Best of Both Worlds: Improving Gmail Labels with the Affordances of Folders

Kerry Rodden

Google

1600 Amphitheatre Parkway

Mountain View, CA 94043 USA

krodden@google.com

Michael Leggett

Google

1600 Amphitheatre Parkway

Mountain View, CA 94043 USA

mleggett@google.com

Abstract

Gmail's filing system for email conversations is based around labels, which are more flexible and powerful than folders. With its original user interface, many users did not discover labels, and wondered why Gmail had no folders. The Gmail team redesigned the user interface for labeling to make it more discoverable and understandable, and to add the most useful functionality of folders. The new design works for the simple use case (a conversation with only one label), while still making the more complex use case (multiple labels) easily available. It has been launched to millions of users worldwide and has resulted in much higher adoption of labels, especially by new users of Gmail.

Keywords

Email organization, personal information management, folders, labels, tags

ACM Classification Keywords

H.5.2 [Information interfaces and presentation]: User Interfaces—evaluation/methodology, user-centered design

General Terms

Design, Human Factors, Measurement

Introduction and related work

Email is used by hundreds of millions of people throughout the world, for both business and social communication. Email systems share a common workflow: new mail arrives in an inbox for the user to read and perhaps act upon. When done, the user has to choose what to do with each message – for example, leave it in the inbox, delete it, or file it.

Previous research studies have found that while users' email organization practices vary widely [10] [13], the majority of users choose to file at least some of their email. For example, one survey study of work email users found that 67% reported filing at least weekly [12], and another found that 27% of messages were filed [5].

The traditional filing method is to use *folders*. Email messages can be moved into a folder, which simultaneously removes them from the inbox. Studies have found many issues with folders as they are commonly implemented. An email may naturally belong in multiple folders, but since it can only exist in one folder at a time, the user is forced to decide on a single location. This can be both inefficient [2] and overwhelming [4]. One solution to this problem is to introduce *labels* [12] or *tags* [11], and enable multiple labels or tags to be applied to the same item.

Google's Gmail web application [8] uses labels [9]. When it launched, Gmail also implemented a number of other non-traditional features in response to problems with the traditional workflow. First, Gmail groups related messages together into *conversations* (or threads), containing an initial message plus any subsequent replies with the same subject line. In traditional email systems, these messages are separated from each other and may appear all over the inbox. Second, Gmail implemented an *archive* feature, which removes a conversation from the inbox without deleting it. This saves users the trouble of having to decide where to file the conversation, making it available for re-finding via search. The Gmail team did not implement folders, expecting that they would be redundant given the low effort involved in using "Archive", plus the availability of labels for those who still wanted to explicitly file conversations. Labeling and then archiving a conversation is equivalent to moving it to a folder.

However, many users did not discover the label-and-archive combination, and so did not realize that Gmail offered any way to file conversations [3]. Some new users would abandon Gmail altogether because it was apparently lacking their familiar folders [11]. Other users simply found label-and-archive inefficient, given that it requires two steps instead of one. In feedback from customer support channels and surveys, one of the most popular feature requests was to add folders to Gmail.

Clearly, the original implementation of labels was not meeting the needs of all users. 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.

So the Gmail team redesigned the user interface for labeling. The goal was to add the basic functionality of folders, without confusing or annoying users who were happy with labels, and without introducing too much

complexity or redundancy into the user interface. The ideal solution would combine the simplicity and familiarity of folders with the flexibility and power of labels.

Bluemail [11] is the only other email system we are aware of that has taken on a similar challenge. Its designers chose to add both conventional folders and a tagging interface, keeping the folder and tag names in two separate lists. The Gmail team started out in a similar direction, but, following a number of design iterations, ended up launching a more integrated solution.

Combining folders and labels

Because of the existing users who were happy with labels, it was not an option for the team to simply switch to folders. The team considered changing the name "labels" to "tags", given the wide use of this term elsewhere on the web. However, early usability studies found that many email users were confused by this term. This was not because of its connotation with social sharing, but rather because users who were more familiar with email client applications (such as Outlook) believed that tags were primarily related to filtering or prioritization of email, rather than filing. The team decided that the term "tags" offered no significant advantage over "labels", and switching was not worth the risk of confusing existing users of labels.

The team's early designs incorporated both folders and labels, as separate lists. These were implemented in prototypes that were iterated on via usability studies. Separate lists worked well from a usability perspective: users who were looking for folders or labels were able to find each of them when needed. But it became

increasingly clear to the team that many users did not see a strong distinction between the concept of a folder and the concept of a label. It is also relatively rare that a user will apply more than one label to a conversation [3]. The team was concerned that, over time, the two sets of names would overlap, giving users even more choices to make than before. For example, when creating a new category, they would also have to decide whether it should be a folder or a label.
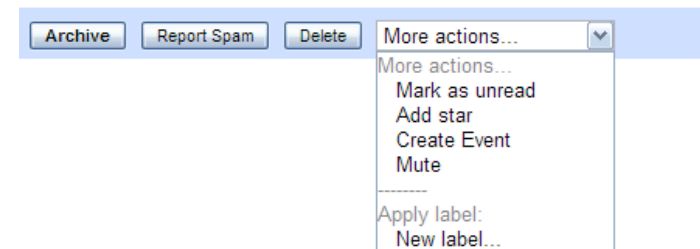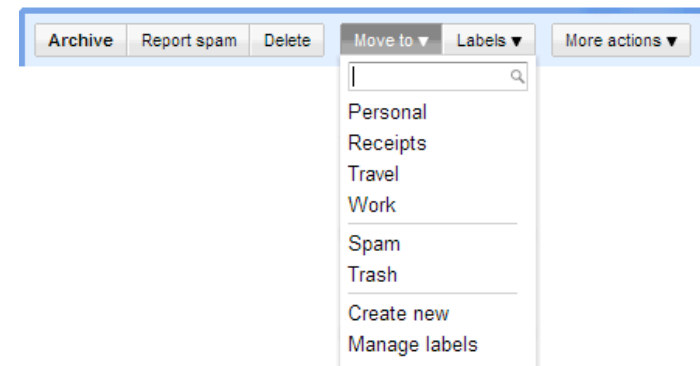

Figure 1: The original way of applying a label in Gmail (top) and the new "Move to" drop-down (bottom).

The main perceived advantage of folders is that they make it easier to get mail out of the inbox [3]. The team realized that they could provide this benefit by simply making it easier to label-and-archive in a single step. They designed a "Move To" drop-down (Figure 1), which both labels and archives an inbox conversation in one action. (More generally, it removes the label corresponding to the current view and applies the new one. This works for "Inbox" because it is a special label applied to new mail by the system.) It can be used either from within a conversation, or when viewing a list of conversations, by first selecting one or more of them. A second drop-down, called "Labels", was added for the benefit of people who were already using labels, as well as to provide more fine-grained control such as adding or removing several labels at once.

These two drop-downs combined to make labeling more visible and more easily accessible than before, when the menu option to apply a label was hidden behind a drop-down called "More actions" (see the top part of Figure 1).

The addition of "Move to" meant that the team did not need to list both folder names and label names in the user interface, and could simply continue to use a list of label names. This list was originally titled "Labels" (Figure 2), but this was confusing to users who were looking for folders and did not recognize the term "labels", leading to the perception that they could not file their mail in Gmail.
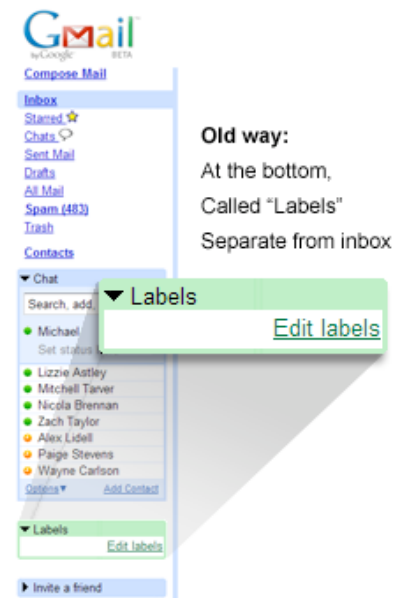

Figure 2: The old left navigation bar.

The team therefore decided that the list of label names would not have a title (Figure 3). The idea was that an untitled list of labels in a familiar location in the user interface would have the right affordances to match the expectations of users, regardless of whether they were looking for folders or labels. To help encourage this understanding, the team decided to give everyone without labels (including new users of Gmail) a set of 4 of the most popular: "Personal", "Work", "Travel", and "Receipts".
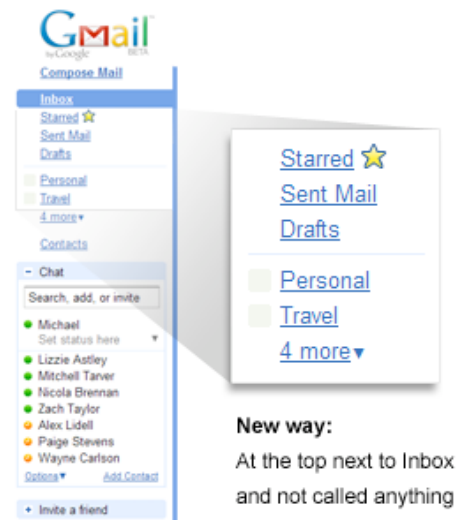
Figure 3: The new left navigation bar.

Because some existing users of labels may have had a long list of them, some of which they were not using actively, the team also designed a one-time process for launch that hid unused labels behind a drop-down ("8 more" in Figure 4). Users can also choose which labels will be always visible and which ones will be hidden. This makes it easier for users to create and apply labels without having to worry about how many they have or the space they are taking up, which is often considered to be one of the advantages of tags.
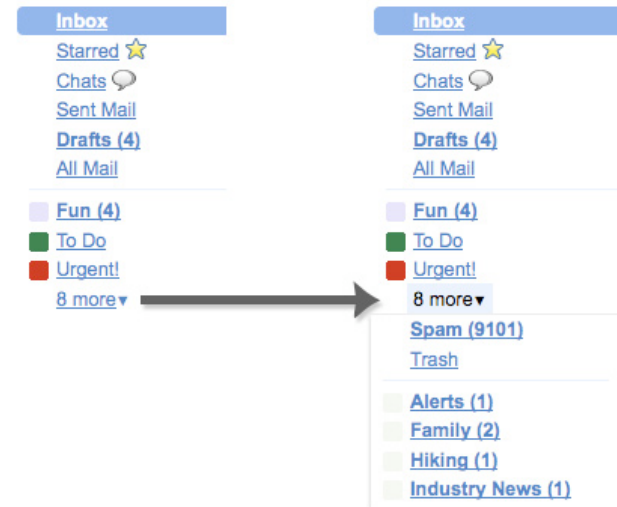

Figure 4: Hiding some labels behind a drop-down.

Finally, the team added two more features to make labeling more convenient: drag and drop, and keyboard shortcuts. Dragging and dropping an email onto a name is equivalent to using "Move To" (see Figure 5), and dragging a name onto an email is equivalent to using "Labels" (see Figure 6). Keyboard shortcuts can improve efficiency of labeling for users who choose to use them.
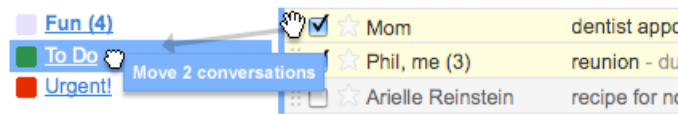
Figure 5: Drag and drop - conversation to label.
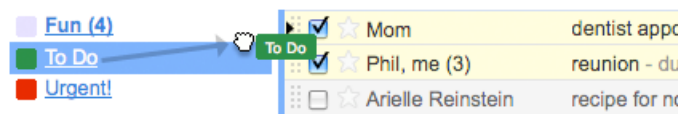
Figure 6: Drag and drop - label to conversation.

Other aspects of filing remain the same as before these design changes. For example, when new messages arrive that are part of a filed conversation, that conversation will re-appear in the inbox, but will still have any previously applied labels. The user can then simply archive the conversation to get it out of the inbox while still retaining the previous labels. Deleting a conversation deletes it from everywhere, not just from the label (because there is still only one copy of the conversation), and deleting a label does not delete the conversations inside it.

One refinement made as a result of user studies was with regard to displaying the labels that have been applied to a given conversation. In the new design, Gmail will show every label that a conversation has, in whatever context it is being viewed. For example, when viewing the list of all conversations with the label "Travel", "Travel" is shown before the subject line of each. In previous versions, this label would have been hidden, because showing it on every conversation seemed redundant. This was changed after user studies, because the lack of consistency was confusing

to users – if a conversation is shown to have the "Travel" label in one view, this label should not be missing from other views.

With this new design, people who are looking for a simple, folder-like way of filing email can use "Move To" and/or drag conversations to names. Those who want to do more complex filing, such as applying multiple labels to a single conversation, can use "Labels" and/or drag names to conversations.

Iteration and launch

These new concepts were subjected to several rounds of usability testing while under development (with findings incorporated into the design as it progressed, as described in the previous section), and iterated on before being completely implemented. They were then launched internally, so that the thousands of users inside Google could test the new features in their own work email accounts. The team gathered feedback from internal users via surveys. After internal testing and further refinement, the team was confident in the basic usability and usefulness of the new features, and chose to launch them to 100% of users.

For technical reasons, the designs were launched in two phases. Phase 1 launched in February 2009 and included the "Move To" and "Labels" drop-downs, as well as keyboard shortcuts for these actions [6]. Phase 2 launched in July 2009 and included the new left navigation bar, plus drag and drop [7]. Because the phase 2 launch involved moving sections of the interface around, it was announced to all users via a one-time call-out that they could dismiss (see Figure 7).
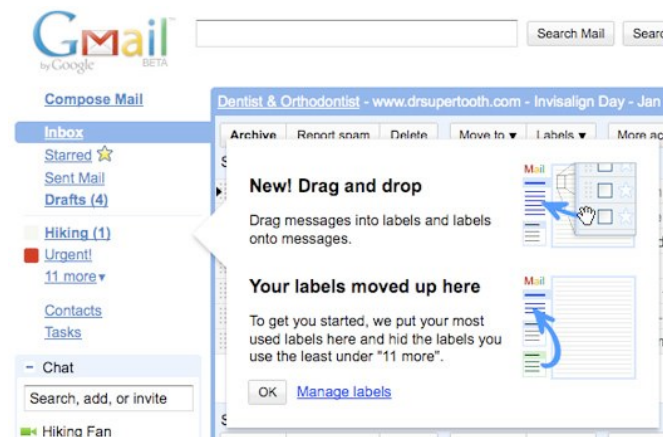
Figure 7: Call-out to explain the changes to the left navigation bar.

Post-launch evaluation

The final step was to analyze the outcome of the external launch to millions of users around the world, and in particular, to analyze aggregated usage data[1] to determine whether users were actually adopting and using labels.

This final step is one that is usually missing from published accounts of user interface innovations in email organization, which tend to be based on small samples of users in research labs. It is hard to evaluate such innovations on a large scale, because experimental email systems tend not to be adopted by

---

[1] All analysis was done in accordance with the Google and Gmail Privacy Policies (http://www.google.com/privacy). The data was anonymized and also aggregated across a large number of users, not analyzed on an individual basis. The data sets were filtered according to dates of usage, and account creation date. No message content or contact information was analyzed.

large numbers of users, especially outside of a work context.

There are three main ways to use labels in Gmail: applying them when organizing email, clicking on them when later trying to re-find conversations, and clicking on them when they contain unread mail (to which the label has been automatically applied, by a filter). Because these design changes were mostly targeted at the organization phase, we focused our analysis on applying labels.

Overall usage
We tracked usage of label-related actions over time, from before the phase 1 launch until the time of writing. At each time point, we considered only actions done by users who were 7-day active (i.e. had been active in Gmail on at least one of the past 7 days). The set of label-related actions included applying labels, creating labels, and clicking on labels. We calculated the percentage of 7-day active users who had done each of these actions at least once in the past 7 days, as well as the average number of times per user that each action was done.

We saw a large relative increase in the percentage of 7-day active users who applied at least one label. Following the phase 1 launch ("Move to" and "Labels" drop-downs), this number was consistently about 20% higher than before. Following the phase 2 launch (the new left navigation bar) it is now about 50% higher than it was before any of the design changes. The overall number of labels applied has increased substantially, while the average number applied per user has stayed about the same.
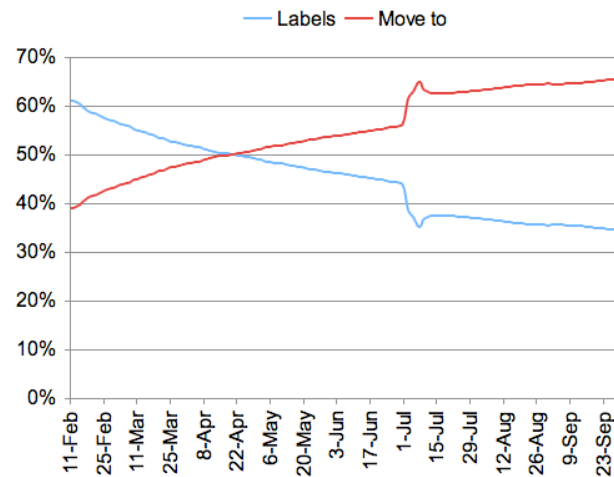
Figure 8: Percentage of labeling actions done via the "Labels" and "Move to" drop-downs, from just after their launch (in Phase 1) to the time of writing. "Move to" is now used almost twice as often as "Labels". The spike in the graph corresponds with the launch of Phase 2 (the new left navigation bar).

Over time, relative usage of the new "Move to" drop-down has surpassed that of "Labels" (see Figure 8). We believe that this is because the "Labels" drop-down was immediately adopted by existing users of labels, while those who were new to labeling discovered it via "Move to". Over time, existing users of labels are discovering that "Move to" is more convenient than a two-step label-and-archive.

Drag and drop and keyboard shortcuts are being used by only a small percentage of users at present, so these additions have not contributed substantially to the findings mentioned above.

To examine possible errors in labeling, we also analyzed usage of the undo action following application of a label [1], in relation to the number of labeling actions performed. We found that usage of undo rose slightly immediately following each of the two launches, but soon dropped back to its previous level. We attributed the temporary increase to users experimenting with the new features.

We also found that more users are clicking on user-created labels, especially with the phase 2 launch, which moved the list of labels higher up on the page. The percentage of 7-day active users clicking on a user-created label is now double what it was prior to this launch.

Impact on new users of Gmail
It was particularly important to assess the impact of these features on new users of Gmail, given the team's goal of encouraging them to understand and discover that Gmail does offer a folder-like functionality. In addition, the extensive testing on internal users did not give much opportunity to examine impact on new users. We did a more detailed analysis of whether users were applying labels in their first two weeks after signing up for Gmail.

After phase 1 was launched, new users were consistently about 50% more likely to apply a user-created label during their first 2 weeks of usage. After phase 2, this has risen to a 100% increase – so new users of Gmail are about twice as likely to try labels than they were prior to the new designs. As with all users, the average number of labels applied per new user has stayed about the same.

Feedback from other sources
Of course, it is important to triangulate these findings with input from other sources such as customer support channels, and satisfaction surveys. These have also been largely positive. The major source of complaints was the fact that a related optional feature (allowing the list of labels to be moved to the right-hand side of the page, instead of the left) had to be disabled for technical reasons. There was also some confusion about what had happened in cases where system labels (like Spam and Trash) were automatically hidden, in the phase 2 launch. The team will continue to monitor feedback and use it to inform future design decisions.

Discussion and future work
Despite the growth in adoption of labels following these design changes, there are still a substantial number of Gmail users who never use labels. The team's goal is not to persuade every Gmail user to use labels, especially because Gmail has always placed a strong emphasis on using search to quickly re-find conversations without the need for organization. Instead, the team hopes to ensure that every Gmail user who has the need for this feature is able to discover and use it successfully. To establish a meaningful target for usage, it would be useful to know the approximate size of this group of users.

Further qualitative studies, and in particular, longitudinal field studies, will be very helpful in gaining a deeper understanding of how these new designs are being used. For example, we would like to understand more about how the folder-like and label-like features are used together. Also, users' organizational strategies are likely to change over time, as they receive and organize more mail [2]. 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. It would also be useful to be able to measure how many conversations have single or multiple labels.

The quantitative evaluation reported here relates mostly to organization of email, and not to subsequent re-finding of those conversations. From other sources (e.g. usability studies, customer support), the team has no reason to suspect that users are having any problems here, but it would be useful to be able to monitor this over time. From usage data alone, it is hard to distinguish successful and unsuccessful label clicks, but it may be possible to make more progress in this area, e.g. perhaps several clicks on different labels in quick succession would be an indicator of not finding the desired item. Again, further qualitative research would be helpful here.

This simple folder-like implementation is missing one common feature of folders – the ability to organize them into a hierarchy. Some users like to use hierarchical organizations as a way of expressing the relationships between categories of items [3].   But hierarchies have also been found to be problematic, particularly at the re-finding stage – for example, items become harder to find in deeper categories [2] – and so it is still unclear whether Gmail should endeavor to provide them.

The team expects that the pattern of combining folders and labels will generalize beyond Gmail, as well as beyond the specific domain of email. Other

implementations will provide the opportunity to study this.

Conclusions

The Gmail team has advanced the state of the art in user interfaces for email filing. They have launched a design that provides the most important affordances of both labels and folders, giving users the flexibility to use them in either way, depending on their needs. The new design is being used successfully by millions of users around the world, and in particular, it has resulted in new Gmail users being twice as likely to discover its filing features. The design works in the simple use case (a conversation with only one label), while still making the more complex use case (multiple labels) easily available. A key insight in the design process was to avoid loaded terminology and focus on affordances. The team expects that this pattern will be generalizable beyond Gmail, as well as beyond the specific domain of email.

Acknowledgements

References

[1]   Akers, D., Simpson, M., Jeffries, R., and Winograd, T. Undo and erase events as indicators of usability problems. *Proc. CHI'09*, ACM, 2009, pp. 659-668.

[2]   Bälter, O. Keystroke level analysis of email message organization. *Proc. CHI'00*, ACM, 2000, pp. 105-112.

[3]   Civan, A., Jones, W., Klasnja, P., and Bruce, H. Better to organize personal information by folders or by tags?: The devil is in the details. *Proc. ASIS&T Annual Meeting*, Wiley, 2008.

[4]   Dabbish, L.A., and Kraut, R.E. Email overload at work: An analysis of factors associated with email strain. *Proc. CSCW'06*, ACM, 2006, pp. 431-440.

[5]   Dabbish, L.A., Kraut, R.E., Fussell, S., and Kiesler, S. Understanding email use: Predicting action on a message. *Proc. CHI'05*, ACM, 2005, pp. 691-700.

[6]   Eklund, E.A. New ways to label with "move to" and auto-complete. *Official Gmail Blog*, February 2009. http://gmailblog.blogspot.com/2009/02/new-ways-to-label-with-move-to-and-auto.html

[7]   Gajda, D. Labels: Drag and drop, hiding, and more. *Official Gmail Blog*, July 2009. http://gmailblog.blogspot.com/2009/07/labels-drag-and-drop-hiding-and-more.html

[8]   Gmail. http://mail.google.com/

[9]   Leggett, M. The evolution of Gmail labels. *Official Google Blog*, July 2009. http://googleblog.blogspot.com/2009/07/evolution-of-gmail-labels.html

[10] Mackay, W. Diversity in the use of electronic mail: A preliminary inquiry. *ACM Transactions on Office Information Systems*, Vol. 6, No. 4 (1988), pp. 380-397.

[11] Tang, J.C., Wilcox, E., Cerruti, J.A., Badenes, H., Schoudt, J., and Nusser, S. Tag-it, snag-it, or bag-it: Combining tags, threads, and folders in e-mail. *Ext. Abstracts of CHI 2008*, ACM, 2008, pp. 2179-2194.

[12] Venolia, G., Dabbish, L., Cadiz, J.J., & Gupta, A. *Supporting email workflow*. MSR Tech Report, MSR-TR-2001-88, 2001.

[13] Whittaker, S. and Sidner, C. Email overload: Exploring personal information management of email. *Proc. CHI'96*, ACM, 1996, pp. 276-283.