Merging Methodologies:
Combining Individual and Group Card Sorting

Robert L. Thomas and Ian Johnson

Liberty Mutual, Personal Insurance, Boston, MA
{robertl.thomas,ian.johnson}@liberty.mutual.com

Abstract. This paper presents a case study detailing how we combined individual card sorts with focus groups and group card sorting to improve the content hierarchy and organization of www.libertymutual.com, the personal insurance website of Liberty Mutual, which customers can visit to get an insurance quote, service their insurance policies, or find insurance-related information. We analyzed quantitative and qualitative data from 26 participants, on which we based our recommendations for a new hierarchy and site structure. Our paper will show how the results from the individual and group sorts differed, how the individual exercise informed the group exercise, and how the group exercise informed the recommendations. We believe this combination of individual sorting, group sorting, and focus group discussion makes this methodology unique.

Keywords: Card sorting, design methodology, information architecture, usability testing, user-based testing, content hierarchy, content organization.

1   Introduction

1.1   Background

The Liberty Mutual web site, www.libertymutual.com, had progressed from a small site with limited content to a large site with multiple lines of business as well as supporting information, insurance resources, and tools. The navigation suffered from having a limited number of options. It did not have a traditional navigation bar with drop-down options; it simply displayed nine categories in the top right corner of each page (see Figure 1). Four of those nine categories took users to a different website. We were also unsure whether some of the options, such as Member Rights, were top-level items that users would most gravitate toward.
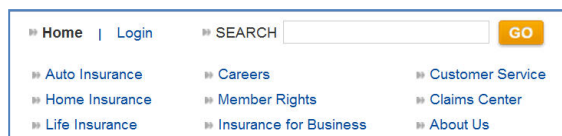

Fig. 1. Navigation from previous www.libertymutual.com website

1.2   Goals

We needed to create a new navigational hierarchy that would enable our users to more easily find specific options under main categories, and enable our team to provide new categories for supporting information. Our high-level objectives were to:

- Determine how current content hierarchies and labeling met users' needs
- Make the navigation labels understandable
- Make content discoverable
- Make the site attractive to users

In terms of visual hierarchy, we knew from our usability tests that participants scrolled down the page as they completed tasks and ignored the main nine links on the top right of each page.

1.3   Method Selection

A *card sort* is a categorization exercise, where participants group physical or virtual cards together based on their relationship to one another. By definition, an open card sort would allow our participants to name each of the categories they grouped together, thereby exposing where users would expect to find information, and what category names users would look for.

Various card sort methodologies have been implemented in the user experience field, including:

- Repeated card sorting
- Delphi and modified Delphi card sorting
- Focus group card sorting

Repeated card sorting is a "variation on open card sorting that involves the repeated sorting of a set of items to understand underlying dimensions or characteristics of a product or service" [1]. While this methodology intrigued us, we were not as interested in cataloging dimensions such as trustworthiness/untrustworthiness and attractiveness/unattractiveness as we were interested in how our users would group information into different categories.

Paul has done work on *Delphi* and *modified Delphi card sorting*, in which "participants build on the results from previous card sorts; instead of asking each participant to start from scratch, participants iteratively improve a proposed hierarchy, which may or may not be started by a subject matter expert" [6]. While Paul successfully reduced "the time to conduct the study and analyze the results" and also "lowered the costs of conducting a study" [6], we were not inhibited by these factors. We were also concerned that having subsequent participants work on modifying a card sort organized by others before them would color their own categorization and organizing principles. Finally, we did not think that running a card sort with a recommended participant pool of "8 to 10 people" would generate enough data to

support a major overhaul to the site navigation, because user experience research experts recommend using at least 15-20 participants for a card sort to generate enough data to substantiate findings and recommendations [5, 9, 10]. However, Tullis & Wood used "a modified version of WebSort to look at random sub-samples of different sizes from [a] full dataset of 168 participants," and concluded from their research "that it may not be cost effective to spend resources to gather information from more than 20-30 participants in a card-sorting study" [10].

Hawley ran *focus group card sorting*, inviting "8–15 participants who are representative of the target audience to a conference room, with a computer for each person… First, ask participants to complete the card sort individually. Then, lead a discussion with participants, regarding the organizational strategies they used during the card sort" [3].

Of all the new card sort methodologies we researched, focus group card sorting provided innovative ways for us to combine quantitative and qualitative data. As Hawley states, "…with an online study, designers may miss out on the insights and comments users can provide in person. Interpreting users' intentions from the data they submit online can be challenging. A focus group card sort is a method that leverages the best of both online and in-person techniques" [3].

This idea intrigued us. As far as we knew, no one had invited up to 15 participants to do an individual card sort *and* followed it up with a group card sorting activity. If Nielsen recommends 15 participants in a card sort study, and Tullis up to 30, we could satisfy the higher criteria for number of participants (25-30), and employ a new methodology, to get the quantitative data we needed as well as rich qualitative data that an individual online sort could not provide.

1.4   Tool Selection

Card sorts are typically conducted in one of two ways: by having users interact with physical index cards, which get sorted into stacks; or by using online card sort applications (for example, WebSort and OptimalSort), which allow users to group items into lists or clusters. We had used physical index cards when running card sorts in the past, but that format presented issues when analyzing the data. We had to input the data ourselves into unsupported programs such as USort, which we knew could not analyze more than 50 labels, and we had 79 labels.

We had recently used an online card sort application, WebSort, and found it saved us from the tedious task of manually entering data from participants, and also provided data that could be downloaded and analyzed in an Excel spreadsheet. However, we were concerned that some participants might have difficulty in using the WebSort interface, and that confusion would color the data. To mitigate that effect, we created a five-minute demo that enabled participants to do a mini card sort and thus learn how to use the WebSort interface.

2   Methodology

After our survey of tools and techniques, we employed a methodology that combined online card sorting, physical card sorting, and focus group card sorting. We came up with 79 labels for the open card sort, based on primary and secondary content that our customers frequently accessed; we also included labels for content that we knew would be added to the site in the near future. As recommended by Donna Spencer [8], we wanted to watch out for "overuse of a particular word." We rewrote many of the labels (for example, "Auto Insurance Coverages," "Car Insurance in Your State," "Senior Driving Information") so that participants would not automatically group together labels that used the same terminology.

We conducted the card sort at a neutral off-site facility in September 2010, so that none of the participants knew the company conducting the card sort. We recruited 30 participants; 26 showed up. We recruited a mix of customers and non-customers, following standard Liberty Mutual recruiting criteria. We conducted one session with 13 participants in the morning and a second session with another group of 13 in the afternoon. Each person was provided a laptop for the individual sort, and sat at a table with 3 or 4 other participants with whom they would complete the group sort. We staffed each table with its own facilitator and note taker. Each session lasted two hours.

2.1   Individual, Online, Unmoderated, Open Card Sort

During the first hour, our participants used their laptops to complete an online, unmoderated, individual card sort, using WebSort. We provided a 5-minute demo of an open card sort that our participants completed on WebSort so that they could first learn how to use the online application. They then proceeded to complete the online open card sort with our 79 labels using WebSort (see Figure 2). There were also open-ended follow-up questions that our participants completed online after they finished the open card sort. All participants completed their work within 45 minutes.
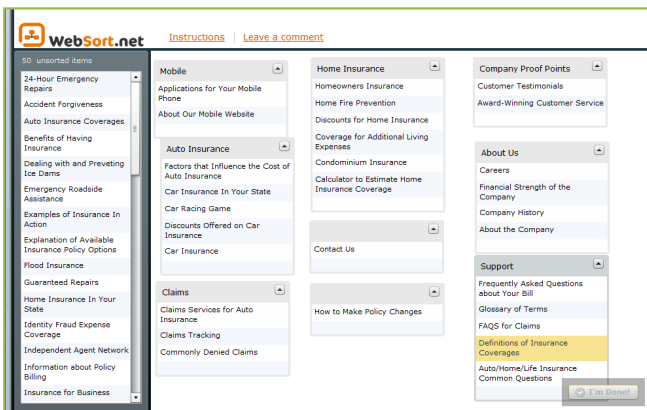

Fig. 2. Screenshot from WebSort. The dendrographs are M MODERATOR GUIDE] taking notes. front of them during teh eing captured.e results on WebSort and validated

2.2   Team-Based, Physical, Moderated, Open Card Sort

During the break, we removed the laptops and laid down brown paper to cover each of the tables. Then we brought our participants back in and asked them to sit at the same tables they were at before. Next, we evenly distributed, among participants in each group, a number of Post-It Notes that had printed on them the same 79 labels we used in the individual card sort. We asked participants at each table to work together and complete an open card sort as a team: that is, to organize the labels into groupings that made sense to them, and then to provide a name for each group of labels. As each team worked together, our facilitators moderated each team's activity: to encourage discussion, to ensure all members of the team were contributing to the conversation and activity, and to ensure no one person or persons were dominating the conversation and activity. Each table also had an assigned note taker to capture qualitative feedback and group discussion.

We asked each team to give a short presentation after they finished, in which a spokesperson for the group explained how and why they organized the information the way they did. We photographed the results of the Post-It Notes exercise for each group, and afterwards manually entered each group's results into WebSort.

2.3   Focus Group Discussion

We then followed up with questions to each group to facilitate a discussion. The moderator asked specific questions. Participants sat in their original seats, with the completed group card sort results still in front of them during the discussion. Supporting research team members (note takers and facilitators) were standing on the sides of the room, taking notes. These were the discussion topics for each group:

- How do you like to have your information organized (for example, by subject, process, business group, information type)?
- How many categories should there be? How many are too many? Too few?
- Were there any items that were difficult to categorize?
- What groups of items did you have difficulty naming?
- Were there any items you did not understand?

3   Analysis and Results

Our 26 participants created 210 unique category labels during their individual sorts. Our 6 groups of 4-5 participants used the same cards for the group sorts, and created 46 unique category labels. Some participants and groups used the same labels for categories (for example, "About Us," "Types of Insurance"), while others used different labels for similar concepts ("Resources," "Insurance Tips," "Helpful Tips & Suggestions," "Educational Tools").

After the 26 individual and 6 group card sorts were completed, we began analyzing the data. Three members of our team analyzed the individual card sorts, and three

other members analyzed the group card sorts. As suggested by Spencer [7], we first standardized the main categories and applied consistent naming conventions. We gave categories with similar names or concepts a consistent name and combined groups where participants used the same basic concept but a slightly different label.

Second, each member of each team separately combined similar standardized categories. Each member then presented and explained their categorization, and then each team agreed upon and arrived at our totals of standardized categories (see Tables 1 and 2).

Table 1. Final list of 19 online, individual standardized categories

Standardised category	Sorters who used this	Total cards in this category	Unique cards	Agreement
Miscellaneous	38	266	79	9%
FAQs	18	174	49	20%
Types of Insurance	29	265	44	21%
About the Company	26	289	47	24%
Advice/Tips	21	189	35	26%
Mobile	11	35	12	27%
Testimonials & Customer Feedback	7	37	19	28%
Video/Interactive	11	49	16	28%
Insurance 101	6	20	11	30%
Billing	10	51	16	32%
Benefits/Perks	8	85	33	32%
Your Policy	5	49	28	35%
Claims	18	128	20	36%
Home/Property Insurance	14	158	26	43%
Life Insurance	6	17	6	47%
Agent	10	19	4	48%
Savings	4	23	10	58%
Auto/Car Insurance	15	197	22	60%
Careers	3	3	1	100%

For example, in the individual sorts, participants put an average of 13 cards (197/15) into the "Auto/Car Insurance" category and used 22 different cards. This category has an agreement number of 0.6, meaning that 60% of participants put the same 13 cards in this category.

Table 2. Final list of 11 in-person, group standardized categories

Standardised Category	Sorters Who Used This	Total Cards in This Category	Unique Cards	Agreement
Customer Service	9	72	29	28%
Tools & Resources	8	66	30	28%
Coverages & Benefits	4	34	22	39%
Insurance Products	6	80	30	44%
Interactive Media	4	30	16	47%
About Us	8	65	16	51%
Home Insurance	3	26	14	62%
Claims Center	6	45	11	68%
Find a Local Agent	4	11	4	69%
Auto Insurance	3	40	18	74%
Insurance Basics	1	5	5	100%

For example, in the group sorts, participant groups put an average of 13 cards (40/3) into the "Auto Insurance" category and used 18 different cards. This category has an agreement number of 0.74, meaning that 74% of participant groups put the same 13 cards in this category.

We saw differences between individual and group card sorts as we compiled the standardized categories. While individuals tended to create a greater number of minor categories, the groups tended to unite the minor categories into much broader categories. This difference might have been a product of analyzing fewer data points from the group vs. individual card sorts, or learned behavior and group think as individuals moved into the group card sort exercise. However, during the group card sorts and follow-up discussions, we discovered that the individual card sort exercise informed the group card sort exercise.

The group card sort exercise provided a forum for participants to discuss those items and categories that were confusing or unclassifiable to them, and enabled participants to relegate these previously unclassifiable items to defined categories, instead of bucketing them to their own minor categories such as those identified in Table 1 ("Advice/Tips," "Benefits/Perks," "FAQs," "Mobile," "Savings," "Testimonials & Customer Feedback," and "Video/Interactive"). This was a beneficial insight from the group card sort. Groups stated to us that their discussion and give-and-take enabled them to reach agreement on major categories to put these cards into.

Third, we compared the grouping of labels between the online, individual studies and the in-person group studies to detect any patterns. WebSort uses participant data to perform cluster analyses and generate "tree diagrams" or dendrograms [2], which display not only the relationship between items but the strength of that relationship, or how frequently the items are associated. The dendrograms showed differences between the individual and group results. The dendrogram for the group results showed closer relationships between content grouped together, but that could be an artifact of having a fewer number of groups (6) than number of individuals (26). The individual results showed weaker relationships between content grouped together.

Both individuals and groups wanted to separate information from tasks. For example, the Auto Insurance (or Auto/Car Insurance) category was focused on tasks, such as our label for "Tool to Estimate Car Insurance Coverage." Information on benefits and features – such as our labels for "24-hour Emergency Repairs," "Accident Forgiveness" and "New Car Replacement" – were isolated into a separate category, "Coverages & Benefits" (or "Benefits/Perks"). This revealed a key finding that both the individuals and the groups organized the content by task (for example, "Tool to Estimate Car Insurance Coverage") *and* by information (for example, "Accident Forgiveness").

Next, we created site maps to organize the results from the individual and group card sort exercises. What this visualized was that individuals had sorted items into a narrow and deep hierarchy. The site map created from the individual card sorts consisted of 3 major categories – Our Company; Policy, Billing, and Advice; Insurance; and 2 utility categories – Careers and FAQs. However, what the group card

Table 3. Group card sort categories and our final navigational categories

Group Card Sort, Major Categories	Our Final Navigational Categories
About Us	About Us (moved to shared navigation)
Auto Insurance	Auto Insurance
Home Insurance	Home Insurance
Insurance Basics	Insurance Resources
Insurance Products	
Claims Center	Claims Center
	Customer Service
	Life Insurance
Group Card Sort, Utility Categories	**Our Utility Bar Options**
Life Insurance	
Online Account	Customer Login
Find a Local Agent	Find a Local Agent

sort revealed was that our groups had sorted items into a medium navigational hierarchy, with 6 major categories, or double what our individual card sorts showed us – About Us, Auto Insurance, Home Insurance, Insurance Basics, Insurance Products, and Claims Center – and 3 utility categories: Online Account, Find a Local Agent, and Life Insurance, as shown in Table 3.

As a result, we created a 6-option navigational bar (see Figure 3). We moved About Us into a top-level navigation bar that all Liberty Mutual websites share. We moved Life Insurance to the same level as Auto Insurance and Home Insurance, based on the objectives of both our participants and our stakeholders, who wanted all insurance products at the same categorical level. We also created a utility bar above the 6-option navigational bar, which included Online Account (Customer Login) and Find a Local Agent.
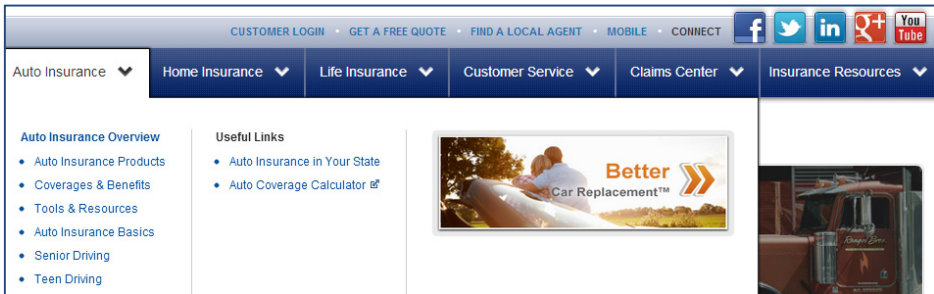

Fig. 3. Navigation from current www.libertymutual.com website, showing the utility bar and drop-down options for one category, Auto Insurance

4   Lessons Learned

One of the drawbacks of our study, and an opportunity for future studies, would be to have a better mechanism for collecting qualitative feedback from the smaller group card sorts and the larger group discussion. While our methodology included having a facilitator and a note taker for each group, the documentation of the raw, qualitative data was limited. The following factors could have contributed to it:

- The close proximity of the three groups per session and the groups having to talk over one another made it difficult for the note taker to listen
- The close proximity of the tables made it difficult for the note taker to navigate around the table to hear all participants in the group

Some solutions that could have improved the collection of the quantitative data include the following:

- A larger research room (or separate breakout rooms for the group sorts) would allow note takers and facilitators to walk around and avoid cross-talk from other groups
- Audio and/or video-recording equipment to record the conversations for each group would allow the research team to perform a content analysis of the raw data after the sessions

5   Conclusion

In conclusion, the combination of individual card sorts, group card sorts, and focus group discussions allowed us to obtain qualitative and quantitative data in one day with a large sample size but without a major investment in research budget or time. While the final recommendation for the new LibertyMutual.com site navigation largely followed the content organization from the group exercises and the focus group discussion, the individual card sort activity established a baseline of knowledge for each participant, aiding in productive negotiations and discussions for the group exercises. Also, by having the individual results, we were able to validate the differences between the results of the post-group exercise discussions with the quantitative similarities and differences of the individual and group card sort results.

Two years later, the site navigation has had a positive effect on the business. Key performance indicators and website analytics demonstrate that customers are actively using the new navigation to complete their tasks, rather than ignoring it as they did in the earlier version. In multiple usability testing studies since the navigation was updated in 2011, participants have commented positively on the website navigation, rating it as being one of the best attributes of the site. Third-party research panels have also favored the new site navigation: in its 2012 Insurance Website Evaluation Study[SM] Best Practices guide, J.D. Power and Associates commented that "Liberty Mutual provide[s] menus that include upper-funnel and lower-funnel shopping tools," allowing shoppers to easily find relevant shopping tools on any page [4].

References

1. Curran, M.J., Rugg, G., Corr, S.: Attitudes to Expert Systems: a Card Sort Study. The Foot 15(4), 190–197 (2005)
2. Dong, J., Martin, S., Waldo, P.: A User Input and Analysis Tool for Information Architecture. In: CHI 2001 Extended Abstracts on Human Factors in Computing Systems, pp. 23–24 (2001)
3. Hawley, M.: Extending Card-Sorting Techniques to Inform the Design of Web Site Hierarchies. In: UXmatters (2008),
   ```
   http://www.uxmatters.com/mt/archives/2008/10/
   extending-card-sorting-techniques-to-inform-the-design-of-
   web-site-hierarchies.php
   ```
4. Power, J.D., Associates, 2012 Insurance Website Evaluation Study[SM] Best Practices. The McGraw-Hill Companies, Inc. (2012),
   ```
   http://www.jdpower.com/content/press-release/xRjzSrq/
   2012-insurance-website-evaluation-study.htm
   ```
5. Nielsen, J.: Card Sorting: How Many Users to Test (2004),
   ```
   http://www.nngroup.com/articles/
   card-sorting-how-many-users-to-test
   ```
6. Paul, C.L.: A Modified Delphi Approach to a New Card Sorting Methodology. Journal of Usability Studies 4(1), 7–30 (2008)
7. Spencer, D.: Card Sort Analysis Spreadsheet (2007),
   ```
   http://rosenfeldmedia.com/books/cardsorting/
   blog/card_sort_analysis_spreadsheet
   ```
8. Spencer, D.: Card Sorting: Designing Usable Categories. Rosenfeld Media, Brooklyn (2009)
9. Spencer, D., Warfel, T.: Card Sorting: A Definitive Guide (2004),
   ```
   http://boxesandarrows.com/card-sorting-a-definitive-guide
   ```
10. Tullis, T., Wood, L.: How Many Users Are Enough for a Card-Sorting Study? In: Proceedings, Usability Professionals Association (UPA) 2004 Conference, Minneapolis, MN (2004)