Artifact Usefulness and Usage in Sensemaking Handoffs

Authors

Nikhil Sharma
School of Information, University of Michigan
1075 Beal Ave, Ann Arbor, MI-49286
Email: nsharma@umich.edu

George Furnas
School of Information, University of Michigan
304, West Hall, Ann Arbor, MI-48103
Email: furnas@umich.edu

The complexities of sensemaking suggest that collaboration should be difficult, requiring a rich ecology of collaboration support. This can be a problem for handoff sensemaking, where one person must continue where another has left off, sometimes with only material artifacts as the basis of the handoff. A detailed analysis of essential attributes of sensemaking tasks, and elements identified in the computer supported collaborative work literature were combined to yield insight into handoff sensemaking and guide empirical work. A lab-study showed that handoffs relying only on artifacts from previous sensemaking could be successful. The lab studies also indicated timing and quality affects on the sensemaking handoffs, with different quality materials used differently, and early efforts possibly being particularly difficult to hand off. Design of support for sensemaking handoffs will have to take such effects into account.

Introduction

Sensemaking is often done collaboratively. While this collaboration can take place in many modes, the research focus here is on handoffs. Handoffs are a form of serial-asynchronous collaboration where provider's unfinished worked is picked up and completed by a recipient. The goal of our research has been to understand the unique nature of sensemaking and the

various factors that affect sensemaking handoffs. The ultimate intent is to use this understanding to provide implications for design of sensemaking handoff support systems.

We have employed a sequence of approaches, which we sketch in this paper. First, an examination of sensemaking theories and scenarios helped to identify several critical attributes of sensemaking. These were then further examined in the light of general research on collaboration to suggest how the unique nature of sensemaking would impact its collaborative handoff. This framing then led to several empirical studies beginning with an exploratory field study of sensemaking handoffs in computer support helpdesks. This was followed by lab studies focusing on specific aspects of sensemaking handoffs including some temporal characteristics and aspects of handoff-material. The two lab studies are detailed in this paper which focused on the handoff artifact usage during sensemaking handoffs.

Essential Attributes of Sensemaking Tasks

The first step in understanding sensemaking handoffs was to understand the unique nature of sensemaking in general and how this would impact sensemaking handoffs. We were interested in studying activities that exhibit serious amounts of difficult sensemaking and hence might require support. Consideration of theories and examples provided us attributes for identifying such activities. We used the sensemaking theories of Russell et al (1993), Weick (1996) and Dervin (1998) and examined prototypical sensemaking scenarios like the diagnosis of 1999 West Nile Virus (WNV) outbreak in Queens, NY, as well as tasks used in the sensemaking literature. The result was a detailed set of attributes, falling into two groups, that has been useful for conceptual analysis, and as a practical guide in the choice and modification of tasks for sensemaking studies. We present this set briefly below.

Activities high on all attributes should be prototypical examples of tasks involving substantial sensemaking; those low on all attributes should not. Note that sensemaking can take place in lightweight forms in even activities with low scores on some of the attributes.

Knowledge Structure Creation Attributes

Most definitions of sensemaking (for example Russell et al (1993), Weick (1996) and Dervin (1998)) acknowledge the creation of some knowledge structure. This essential characteristic of sensemaking tasks can be elaborated using three attributes. Two of these attributes are aspects of the knowledge structure that is created and the third is a capability of the resulting structure. Since knowledge

structures include a schema for organizing information as well as information that is encoded into the schema (Russell, 1993), it can be argued that structure creation can have two aspects that can cause the need for sensemaking: either a novel representation may be required or information encoding may be difficult. The presence of one or both of these two aspects can be used to argue the case for existence of non-trivial sensemaking. What we call "broader applicability" is a capability supported by the created knowledge structure and its presence is a strong indicator of sensemaking.

1. **Representation novelty requirement.** The first attribute in this group reflects the idea that sensemaking activity differs from routine activity in that new structures of knowledge ("novel representations") must be created to provide the understanding needed for tasks at hand. The notion of novelty is not simple since the knowledge structure for sensemaking is never created from scratch; it is often at least partially appropriated from elsewhere (Qu, 2005). If the sensemaker has access to good pre-existing representations, the need for novel representation, and consequently the need for sensemaking, is reduced. Representation ideas can also come from the sensemaker's own existing knowledge as well as from representations created by others working on a similar task (Qu, 2005). Thus when evaluating a task for the degree of novelty required, it is recommended that a researcher articulate: (1) some approximation to the knowledge or understanding that must ultimately be achieved, (2) the sensemakers' existing knowledge and their access to representations created by others, and (3) some assessment of the amount of new work needed to move from (2) to (1). The larger that amount of work entailed, the stronger the case that the task involves substantial amounts of sensemaking.
2. **Encoding Difficulty.** The second structure creation attribute associated with substantial sensemaking involves the extent of non-trivial encoding required. Encoding was the term used by Russell et al. for the process of putting information specific to the task instance at hand into the representation or framework the sensemaker is trying to use for the task. Encoding information into a good representation can be difficult, and sensemaking can consequently be more substantial, for at least three reasons. First, the overall relevance of various information at hand to the current task may not be known (e.g., a detective trying to make sense of a case wondering, "Is Joe involved in this at all? Is Al?") Second, the precise

relationship of known-be-relevant information to the representation in use may not be known (detective: "I am sure Joe is involved in the plot, but I do not yet know how.") Finally, encoding may be difficult when the roles of items cannot be evaluated independently, and instead many pieces of information need to be compared simultaneously for a match against many parts of the representation. ("Joe, Al, and Mike are both involved, but who is calling the shots, and who is going to actually do the deed?"). Insofar as a researcher can articulate how any of these difficulties arise, they have a stronger case that degree of sensemaking involved is high.

3. **Broader Applicability.** The third structure creation attribute concerns an emergent capability of the final structure. Sensemaking differs from problem solving per se, in that it creates understanding, not just a solution. For example, one might simply stumble upon a solution by luck or brute force and in that sense "solve" a problem, while never having really understood, i.e., made sense of, it. We take as one core criterion for sensemaking, that understanding is achieved. Genuine understanding is inherently generative and captures regularities in the situation in a way that supports many inferences, including those yielding a solution. Importantly, however, it supports other inferences as well. True understanding can be used to solve a whole suite of related problems, not just the original one. Therefore one candidate operationalization of understanding (i.e., sensemaking accomplished) is a notion of "broader applicability." Representations or knowledge structures may not always be created with the intent for multiple uses and broader applicability, but if a researcher can make the case that some activity builds a capability that allows success in broader re-use, in multiple scenarios, they have substantive evidence that such generative structure was created and that the activity entailed considerable sensemaking.

Process Complexity Attributes

Sensemaking is inherently difficult and even stressful, due in part to the complexity involved in making sense of a new situation. Two possible reasons for complexity in sensemaking are: (1) the difficulty in searching the space of possible representations and (2) the interdependent nature of sub-tasks involved in the sensemaking process. We have elaborated these two aspects as attributes of complexity. These aspects were drawn in part from Funke's (1991) work on

complex problem solving.

1. **Representation Search Space.** The first process complexity attribute whose presence indicates more substantial sensemaking concerns the nature of the "space" of possible representations through which the sensemaker must "search" to find one suitable for the task. The space of representations used in a prototypical sensemaking scenario can be difficult to search, and hence the sensemaking more substantial, for at least three possible sets of reasons. First, are factors contributing to combinatorial complexity in the design space, arising from a high number of representation elements, difficulty in identifying possible elements, and any interdependence of these elements. Second, are problems in evaluating candidates in the search space: difficulties of observation, manipulation, or assessing heuristic search value of items. Third are dynamic complications, where a continually evolving situation is forever changing the problem to be solved, making the space of relevant representations itself dynamic. The stronger the case made that these three aspects are present in the task, the more difficult the needed representation space will be to search, and hence the better the case that the task will involve non-trivial amounts of sensemaking.
2. **Subtask Interdependence**. Russell et al (1993) characterized sensemaking as an "interlocking set of different types of subtasks." Sensemaking tasks are complex because these sub-tasks are simultaneously active, in interleaving threads that are closely coupled. To argue that a task involves serious sensemaking, a researcher should be able to identify simultaneous threads of activities that must rely on information from each other to guide them, and that considerable coordination and communication between the activities is needed for their successful execution.

These sensemaking attributes may not be exhaustive, but by trying them against a variety of cases, including several from the literature, they have proven useful in distinguishing tasks and situations that involve higher amounts of sensemaking from those that do not.

Crucial Elements for Sensemaking Handoffs

The two preceding groups of essential attributes of sensemaking tasks can be combined with insights from CSCW research to help focus on the special challenges of collaborative

sensemaking, and handoff of sensemaking in particular. For example, the CSCW literature suggests that due to the complex and closely coupled nature of sensemaking activity, it will succeed in a collaborative mode most easily when the certain elements are present. Studies by Olson & Olson (2001) and Klein, Moon & Hoffman (2006) suggest that these elements include a strong intent to collaborate and high common ground (Clark & Brennan, 1991). Further research yields a whole ecology of other elements key to complex collaborative tasks: good awareness information of collaborators (Dourish & Belotti, 1992), a shared physical space (Suchman, 1996), additional communication channels (Patterson & Woods, 2001) and handoff artifacts.

These collaboration elements can be considered explicitly in terms of how they address the identified attributes of tasks requiring serious sensemaking. For example, common ground and communication channels support sensemaking's closely coupled activities whose definition and coordination may require frequent callbacks to the handoff provider. Common ground can also help reduce the difficulty of the search-space of representations by increasing the acceptance of provided representations, information and search heuristics. Awareness and shared-space can inform the recipient when clarifications can be requested and can also support the tracking of dynamic situations, which can reduce the interdependence attribute of a task. The handoff artifact can enable the transfer of representations, information and search heuristics. Additional communication can help by allowing clarifications of representation and information handed-off, and by allowing the tracking of dynamic situations.

Laboratory studies of Sensemaking: Artifact Use

This examination of the unique nature of sensemaking and the crucial elements required for its handoff positioned us more strongly to explore sensemaking handoff empirically. This section briefly reports the findings of two laboratory studies of sensemaking handoff conducted so far.

Insights from an earlier field study (Sharma, 2008) were used to guide two laboratory studies to allow better control of the tasks and the nature of the hand-off collaboration. Handoffs in the field study (Sharma, 2008) were accompanied by broad collaborative support. We wanted to examine how much was possible without much collaborative-support, an important kind of boundary case. Handing of artifacts of sensemaking in progress also form the core of a practical technology vision we call "wide area sensemaking" where such artifacts would be made available via the web to others anywhere working on similar sensemaking problems.

Therefore, the lab studies investigated whether sensemaking material can be helpful on its own, with other major collaboration elements kept to a minimum. In the experimental conditions there was little common ground, no shared physical space, no available awareness information, and no option for additional communication. The lab-studies were actually trying to ascertain the usefulness of handoff material on its own when other collaboration elements are lacking.

The first study examines if sensemaking handoff material can be useful even when common ground, awareness information and additional communication are absent. The second study explores how recipients use high and low quality sensemaking handoff materials differently. This study focuses on the use of handoff support material, noting differences in usage depending on the quality of the material.

Study Task

The studies reported here tested the performance of students sharing sensemaking information in an online searching and sensemaking task. Choosing amongst a complex set of products has been considered sensemaking by Russell et al (1993) so choosing a camcorder with the help of information on the web was the task used here. Its appropriateness was also supported with an analysis using the attributes described earlier. The task was found to have both significant structure creation and complexity, which implies that sensemaking needs to be accomplished. The unfamiliarity with the product in question is the biggest driver of sensemaking here. The following task was presented to the participants of both studies:

Your friend's father is an avid traveler who goes on vacations frequently. That's why your friend thinks a digital camcorder be the perfect gift for him. He is also a serious photography enthusiast and he would make movies not just for memories but also to create travel movies that provide a medium for his artistic expression. The product's typical use will be on vacations, but it will sometimes be used for making home videos. Your friend needs help in buying the gift. Use the provided resources to search for the most appropriate camcorder for him and recommend a place to buy it at the best price. Your friend is willing to spend up to $500 for the camcorder but will go slightly over budget for a good camcorder.

The subjects were also told that they will need to fill in a post-experiment questionnaire which will include questions about the justification of their choice.

Experiment Details

As mentioned before, the participants were presented a scenario where they had to search online and recommend a camcorder to be used by a friend's father. This indirect task, recommending for a friend's father, was used to in part to standardize the task scenario, but also, importantly, to encourage the participants to externalize both their work and the rationale for their final choice. Time allowed for the task was one hour, after which participants had to make their final decision.

At the end, the participants individually answered a questionnaire about their search process and their acquired knowledge of camcorders. The questionnaire had three main sections. The first section dealt with demographic and background information. The second section was composed of questions related to camcorders. The purpose of these questions was to gauge the increase in the participants' understanding of camcorders and the subsequent broader applicability of their understanding. For every question they indicated whether they knew the answer before the online-search. The last section was related to participants' self-evaluation of their effort, the process of collaboration and feedback regarding their partners or sensemaking material that they received.

In order to enable their friend to understand their choice, the participants were told to document their search. The information they collected about camcorders was to be saved in a way that would be usable by their friend later. They were told to bookmark all important pages they visited. They were also asked to organize their bookmarks into appropriate categories or folders. They were provided paper/pen and a word processor (MSWord) so that they could make additional notes if they wanted, during the task to supplement the bookmarks.

Equipment

Subjects used two identical Dell D800 1.6GHz notebook-computers running Windows XP, with attached mice. The computers had 15.4-inch diagonal displays (1900x1200 pixels) and 11Mb wireless Internet connections. Subjects had access to a word processor (MSWord), an Internet browser (Internet Explorer) and scratch paper to make notes

Experiment 1 on Handoff Effectiveness

The first experiment used a between-subject manipulation to evaluate if sensemaking handoff material can be useful even when common ground, awareness information and additional communication are absent.

Participants

A total of 30 participants were recruited though email sent to students at a large mid-western university. Sixty eight percent of the participants had technical educational backgrounds (engineering, cognitive science, information, economics and management) and thirty-two percent had non-technical educational backgrounds (education, arts, planning, languages and humanities). Only those who had neither bought a camcorder nor searched online for one before were asked to participate. Of the 30 total participants, 16 were male and 14 were female. The average age was 26 years with a range from 19 to 39. All but two participants had completed their bachelor's degrees and all participants had shopped online at one time or another.

Experiment Conditions and Groups

There were three experimental conditions, all of which involved the same camcorder task. Thirty participants were randomly assigned to the three conditions:

1. **Control group.** In this condition, participants (N=10) completed the camcorder recommendation task alone.
2. **Hand-Off Collaboration**. In this condition, the participants (N=10) were provided a set of bookmarks, in the form of an "exported webpage," and accompanying notes made by a randomly chosen previous participant from the control group. They were informed that they could use the provided bookmarks and notes to aid themselves in the task if they wanted to, but they still had to create their own, separate collection of notes and bookmarks.
3. **Synchronous Collaboration.** In this condition, two people completed the task side by side in the same room on separate computers. Thus the participants (N=10) worked in 5 pairs. During the task they could collaborate by exchanging notes and links verbally or via instant messaging (IM/chat). They were informed that they could help each other in any way they wished, but

had to create their own, separate collections of notes and bookmarks. They were also told that they were not required to agree on their final choices.

It was expected that synchronous collaboration would perform the best since the collaboration mode allowed double the 'work hours'. It was expected that the handoff condition would perform better than the control group but not better than the synchronous group because handoff participants had no access to additional communication with the providers. It was expected that the handoff group would do better than the control group because they had access to representations created by earlier sensemakers to guide their sensemaking.

Results

The basic dependent measure used here was the quality of the final recommendation chosen by the participants. Two independent experts made a list of camcorder criteria reflecting the profile of the hypothetical user and budget. The experts generated 22 and 29 important features respectively, out of which 20 were common. Every camcorder could either score low (1 point), medium (2 points) or high (3 points) on each of these features. Experts also gave the features an importance weight from 1 to 10. The correlation of the weights between the experts was 0.6. Of the 29 total features generated, 2 were not found in any of the camcorders selected by participants. Since they would be irrelevant to scoring, they were dropped. All 27 remaining features were used, and given either the average weight if mentioned by both experts or the corresponding individual weight if mentioned by only one. These weighted components were added up to create an overall Choice Quality score (CQscore). The 18 different camcorders chosen by the 30 participants in the study ranged in CQscore from 188 to 255.

The subjects' final camcorder choices were analyzed to see if collaboration had an impact on quality, as indicated by the CQscore. Mean CQscores were calculated for all three groups and t-tests were performed to determine if differences in means were significant. The data are displayed in Table 2 below.

Table 1. Mean Choice Quality Scores in the Three Conditions

Group/ Condition	Mean (Std.Err)
Group I (Control)	214.9 (5.82)
Group II (Hand-Off)	235.1* (5.65)
Group III (Synchronous)	232.8* (4.55)

* differs from Group I (Control) at the p<.02 level of confidence

The mean CQscores in collaboration groups II (Handoff) and III (synchronous) were significantly higher than the control group, with $p<0.011$ and $p<0.013$ (one tailed t), respectively. There was no statistically significant difference in CQscores between the handoff and synchronous groups ($p>0.75$).

The other basic dependent measure was the participant's score on the post-session knowledge questionnaire, interpreted with subjects' self-report: The participants were given a point for each correct answer, provided they indicated that they did not know the answer from prior knowledge. Overall, the learning scores ranged from 0 to 17 with a mean of 5.3 (S.D. = 4.3). Mean learning scores were calculated for all three groups and t-tests were performed to determine if differences in means were significant. The data are displayed in Table 3 above. Although in the post-experiment questionnaire all groups asserted learning various facts from the exercise, there was no differential effect of condition: the mean scores in groups I, II and III were not significantly different at the 0.05 level. The subjects' behavior and attitudes showed that the predecessor's bookmarks actually helped the recipients. In the Hand-Off collaboration condition, 80% of the participants indeed used the stranger's bookmarks, visiting 32% of them on average (SE=10%). The subjects generally rated the bookmarks as quite understandable (average 4.25 on a scale of 5, SE=0.31), and those who considered the bookmarks more useful, visited a higher percent of them ($r=0.93$, $p<0.002$) and expressed a lower need for more time ($r=-0.9$, $p<0.014$). Although the Handoff Collaboration helped, the specifics of the performance of the two collaborators (the subject creating the bookmarks and the subject receiving the hand-off) were not strongly linked: The final CQscores between the two were not significantly correlated ($r=0.33$, $p>0.35$), and none of the subjects made the same choice of camcorder as their predecessor. In Synchronous collaboration condition, there was a stronger linking of performance: there was a significant correlation

(r=0.66, p<0.038) between the CQscores of the participant and their partners.

Table 2. Mean Learning Scores in three conditions

Group/Conditions	Mean (Std. Err)
Group I (Control)	5.9 (1.63)
Group II (Asynchronous)	4.3 (1.02)
Group III (Synchronous)	5.7 (1.45)

(No pair-wise differences statistically significant.)

Discussion

The results showed that performance was better in an information gathering and sensemaking task when collaboration was involved. The handoffs here had many of the crucial collaboration elements lacking. Since the providers did not know the recipients and were told a subsequent person might use their bookmark, the intent to collaborate could not have been high. Many of the participants were graduate students in the university and thus may have some common ground. However they were again strangers and with varying backgrounds that suggests that common ground was low. There was no option of additional communication and spatial or other awareness in the laboratory setup of the handoff condition. The only element present was the handoff material.

Both synchronous collaboration and use of handoff material prepared by others resulted in better performance in the online search and sensemaking task. One possible reason why the handoff material was useful for the recipients was because the material was the outcome of a nearly complete sensemaking effort by the provider. Most recipients also rated the provided bookmarks as good quality. Thus the result of the recipients' work was due to two person-hours, something that can also explain why synchronous collaboration was better. This finding is a encouraging for handoffs since it suggests that a handoff from a nearly completed sensemaking work can be nearly as helpful as having another collaborator.

Another interesting observation was that participants in the hand-off condition sometimes did not seem to start with the provided bookmarks, rather they started on their own and came back to the handed bookmarks after a few searches. There could be several reasons. One of the more intriguing is that perhaps they were not "ready" to use them. Perhaps they needed to explore a bit themselves before they could know how to interpret the provided material or assess its value. This possible reluctance to start with and completely depend on the handoff material might also have contributed to the fact that participant pairs in the handoff condition had different recommendations. However since the above finding regarding the pattern of usage of handoff materials was just observed by the experimenter and because the experiment setup did not involve a close observation of material usage pattern, another study was conducted to do so which is explained in the next section.

Experiment 2 on Handoff Material Quality and Use

While the first lab study shows that handoff material can be useful, it raises the questions: does the quality affect what the material is useful for? And does the quality affect when and how a material is used? The issue of the effect of quality on the use of handoff material by recipients was further explored in the lab in the second study described here. In many ways, the second study was an open-ended exploration to see what was going on in detail with handoffs in sensemaking and the focus was not on collecting and comparing performances of many subjects as in the first study. To see how the influential Russell et al. (1993) model of sensemaking fit in the second participants' use of the bookmarks, detailed minute-by-minute observational data on users' behavior was collected. The second study also tried to find how the pattern of handoff material use might differ as a function of the quality of the material.

As mentioned before, one of the goals of this research is to investigate the use of information systems to share sensemaking work. The sensemaking work shared by people might also vary in quality and it is useful to investigate how differently the work with varying quality will be used by subsequent sensemakers.

Participants

A total of eight participants were recruited though email sent to students at a large mid-western university. Five participants had technical (Information, human computer interaction) and 3 had non-technical educational backgrounds

(languages, political science). Of the 8 total participants, there were equal numbers of males and females. The average age was 27 years with a range from 22 to 51. All but 3 participants had completed their bachelor's degrees and all participants had shopped online at one time or another.

The low number of participants cannot give statistical power to the study but since it allowed for detailed and minute by minute activity of the sensemakers post handoff and, the study can offer rich insights into handoff material usage and can guide further large-scale studies.

Experimental Conditions

The participants performed the same camcorder recommendation task used in the first study (handoff condition) with two exceptions. First, while some of the subjects in the previous study had gotten notes as well as bookmarks (depending on whether their randomly chosen "provider" generated notes); in this study we used material from people in the first study who had in fact not generated notes – only bookmarks. That is, only bookmarks were handed-off to the current subjects. Second, to simplify the task and focus on the role of the bookmarks in their sensemaking, the current subjects were not required to make bookmarks or notes of their own; they just had to come up with a recommendation for a camcorder for the profile (friend's father).

Each minute while performing the task, the subject's behavior was assigned two codes by the experimenter sitting with the subject, one for their activity (G=looking at general information, M=looking at specific models, S=selecting a model), and one for the type of website they were looking at (Handed-off Bookmark, Buying Guide website, Seller site, Review site).
Participants were randomly assigned to two conditions, each performing the same camcorder task, but differing in the quality of the bookmarks provided.

1. **High Quality Bookmarks.** (n= 4) A single set of bookmarks was chosen from those generated in the previous study that had been given very high ratings by earlier users. (Understandable=5/5, Useful=4/4, Better than own=4/4. An independent domain expert also gave this set of bookmarks a 5/5 rating for overall helpfulness). On inspection, these bookmarks appeared systematically organized (two levels), both the links and their groupings

were well labeled, there were several general links, and the groupings appeared in a coherent order.

2. **Low Quality Bookmarks.** (n= 4) A set of bookmarks was selected that had a comparable number of links to the High Quality set, but had low ratings from subjects in the first study, was not organized into groups, and was not carefully labeled. (The independent judge gave these a 3 rating on overall helpfulness.)

After completing the task the participants were asked a series of questions about their usage of the provided bookmarks.

Results

The time stamps of various webpage visits were normalized to a (0, 1) range, by dividing by the subject's overall time. These normalized timestamps tell what proportion of the way through each subject's session the sample of activity occurred. The mean timestamps of these three categories of activity are in Table 6. Subjects overall spent about 10% of their time using the provided bookmarks. This did not vary significantly between groups or between individual subjects. The overall use of the bookmarks was sporadic, spread throughout much of the session, though on average they tended to be consulted a bit earlier than other websites: the mean normalized timestamp for consulting bookmarked sites was 0.404 compared to 0.522 for other sites (significantly different, $p<0.027$).

Table 3. Bookmark Quality and Use

	High Quality Material	Low Quality Material	Significance.
Mean Normalized Time-stamp	0.277	0.567	*p = 0.0055*
Time to First Look	1.25 min	2.5 min	*n.s.*
Time from First Look to First Use	0 min	8 min	*Trend: p = 0.07*
Time After Last Use	34 min	10 min	*Trend: p = 0.07*

There was no significant difference in the Choice Quality score as a function of the quality of the bookmarks handed-off, but there was a significant difference in the way the High and Low Quality bookmarks were used. Links in the high quality

bookmarks were followed early and those in the low quality ones followed late.

This can be seen in several ways. For example, the mean normalized time-stamp for bookmark use was considerably earlier for High Quality bookmarks (Table 4, row 1). Also, two trends in the data were relevant in an intriguing way. As one would hope, there was no significant difference (Table 4, row 2) between groups as to when they took their first look at the provided bookmarks. Only after that point could the bookmark quality make a difference. However, people who were handed-off high-quality bookmarks tended to use them use them right away (Table 4, row 3). Moreover, people with the good bookmarks used them and were done with it, whereas people with bad bookmarks still consulted them increasingly, up to the end. (Table 4, row 4). This is also illustrated in Figure 1 below, showing bookmark use in successive thirds of the subjects' session time. Use of High Quality bookmarks was all in the first two thirds, while Low Quality bookmarks were used more and more towards the end.

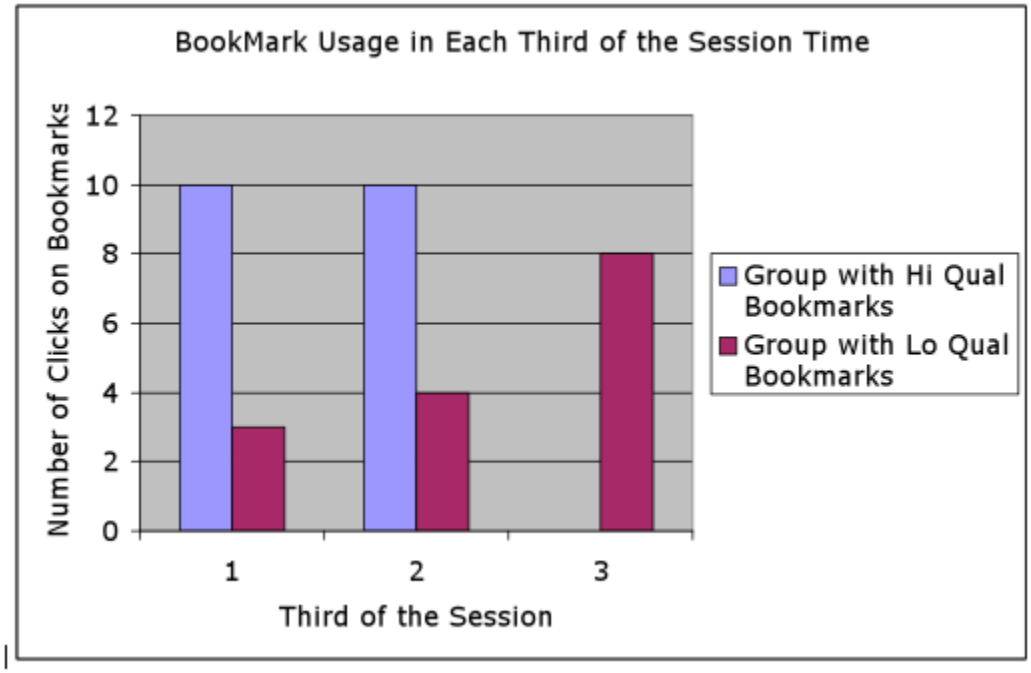

Figure 1. Bookmark usage in each third of the session time for the high and low quality bookmark groups. (c2(2)=13.91, p<.001)

Discussion

Examination of the data led to an interesting pattern of bookmark use which can

be summarized thus: everyone looked early, then those with High Quality bookmarks used and finished with them, but those with Low Quality bookmarks, while ultimately using them just as much, waited until nearer the end of their time.

According to the Russell et al model (1993), sensemaking involves two major subtasks: (1) coming up with a good representation or framework for the information to be used in a task, and (2) encoding instances of that representation based on particular data in the world. If a person is working entirely alone, they must produce their own representation, by deducing from their own background knowledge and inducing from instances in the context of the task. If the person has access to the results of professionals who have made sense of things and authored guides, the person can get some help coming up with a good representation from them. Presumably this is why in the data analysis it was found that the sites consulted early tended to be general overview sites as well. If there is output from some relevant amateur efforts that can be handed off, these too can provide guidance for finding a good representation if they are of sufficient representational quality. The High Quality bookmarks in this study were presumably valuable in this way: they clustered the bookmarks sensibly for the task, named both the bookmarks and the clusters well and even presented the clusters in an order that made sense for the task (General, Models and Sellers). Furthermore, having the links clustered accordingly, those links could be used systematically in gathering information to be encoded. For these reasons, the good bookmarks were immediately useful, were followed quickly, and used fully. Later work just carried on where these left off.

In contrast, the low quality bookmarks had no such valuable structure. Subjects in that condition were left much more on their own to come up with a representation by a combination of induction from instances and any hints the subjects might find on the web for general issues to consider. Only after they had done so were the Low Quality bookmarks helpful. Even then, after the subjects had a representation of their own, without clustering and good labeling in the provided bookmarks, it was not clear which links would be useful for providing information relevant to what topics in the subjects' personally created representation. Thus it was likely that the links were of most use in providing a check on their own work, regarding the adequacy of their representation, and completeness of coverage for encoding. In fact, two of the subjects finished in exactly this way, spending their

final 5-10 minutes taking their most extensive look at the provided bookmarks just before making their decision. Apparently the quality of even the poorer bookmark set was adequate for this.

The small numbers in each group make between group comparisons at most suggestive in this study. However this study points to the usefulness of detailed analysis of handoff material usage. The larger-scale study is planned to substantiate the results of this study.

Conclusion and Summary

The handoff of sensemaking is difficult and challenging. Essential attributes of sensemaking highlight its high complexity and interdependence. The collaboration literature suggest that such activities should require rich support, such as common-ground, awareness, shared-space and additional communication. These elements, however, can be lacking in many handoff situations, where the material artifacts become the only basis of handoff. Nonetheless, we have found that sensemaking is successfully handed off in the field, and in the lab it can sometimes be as effective as simultaneous collaboration. In both circumstances there was evidence that the quality of the handed-off material and the timing of the handoff were important. Poor quality material seemed to be used at different times and in different ways from good quality material. Handoff in moderately early phases was avoided in the field and may turn out to be disproportionately unhelpful in the lab. All of these findings suggest that design for better handoff sensemaking may indeed be possible, but will require subtle support of these timing and quality aspects.

Acknowledgements

This work was funded by NSF grant IIS-0325347-ITR

References

Clark, H.H., & Brennan, S.E. (1991). "Grounding in communication." In L. Resnick, J. Levine, & S. Teasley (Eds.) Perspectives on socially shared cognition Washington, D.C.: American Psychological Association, pp. 127-149.

Dervin, B. Sense-Making Theory and Practice: An overview of user interests in knowledge seeking and use. (1998) Journal of Knowledge Management, 2(2), 36-46.

Dourish, V. Belotti, "Awareness and Coordination in Shared Workspaces", in Proceedings of CSCW `92, ACM Press, Toronto, Canada, pp.107-114, 1992.

Funke, J. (1991). Solving complex problems: Human identification and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), Complex problem solving: Principles and mechanisms (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum Associates.

Furnas, G. W. “Representational Change in Sensemaking,”. Position Paper for Workshop on Sensemaking. CHI 2008, Florence, Italy.

Klein, G. Moon, B. and Hoffman, R.R. “Making Sense of Sensemaking 1: Alternative Perspectives,” Intelligent Systems, vol. 21, no. 4, 2006, pp. 70–73.

Olson, G. M. and Olson, J. S. (2001) Distance Matters. Human-Computer Interaction 15, 139-179.

Patterson, E.S. and Woods, D.D. Shift Changes, Updates, and the On-Call Architecture in Space Shuttle Mission Control, Computer Supported Cooperative Work, v.10 n.3-4, p.317-346, December 2001

Qu Yan, Furnas George W. Sources of structure in sensemaking. CHI Extended Abstracts 2005.

Russell, Daniel M., Stefik, Mark J., Pirolli, Peter, Card, Stuart K. (1993) "Cost structure of sensemaking" Proceedings of the CHI '93. 269-276.

Sharma, N. Sensemaking Handoff: When & How. Proceedings of the 71st ASIS&T Annual Meeting, vol 45, pp 73-84. 2008.

Suchman L., Constituting shared workspaces. In Y. Engestrom & D. Middleton (eds.) Cognition and Communication at Work. Cambridge University Press, Cambridge, MA, 1996.

Weick, K. E. (1996). Sensemaking in organizations. Newbury Park, CA: Sage