89$, $p=0.059$.

Possible Age & Computer Literacy Effect

RQ8: Does age and computer literacy affect storage and retrieval behavior?

We used this question to test the validity of our findings. If, for example, younger participants tended to use tags or multiple classification more than older participants, this might suggest that our findings were confined to older users and that preferences change over time toward these behaviors, or that older users are less comfortable with the use of tags. To test the *Gmail* study results on this aspect, we conducted four Pearson correlation tests to examine the effect of age and computer literacy on (a) average percentages of tag-label usage (b) estimated tag-label search percentages (c) average percent of multiple classification (d) estimated search percentages. None of the results was significant.

Discussion

Overall, our results show a strong preference for folders over tags, for single over multiple classification, for hierarchical over flat storage and for location-based over non-location retrieval. In this section, we will discuss and attempt to explain each of these findings.

Folders vs. Tags

Results show a strong preference for folder over tag usage in PIM for both storage and retrieval. Regarding storage: in the *Gmail* study, 67% of all labeled messages were folder-labeled compared to 33% tagged-labeled (conforming with Rodden & Leggett, (2010)); moreover, the main reason for tagging behavior reported was the wish to keep the emails in the Inbox, rather than the need of multiple classifications.⁵ In the *Windows 7* study, 96% of the files were stored in specific folders (rather than default ones) and only 6 participants (16%) tagged any of their files during the test phase. Retrieval results are even more striking: In the *Gmail* study, an average of 16% of retrievals were estimated to be conducted by opening a label compared to only 3% using tag search. And in the *Windows 7* study, participants estimated that they used navigation for 61% of their retrievals compared to 5% for tag retrieval. Thus, in contrast to previous studies (see Table 1) that presented ambiguous results regarding the superiority of one or the other systems (in storage as well as retrieval), our findings are more unequivocal: based on their behavior, most users prefer folders for storage as well as for retrieval.

Behavior is very often determined by long term experience and habit. This, of course, might have affected participants' choices when one system or the other dominated their past experience. As half of the participants in the *Gmail* study used previous version of *Gmail* which offered tags only, more experience with folders is a weak argument supporting participants' choices. In an effort to overcome possible habit effect in the *Windows 7* study we successfully forced our participants to tag for two weeks prior to the experimental task. Nonetheless, we are aware that in spite of our efforts, habits still work

⁵ This behavior of putting aside unimportant messages in order to manage the Inbox as a to-do list was previously observed in Whittaker, Bellotti, & Gwizdka (2007), Whittaker et al. (2011) and Whittaker & Sidner (1996).

and participants choices may have still be at least partly affected by their past experience. As tags are now implemented more and more, and new users (without previous experience) are exposed to both systems, future studies will be less concerned with the habit effect.

Though we studied participants' behavior in their PIM environment, we made a choice regarding the particular interfaces for the study. This might limit possible generalizations to other interfaces. This is true regarding Windows 7 storage which does not offer the users a tag list (but does offer an autocomplete option), and in particular regarding the retrieval interfaces of both environments which did not include tag clouds and tag checkboxes.

Single vs. Multiple Classification

Several PIM researchers have criticized the hierarchic system for limiting users to a single classification (Bloehdorn & Völkel, 2006; Dourish et al., 2000; Heckner et al., 2009; Hsieh et al., 2008; Lansdale, 1988; Marsden & Cairns, 2003; Quan et al., 2003; Rodden & Leggett, 2010). However, our results indicate that when given the choice, users clearly prefer single classification and use little multiple classification. Regarding classification, in the *Gmail* study, 92% of the labeled emails, on average, had a single label compared to only 8% with multiple labels for tag-labeling participants. In the *Windows 7* study, only 6 participants used tags so no conclusions regarding multiple classification could be drawn (but perhaps if participants were more enthusiastic regarding multiple classification than more of them would have used tags) . Here too, results regarding retrieval are even more clear-cut: In the *Gmail* study, participants estimated that only one message out of a thousand was retrieved by using multiple labels and in the *Windows 7* study, participants (including the 6 who used tags for storage) estimated that they did not use multiple classification at all for retrieval.

There are several possible explanations for a single classification preference:

Multiple classification may be difficult and time consuming. This was suggested by our results: almost half of the participants in the *Windows 7* study indicated that tagging was either difficult or time consuming. If classification is cognitively challenging (Dumais & Landauer, 1983; Lansdale, 1988; Malone, 1983), then multiple classification may increase the difficulty. In situations where several possible

classifications immediately come to the user's mind multiple classification may indeed eliminate the cognitive cost of selection, however in other situations thinking of a single category may be simpler than attempting to think of all possible ones. In addition Pak et al. (2007) suggest that classification by tags is cognitively more demanding because they tend to be item-specific, where folder classification tends to be more generic. A related problem is tagging inconsistency (Civan et al., 2008; Gao, 2011).

Multiple classification may be redundant. Some of our participants as well as those reported in previous studies (Civan et al., 2008; Quan et al., 2003; Wash & Radar, 2007) felt that multiple tags were superfluous. Indeed, results of our *Gmail* study indicate that even when several labels were used for classification, participants tended to use the same label whenever they retrieved the information item. If a single classification is used in retrieval then the rest of the item's classification can be regarded as superfluous (an alternative view is presented at the end of this section).

Single classification is sufficient. Participants in both the *Gmail* and the *Windows 7* studies indicated that one category was sufficient for them to retrieve information items. In Bergman, Beyth-Marom, Nachmias et al. (2008), participants estimated that they remembered the exact location of 74-90% of their files on average (depending on the study), and in Bergman, Whittaker, Sanderson, Nachmias, & Ramamoorthy (2010), participants were able to retrieve 79% of their files without making any mistake along the way (an additional 15% of the files were eventually found using navigation, and only 4% of the files were not found at all). This is not perfect, but may be sufficient for the majority of users.

Multiple categorization makes exhaustive navigation inefficient. This was put forward by one of the participants in the study by Civan et al. (2008): “*You can exhaustively search a set of folders, you can search one folder and if it's not there, you search the next folder and so on. But you can't do the same thing with labels because they overlap so much and you don't want to look at the same things over and over again. That makes search inefficient and redundant.*” (p. 10).

In sum, we suggest that single classification is preferred because it is *simpler* than multiple classification. It is reasonable to assume that a 1 to 1 relationship (information item *a* is in folder *x*) is easier to remember than a 1 to many relationship (information

item *a* can be retrieved by using tags *x*, *y* or *z*). This was expressed by one of the participants in the *Windows 7* study: “*Multiple tags are somewhat cumbersome. One short name for a tag is something that can be remembered and eases retrieval.*” Future cognitive research should test this hypothesis in a controlled laboratory environment.

Simplicity may also explain the behavior of the minority of participants who used multiple classification for storage most of whom used one single tag for retrieval. Multiple classification may be chosen because participants do not rely on their memory and want to ensure easy retrieval (if not under *x* it might be under *y* or *z*) as two of our participants explained. Storing with a number of tags ensures retrieval with each one of them; hence, no multiple tagging is necessary when retrieving. Thus, even when using a single tag for retrieval, multiple classification is not necessarily redundant as it may increase the chances of successful retrieval.

Hierarchical vs. Flat Storage

Both studies indicated a preference for the hierarchical method over flat storage and non-location retrieval. Regarding storage, in the *Gmail* study, 79% of the participants indicated that they would prefer a hierarchical labeling system over the current flat one; in the *Windows 7* study, participants stored 86% of their files in a non-flat hierarchical location. Indeed, in April 2010, in its Lab section, Gmail presented nested labels – the ability to create a hierarchical label structure. The Gmail blog noted that “*A highly requested feature for labels, though, comes from the world of folders: the ability to organize labels hierarchically.*”⁶ Following users' feedback, nested labels became a standard feature in the new Gmail interface released in June 2011.

Location-based vs. Non-location Retrieval

On average, participants estimated that they used navigation for 69% of their retrievals (in the *Gmail* study) and 61% (in the *Windows 7* study) and used non-location retrieval for the rest (31% and 39%, respectively) indicating that introducing users to tags did result in non-location retrieval preference. These results are in line with previous research: In Jones et al. (2005), participants, even hypothetically, refused to relinquish

⁶ Retrieved from the Official Gmail Blog at <http://gmailblog.blogspot.com/2010/04/new-in-labs-nested-labels-and-message.html> on January 1, 2012.

their folders. Gao (2011) had stated that “*the hierarchical structure of categorization may encourage or force users to form a clear mental structure of the information*” (p. 826). A large scale research (Bergman et al., 2010) found that participants retrieved their files from hierarchic folders with very little use of default storage locations such as My Documents. In another large scale study (Bergman, Beyth-Marom, Nachmias et al., 2008), participants estimated that they navigated to their files in the majority of their retrievals (regardless of the search engine used). The paper also suggests possible reasons for navigation preference based on cognitive psychology research (p. 18-20). Finally, Teevan et al.(2004) found a strong preference to retrieve in steps from the general to the specific (teleporting). This is highly supported in folder navigation, but not in tag retrieval (Gao, 2011).

Retrieval Efficiency

Our controlled retrieval task results did not find tag retrieval to be more efficient than other forms of retrieval. On the contrary, it resulted in more retrieval failures and significantly slower retrievals. These results do not conform with (Hsieh et al., 2008; Quan et al., 2003) but are in line with Bergman et al. (2010), which found that participants managed to successfully navigate to 96% of the files from their Recent Documents list within 14.76 seconds on average. The efficiency of the location-based navigation method may be due to millions of years of evolution during which our predecessors navigated in the physical world. We are currently conducting a study to test the hypothesis that folder navigation uses the same cognitive brain structure as physical navigation (namely, the lateral part of the hippocampus).

Tags as “Cognitive Nodes in a Concept Network”

The results of our study indicate that with regard to PIM systems participants prefer folders over tags and single over multiple classification both for storage and retrieval. They also prefer location based over non-location retrieval. Perceived simplicity might explain these preferences. However, simplicity should according to Hsieh et al (2008) cause the opposite result – preference of tags over folders – based on their claim that tags are more compatible than folders with our cognitive system in which semantic processing is based on a flat-structure tag network (Collins & Loftus, 1975). Even if we

accept the tag-like model for *internal* representation, it does not logically derive that this is the best model for *externally* storing our personal information. We believe that there is confusion that should be avoided regarding objectives: The main aim of information item classification is not to externalize our internal representation of these items (Hsieh et al., 2008) or to fully describe them as implied from (Civan et al. 2008), but to support easy, fast and efficient retrieval. Therefore the comparison between methods should not focus on parameters such as which system bears more resemblance to our cognitive system or which gives a more detailed description of the information items. Rather, it should focus on relevant parameters such as users' preferences, retrieval speed and accuracy.

Conclusions

Tags are popular for Web 2.0 content sharing. However, this research shows that when it comes to their own personal information, users prefer folders over tags, and make very little use of multiple classification. Why do users, who use tags in a content-sharing Web 2.0 environment, prefer folders to tags when it comes to PIM? We think that the key issue here is familiarity. When looking for content which other users uploaded to the Web, users could not possibly know where it is located. Therefore, searching by tags seems a much better option. In PIM on the other hand, users are very familiar with their own information organization; after all, they stored it there according to their own subjective needs (Bergman, Beyth-Marom, & Nachmias, 2003, 2008). Moreover, they become more familiar with their organization scheme each time they navigate through it to retrieve their files. Therefore, in the great majority of cases, they are able to retrieve their own personal information quickly and efficiently (Bergman et al., 2010). Thus, maybe the transfer of the tagging approach from Web 2 to PIM systems is questionable.

Instead of borrowing ideas from Web 2.0 and applying them to PIM systems, we suggest a different approach - the user-subjective approach to PIM systems design (Bergman et al., 2003). This approach takes advantage of the fact that in PIM, the person who stores the information items is the same person who retrieves it later on. It suggests that PIM systems should use subjective (user specific) attributes given to the information item during the user-information interaction in order to assist future retrieval. The approach has received empirical support, and been used to develop several novel design schemes to improve navigation (Bergman, 2012; Bergman, Beyth-Marom, & Nachmias,

2008). Two such designs have already been positively evaluated (Bergman, Komninos, Liarokapis, & Clarke, 2012; Bergman, Tucker, Beyth-Marom, Cutrell, & Whittaker, 2009). Future research could compare between design approaches "imported" to PIM from other information fields, such as Web 2.0, with design approaches developed specifically for PIM based on the user-subjective approach.

Acknowledgments

We thank our participants for their time and efforts and Prof. Steve Whittaker for his continuous help. This research was partially funded by the European Union Marie Curie Grant, PERG-GA-2009-248997.

References

- Adrian, B., Sauermann, L., & Roth-Berghofer, T. (2007). In T. Pellegrini & S. Schaffert (Eds.), *Contag: A semantic tag recommendation system* (pp. 297–304). Paper presented at the ISEmantics' 07. JUCS.
- 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 Ph.D. dissertation (in Hebrew), Tel Aviv University, Tel Aviv, Israel.
- Bergman, O. (2012). The user-subjective approach to personal information management: from theory to practice. In Z. Marielba & d.O.J. Valente (Eds.), *Human-computer interaction: the agency perspective* (Vol. 396, pp. 55-81). Berlin / Heidelberg: Springer.
- 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. (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., Komninos, A., Liarokapis, D., & Clarke, J. (2012). You never call: Demoting unused contacts on mobile phones using DMTR. *Personal and Ubiquitous Computing*, 16(6), 757-766.
- Bergman, O., Tucker, S., Beyth-Marom, R., Cutrell, E., & Whittaker, S. (2009). In It's not that important: Demoting personal information of low subjective importance using GrayArea. Paper presented at the CHI 2009 Conference on Human Factors and Computing Systems, Boston, USA.
- Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R., & Ramamoorthy, A. (2010). The effect of folder structure on personal file navigation. *Journal of the American Society for Information Science and Technology*, 61(12), 2426-2441.
- Bloehdorn, S., & Völkel, M. (2006). Tagfs: Tag semantics for hierarchical file systems, 6th International Conference on Knowledge Management (I-KNOW 06). Graz, Austria.
- Boardman, R., & Sasse, M.A. (2004). In "Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management (pp. 583-590). Paper presented at the SIGCHI conference on Human Factors in Computing Systems, Vienna, Austria. ACM Press.
- Bruce, H. (2005). Personal anticipated information need. *Information Research*, 10(3).
- Capra, R.G., & Pérez-Quiñones, M.A. (2005). Using Web Search Engines to Find and Refind Information Computer, 38(10), 36-42.
- Civan, A., Jones, W., Klasnja, P., & Bruce, H. (2008). Better to Organize Personal Information by Folders Or by Tags?: The Devil Is in the Details., 68th Annual Meeting of the American Society for Information Science and Technology (ASIST 2008). Columbus, OH.
- Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428.
- Cutrell, E., Robbins, D.C., Dumais, S.T., & Sarin, R. (2006). Fast, Flexible Filtering with Phlat: Personal Search and Organization Made Easy. In CHI 2006 Conference on Human Factors in Computing Systems (pp. 261-270). Montreal, Canada: ACM Press.
- 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 Trans. Inf. Syst.*, 18(2), 140-170.

- Dumais, S.T., & Landauer, T.K. (1983). Using examples to describe categories. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 112-115). Boston, MA: ACM Press.
- Elsweiler, D., Baillie, M., & Ruthven, I. (2011). What Makes Re-finding Information Difficult? A Study of Email Re-finding. In Advances in Information Retrieval (Vol. 6611, pp. 568-579): Springer Berlin / Heidelberg.
- Furnas, G.W., Fake, C., von Ahn, L., Schachter, J., Golder, S., Fox, K., et al. (2006). Why do tagging systems work?, CHI '06 extended abstracts on Human factors in computing systems. Montreal, Canada: ACM.
- Gao, Q. (2011). An empirical study of tagging for personal information organization: Performance, workload, memory, and consistency. *International Journal of Human-Computer Interaction*, 27(9), 821-863.
- Heckner, M., Heilemann, M., & Wolff, C. (2009). Personal Information Management vs. Resource Sharing: Towards a Model of Information Behaviour in Social Tagging Systems, Third International AAAI Conference on Weblogs and Social Media, ICWSM-09 (pp. 42-49). San Jose/CA: epub:6820.
- Hsieh, J.L., Chen, C.H., Lin, I.W., & Sun, C.T. (2008). A Web-based tagging tool for organizing personal documents on PCs, International Conference of Computer-Human Interaction 2008. Florence, Italy.
- 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. Portland, OR: ACM.
- Kidd, A. (1994). The marks are on the knowledge worker In Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence (pp. 186-191). Boston, MA: ACM Press.
- Kirk, D., Sellen, A., Rother, C., & Wood, K. (2006). In Understanding photowork. Paper presented at the SIGCHI conference on Human Factors in Computing Systems, Montreal. ACM.
- Lansdale, M.W. (1988). The psychology of personal information management. *Applied Ergonomics*, 19(1), 55-66.
- Ma, S., & Wiedenbeck, S. (2009). File management with hierarchical folders and tags, Proceedings of the 27th international conference extended abstracts on Human factors in computing systems. Boston, MA, USA: ACM.
- 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.

- Marsden, G., & Cairns, D.E. (2003). Improving the usability of the hierarchical file system, 2003 annual research conference of the South African Institute of Computer Scientists and Information Technologists on Enablement through Technology: South African Institute for Computer Scientists and Information Technologists.
- Oleksik, G., Wilson, M.L., Tashman, C., Rodrigues, E.M., Kazai, G., Smyth, G., et al. (2009). Lightweight tagging expands information and activity management practices, 27th international conference on Human Factors in Computing Systems. Boston, MA, USA: ACM.
- Pak, R., Pautz, S., & Iden, R. (2007). Information organization and retrieval: An assessment of taxonomical and tagging systems. *Cognitive Technology*, 12(1), 31-44.
- Quan, D., Bakshi, K., Huynh, D., & Karger, D.R. (2003). User Interfaces for Supporting Multiple Categorization. In M. Rauterberg (Ed.), *Proc. of INTERACT 2003* (pp. 228-235). Amsterdam: IOS Press.
- Robbins, D.C. (2008). TapGlance: designing a unified smartphone interface for personal information management, CHI 2009 Conference on Human Factors and Computing Systems Florence, Italy.
- Rodden, K., & Leggett, M. (2010). Best of both worlds: improving gmail labels with the affordances of folders, Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems (pp. 4587-4596). Atlanta, Georgia, USA: ACM.
- Sajedi, A., Afzali, S.H., & Zabardast, Z. (2012). In Can you retrieve a file on the computer in your first attempt? Think to a new file manager for multiple categorization of your personal information. Paper presented at the PIM12 CSCW 2012 Workshop, Seattle, WA, USA.
- Sauermann, L., Grimnes, G., Kiesel, M., Fluit, C., Maus, H., Heim, D., et al. (2006). Semantic Desktop 2.0: The Gnows Experience. In *The Semantic Web - ISWC 2006* (Vol. 4273, pp. 887-900). Berlin / Heidelberg: Springer.
- 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, CHI '08 conference on Human Factors in Computing Systems. Florence, Italy: ACM.
- Teevan, J., Alvarado, C., Ackerman, M.S., & Karger, D.R. (2004). In The perfect search engine is not enough: a study of orienteering behavior in directed search (pp. 415-422). Paper presented at the SIGCHI conference on Human Factors in Computing Systems, Vienna, Austria. ACM Press.
- Voit, K., Andrews, K., & Slany, W. (2012). Tagging might not be slower than filing in folders, Proceedings of the 2012 ACM annual conference extended abstracts on

- Human Factors in Computing Systems Extended Abstracts. Austin, Texas, USA: ACM.
- Wash, R., & Radar, E. (2007). Public bookmarks and private benefits: An analysis of incentives in social computing. *Proceedings of the American Society for Information Science and Technology*, 44(1), 1–13.
- Whittaker, S., Bellotti, V., & Gwizdka, J. (2007). Everything through email. In W. Jones & J. Teevan (Eds.), *Personal information management* (pp. 167-189). Seattle: University of Washington Press.
- Whittaker, S., Matthews, T., Cerruti, J., Badenes, H., & Tang, J. (2011). In Am I wasting my time organizing email? A study of email refinding. Paper presented at the Conference on Human Factors in Computing Systems, Vancouver.
- 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). Vancouver, British Columbia, Canada: ACM Press.

Appendix A

Gmail Questionnaire

Hello,

You are invited to participate in a research project. The purpose of this anonymous questionnaire is to collect data for a scientific research. The questionnaire relates to your email management strategies using the Gmail interface: to the way you store and retrieve your email messages using this interface regarding the possibilities presented to you in the informative video that you had watched a few weeks ago. At the end of the research you will receive a comprehensive explanation. Thank you for your cooperation.

Questions Regarding Marking/Classifying with Labels

1. To what extent do you use labels in Gmail to mark your messages?
 - a. I use labels for most or all my messages
 - b. I use labels for some, but not most, of my messages
 - c. I usually don't use labels to for my messages
 - d. I almost never use labels

(If you chose answer d please go to question 4)

2. How long do you estimate that you use labels in Gmail? _____ months

3. As was explained to you in the video you have two ways of marking/classifying your email messages using labels (after you had already created the labels): in the first way you mark the message by adding tag/s to your email message, then you can find the message in the Inbox and in each of the labels assigned to the message (in the label list located on the left side of the messages). In the second way you move the message from the Inbox into one of the labels (located on the left side of the messages) then you can find it in the label but not the Inbox. What do you do **in the majority of cases?**

- a. Add a label/s to the message
- b. Move the message from the Inbox into one of the labels

Please explain this preference

The Use of Gmail Interface

4. The Gmail interface allows you to go back to messages that you have seen in the past (retrieved them) using several options:

- (a) By scanning the Inbox
- (b) By opening a label in which it was located and scanning its content (e.g. opening the label "studies" and looking for the message you wanted regarding studies)
- (c) By using the search option without using labels
- (d) By using the search option with one label (e.g. typing "label:studies" in the search box to find a letter regarding this).
- (e) Using the search option with two or more labels (e.g. typing "label:studies , label:Yael" in the search box, in order to find a message regarding studies sent by Yael).
- (f) In other ways.

Think of all the occasions on which you retrieved your Gmail messages (e.g. in the last two weeks). Please try to estimate the percentage of occurrences when you used each of these methods.

| | |
|--|---------|
| a. By scanning the Inbox | _____ % |
| b. By opening labels and looking inside them | _____ % |
| c. Searching using the search box without using labels | _____ % |
| d. Searching by typing a single label in the search box | _____ % |
| e. Search by typing two labels or more in the search box | _____ % |
| f. Other (please specify) _____ | _____ % |
| Total | 100% |

5. Did you ever assign two labels or more to a message? Yes/no (if you answered negatively go to question 7)

6. When you assign more than a single label to a message you usually –

- a. Attempt to locate it in the Inbox
- b. Attempt to locate it using a single label, i.e. consistently go to the same label (in the labels list), "open" it and attempt to locate the message inside it.
- c. Attempt to locate the message using different labels at different retrievals, i.e. go to different labels (in the labels list) "open" it and attempt to locate the message inside it.
- d. Search for it using the search box without using tags
- e. Search for it each time with the same label
- f. Search for it with different labels on different occasions
- g. Other (please specify) _____

7. The current Gmail design allows you to apply any number of labels to a message, however, these labels *are not* ordered hierarchically. Try to think of a label system where the labels can be organized hierarchically. For example, a student can have a “courses” label and below it sub-labels with the names of those courses (“introduction to psychology,” “social psychology,” etc.). In that case, if a message is labeled “introduction to psychology,” it would also automatically be labeled “courses.” This

message would be found in both places, in the “courses” label (where all the messages related to courses can be found) and in the “introduction to psychology” label (where all the messages related to this course can be found). If we would like to retrieve a course related message and we can’t recall the specific course, then we can look for it in the “courses” label. And if we remember the name of the course, we can look for it in the specific label. If you had the choice, which system would you prefer?

- (a) A label system that allows for a hierarchy of labels
- (b) I have no preferences
- (c) A label system without hierarchy (the current system)

General Information

We conclude with a few general questions

8. Year of birth _____

9. Gender

- a. Male
- b. Female

10. I am an experienced computer user

- a. Strongly disagree
- b. Disagree
- c. Neither agree nor disagree
- d. Agree
- e. Strongly agree

11. I am familiar with the Gmail interface

- a. Strongly disagree
- b. Disagree
- c. Neither agree nor disagree
- d. Agree
- e. Strongly agree

Appendix B

Windows 7 Study - Retrieval Questionnaire

In the Windows 7 retrieval questionnaire we first asked participants to assess the percentage of times that they used all Windows 7’s basic retrieval options (navigation, shortcuts, search, recent documents and the “arrange by” option) adding up to 100%. As there are several ways to retrieve files using tags (using navigation, search or the “arrange

by” option), we asked the participants additional estimation questions regarding the use of tags in each of these options.

This questionnaire relates to the way in which you retrieve files. In all questions we will present possible retrieval methods and ask you to estimate your frequencies of use in each of these possibilities using percentages. We will also leave you space to specify another method, if there is one. In order for you to give us realistic estimations we ask you to **first** read about all the possible methods, specify your additional method if there is one and only then divide the percentages.

1. Windows 7 operating system allows you to retrieve files that you had save on your computer in various ways: a. **Navigation** to the folder in which the file is located, and then finding the required file in the folder's list of files, b. clicking on a **desktop shortcut** if you had previously created one, c. **searching for it** using the Widows 7 search engine d. finding it in the **recent documents list** of any of the Office programs, e. using the **Arrange by** option, or f. other options. Think of the occasions in which you had retrieved your office documents during the recent month. Try to estimate the percentages in which you had used any of these options (remember to read about all possibilities first, fill section e if relevant, and only then divide the percentages).

| | |
|--|---------|
| a. Navigation to the folder in which the file is located, and then finding the required file in the folder's list of file | _____ % |
| b. Pressing a desktop shortcut | _____ % |
| c. Searching for it using the Widows 7 search engine | _____ % |
| d. Finding it in recent documents list | _____ % |
| e. Using the Arrange by option | _____ % |
| f. Other (please specify) | _____ % |
| Total | 100% |

2. We will now relate only to files that you had retrieved by using **navigation** (question 1, option a). These will be your 100% for this question. At the end of the navigation when you reach the folder containing the required file, you can either sort the folder or use the

default sort (alphabetical sorting of file names). Think of all the navigations that you had made during the recent month and try to estimate the percentage in which you had changed the sorting.

a. I had *changed the sorting* of _____% of the files I retrieved by using navigation (if you answered 0% skip to question 4)

b. I had used the *default alphabetical sorting* for _____% of the files I retrieved by using navigation

Total 100%

3. We will now relate only to files you retrieved by **navigation while changing the sorting of the designated folder** (question 2, option a). These will be your 100% for this question. You can sort your files by: date, format, size, tags or use other sorting. Think of all the navigations while changing the sorting that you had made during the recent month and try to estimate the percentages in which you had used each of the following sorting (remember, first read about all possibilities, fill in section e if relevant, and only then divide the percentage):

a. Sorting by *date* _____% of navigations

b. Sorting by *format* _____% of navigations

c. Sorting by *size* _____% of navigations

d. Sorting by *tags* _____% of navigations

e. *Other sorting* (please specify) _____, _____% of navigations

Total 100%

If you do not use the *Arrange by* option (i.e. wrote 0% in option e of question 1), skip to question 5.

4. We will now relate only to documents which you had reached by using the **Arrange by** option (question 1, option e). These will be your 100% for this question. As you had learned, it is possible to change the default option (which is by folders) to arrange your files by: date, format, size, tags or other attributes. Think of all the files that you had retrieved using *Arrange by* during the recent month and try to estimate the percentage in

which you had used the following options (remember, first read about all the options and only then divide the percentages).

- a. Arrange by *date* ___% of the retrievals using *Arrange by*
- b. Arrange by *name* ___% of the retrievals using *Arrange by*
- c. Arrange by *format* ___% of the retrievals using *Arrange by*
- d. Arrange by *size* ___% of the retrievals using *Arrange by*
- e. Arrange by *tags* ___% of the retrievals using *Arrange by*
- f. Arrange by *other attributes* ___% of the retrievals using *Arrange by*

Total 100%

5. We will now relate only to files that you had retrieved by using **search** (question 1, option c). These will be your 100% for this question. Search possibilities include: searching by file name, search using a word contained in the file, search using date, search using author name, search using format, search using file size, search using tags and search using other file attributes. Think of all the files you had retrieved by using search during the recent month and try to estimate the percentages in which you had used each of these eight options:

- a. Search using *file name* _____% of search retrievals
- b. Search using a *word contained in the file* _____% of search retrievals
- c. Search using *date* _____% of search retrievals
- d. Search using file *author name* _____% of search retrievals
- e. Search using *format* _____% of search retrievals
- f. Search using file *size* _____% of search retrievals\
- g. Search using *tags* _____% of search retrievals
- h. Search using *other file attributes* (specify)_____, _____% of search retrievals

If you did not search by using tags (i.e. answered 0% in option g of question 5) than there is no need for you to answer question 6.

6. In the last question we will relate only to tags which you had **searched using tags**. These will be your 100% for this question. When searching by using tags, you can either

use a single tag or you can use multiple tags (two tags or more) in order to find the required file. Think of all of the files that you retrieved during the recent month and try to estimate the percentages in which you had used each of these options:

a. Using a *single tag* _____% of the search retrievals using tags

b. Using *two tags or more* _____% of the search retrievals using tags

Total 100%