

Generality of Transfer of Navigation Knowledge from Degree-of-Interest Trees

Peter Pirolli and Michelle Gumbrecht

Palo Alto Research Center, Inc.

3333 Coyote Hill Road

Palo Alto, CA 94304 USA

pirolli@parc.com

mgumbrec@parc.com

ABSTRACT

An experiment was conducted to assess the generality of navigation knowledge learned from interaction with the Degree-of-Interest (DOI) Tree. The degree of transfer of learning from the DOI Tree was compared to transfer of learning from a standard Internet Explorer browser. Twenty participants worked with either the DOI or Explorer browser in a Training Phase in which they performed information-finding or comparison tasks requiring navigation through a widely used Web directory. Degree of transfer was measured by effects on task completion time in a Transfer Phase in situations with varying degrees of similarity to the original training conditions. An analysis based on percent transfer scores indicated that participants trained on navigation of an information structure using the DOI showed superior transfer across a range of conditions as compared to participants trained using the Explorer.

Author Keywords

Focus +context, Degree-of-interest Tree, navigation

ACM Classification Keywords

H5.m. Information interfaces and presentation

INTRODUCTION

Human-information interaction (HII) systems for navigating large hypermedia systems such as the World Wide Web may differ in the rate and accuracy with which they provide users with useful information for a given task [10]. They may also differ in how they support learning about navigation paths through information. That is, HII systems may support different degrees of learning of navigation knowledge as a function of experience. In this

paper, we argue that the Degree-of-Interest (DOI) Tree browser is expected to support greater learning of navigation knowledge than standard Web browsers. We then test this hypothesis using a transfer-of-learning paradigm.

DOI TREE BROWSER

The DOI Tree [7, 8] is a visualization of graphical tree structures (Figures 1 and 2). Many information structures are typically arranged into hierarchical trees, such as file systems, organizational charts, and taxonomies. More complex information structures, such as Web sites, may be amenable to a tree-based visualization, as is typically done when depicting Web site maps. The DOI Tree visualization is an example of a *focus+context* technique [5, 6], in which the user's degree of interest is estimated over elements of the tree. Greater display resources, such as display space, are given to parts of the tree with higher estimated degree of interest (the *focus*) and less resources are given to parts of the tree with lower estimated degree of interest (the *context*). The DOI Tree gives more space to tree nodes predicted to be relevant to the user's focus of attention but also displays the remainder of the tree, in compressed spatial form, as context.

The DOI Tree uses a variety of methods to implement a focus+context visualization of hierarchical information [7]. Specifically, it uses *logical filtering* to select nodes to display, based on a DOI function; *geometric distortion* to render nodes at a size proportional to their DOI values; *semantic zooming* to determine the amount of node labeling to present; and *aggregate representations* to elide portions of the tree in the context region. The version of the DOI Tree tested in this study used a DOI function that was based solely on the focal node most recently clicked on by the user, and distance in the tree structure. Specifically, the most recently clicked node, and all nodes on the path from that node to the root, are assigned maximal DOI values. The DOI values for remaining nodes decrease linearly as a function of the degree (number of links) distance from the highest interest nodes.

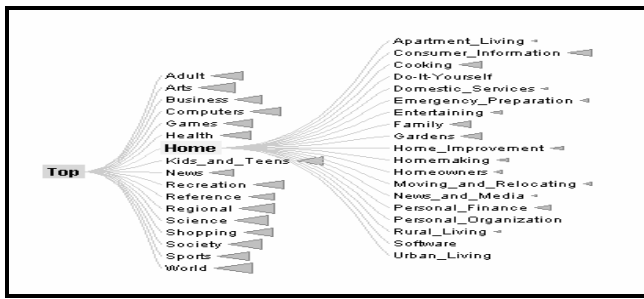


Figure 1. DOI Tree visualization after user has clicked on the node “Home”.

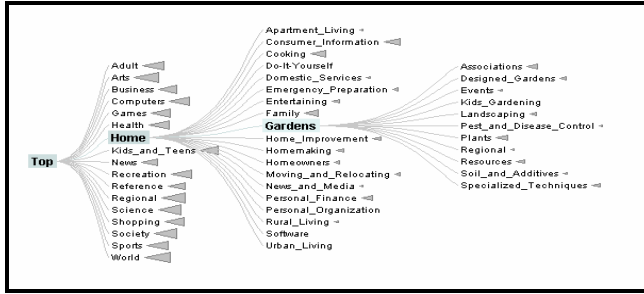


Figure 2. DOI Tree visualization after user has clicked on the node “Gardens”.

Figures 1 and 2 are screen shots of a DOI Tree visualization of the Open Directory Project (<http://dmoz.org>) directory of Web sites organized (primarily) as a hierarchical categorical structure. These examples show the effect of clicking on nodes in the DOI Tree. In Figure 1, the user has just clicked on the node “Home” and in Figure 2 the user has just clicked on “Gardens”. The change in the DOI Tree from Figure 1 to Figure 2 occurs as a smooth animated transition.

PERFORMANCE AND LEARNING WITH THE DOI TREE

Although hundreds of articles have been written about focus+context techniques, there are relatively few empirical evaluations of their effectiveness on navigation performance or learning. In general, those empirical evaluations have shown mixed results. This is consistent with the sobering meta-analysis by Chen and Yu [3] of empirical evaluations (through the year 2000) of information visualizations that suggested small non-significant effects of information visualization on performance time and accuracy. Previous studies [11] of the Hyperbolic Tree¹ browser, which is another focus+context technique, produced mixed results: In one experiment using information-finding and comparison tasks there was no overall superiority of the Hyperbolic Tree over a benchmark standard 2D tree browser. However, in a second experiment focusing on a more constrained set of just information-finding tasks, the Hyperbolic showed a 38%

¹ The Hyperbolic Tree browser was a precursor to the Star Tree product of Inxight.

superiority in performance times. Furthermore, Pirolli et al. [11] noted that the observed performance times tended to be more strongly associated with task differences and individual differences than with differences between browser interfaces. Although some researchers have found similar weak or even negative [15] benefits of focus+context, some have indeed been able to obtain superior performance time and accuracy for focus+context navigation.

Perhaps focus+context techniques help users learn more about tree structures than standard techniques, since more of the tree structure is displayed at any given time. One might hypothesize that users can attend to more of the tree over any given interval of time using a focus+context technique than other standard methods. Indeed, eye-tracking data reported by Pirolli et al. [11] suggested that Hyperbolic Tree users were visually scanning more of the tree structure at a faster rate than the 2D tree benchmark browser. In addition, the Pirolli et al. [11] study suggested that on a restricted set of tasks (those with low *information scent* [12]), Hyperbolic users were showing greater learning benefits. Low information scent tasks were ones in which the node labels on the displayed trees provided such poor cues for navigation choices that users tended to explore many unproductive paths in the tree structure on their way to finding target information. However, with practice, Hyperbolic users tended to show more rapid improvements in their navigation performance when compared to users of the 2D benchmark browser. This result would be consistent with the general notion that focus+context techniques such as the DOI Tree might provide more rapid learning of an information structure than more traditional information visualization techniques.

TRANSFER OF LEARNING

One way to approach the analysis of learning is to identify what knowledge is gained from earlier experiences and how that knowledge *transfers* to future situations. In the psychology of learning, transfer is usually defined as the extension of something learned in one context to new contexts. Variations in the quality and quantity of learning in the original training context can alter the range of contexts improved by that learning experience. For instance, a person who learns the solution to an algebra problem by rote may show an inability to solve other problems whereas another person may learn something deeper (e.g., the meaning of the “associative rule”) and be able to solve a whole class of algebra problems. In the current study, our interest lies in whether the distinct visualizations and interactions supported by different browsers lead to different types of learning that support varying degrees of transfer.

Perhaps the most substantive and comprehensive theory of transfer is the *modified identical elements theory of transfer* [1, 14], which uses the ACT production system theory of cognition to provide detailed predictions about the transfer

of knowledge. This theory has been used, for instance, in human-computer interaction to predict how experience with one text editor will lead to savings in learning another text editor [13, 14]. In essence, the modified identical elements theory predicts that the degree of transfer between an earlier learning context and a later transfer context depends on the knowledge elements acquired in the learning context that apply to the transfer context. The quality of knowledge learned, and the degree of transfer attained, can vary with such factors as individual differences [e.g., 4] and instruction [e.g., 2].

When users interact with an information browser (e.g., a Web browser), such as the DOI Tree or the Internet Explorer, we may question what navigation-relevant knowledge (including skill) they are acquiring, and how that might impact future navigation. One might ask whether the acquired navigation knowledge is specific (e.g., to the specific browser, the specific navigation paths, specific tasks, etc.) or more general. In the study presented here, we employ an experimental design based upon a traditional paradigm for the study of knowledge transfer [9, 14]. We expect this study to provide indications of the degree to which the DOI Tree supports the acquisition and transfer of information-seeking knowledge and the generality of that knowledge.

In this study, we use an experiment designed to test the following levels of transfer, ordered according to our expectation about what constitutes increasing generality of learning:

1. *Same-task, same-browser*, in which users perform navigation tasks through a Web directory and then perform the same tasks with the same browser later,
2. *Same-task, different-browser*, in which users perform navigation tasks through the Web directory in one browser and then do the same tasks using a different browser,
3. *Different-task, same-browser*, in which users perform one set of navigation tasks through a Web directory and then transfer that learning to novel navigation tasks in the same browser,
4. *Different-task, different-browser* in which users learn to navigate a Web directory in one browser on one set of tasks and then transfer to novel tasks on a different browser.

AN EXPERIMENTAL PARADIGM

We employ an experimental paradigm aimed at measuring transfer that derives from an experimental design advocated by Katona [9] and, more recently, Singley and Anderson [pp. 37-41, 14]. The basic logic of the design is illustrated in Tables 1 and 2.

Table 1 shows that users assigned to a Transfer Condition receive some Training tasks A followed by some other kind

		Tasks Presented	
		Training	Test
Condition	Transfer	A	B
	Practice	B	B

Table 1. Basic experimental design for measuring transfer.

		Model for Performance Scores	
		Training	Test
Condition	Transfer	Y_A	$Y_{B_A} = Y_B - T_{AB}$
	Practice	Y_B	$Y_{B_B} = Y_B - T_{BB}$

Table 2. Model for the analysis of performance scores.

of Test tasks B. The expectation is that there is some degree of improvement on B that results from having previously learned A. Table 1 also shows a Practice Condition in which users simply practice B tasks in both Training and Test blocks. The expectation is that repeated practice on a specific set of task situations will, generally, provide an upper bound on performance improvement due to transfer.

As a hypothetical example, in an HCI transfer study one might present users with Tasks A with a spreadsheet and Tasks B with a word processor. Participants in the Transfer Condition would first learn to do tasks on the spreadsheet (A), then would be tested on tasks involving the word processor (B). Participants in the Practice Condition would just repeat tasks using the word processor (B).

Table 2 presents a model that can be used to analyze the transfer effects obtained in the experimental design in Table 1. Assume that the measurements consist of a task completion time variable, Y , which would be expected to decrease with learning.

In Table 2, Y_A is the performance score achieved on Training tasks (A) in the Transfer Condition and Y_B is the performance score achieved on the Training tasks (B) in the Practice Condition. On the Test tasks in the Transfer Condition, Y_{B_A} is the score on B tasks following prior learning on A tasks. On the Test tasks in the Practice Condition, Y_{B_B} is the score on B tasks following prior learning on B tasks.

Table 2 shows that the scores Y_{B_A} and Y_{B_B} in the Test phases of the experiment are expected to include some

improvement, T_{AB} , due to knowledge transfer from A tasks to B tasks, or an improvement T_{BB} due to practice of B prior to testing on B (since the scores are assumed to be performance time these transfer effects are assumed to decrease scores).

The transfer effects can be calculated by computing differences between performance scores observed in various conditions, i.e.,

$$T_{AB} = Y_B - Y_{B_A},$$

and

$$T_{BB} = Y_B - Y_{B_B}.$$

These can be used to calculate what is known as a *percent transfer measure*, $T_{\%}$,

$$\begin{aligned} T_{\%} &= \frac{Y_B - Y_{B_A}}{Y_B - Y_{B_B}} \times 100 \\ &= \frac{T_{AB}}{T_{BB}} \times 100 \end{aligned}$$

This percent transfer measure assumes that the improvement obtained in the Practice Condition (captured in the denominator) is an upper bound. Typically, percent transfer scores will vary from 0% to 100%, although it is possible to obtain negative scores (indicating negative transfer) or scores greater than 100% (indicating supertransfer).

Tables 1 and 2 represent a basic paradigm for the measurement of transfer. The experiment presented below uses a more elaborate design, but it is based on the underlying logic of this basic paradigm.

METHOD

Design

In our study, we were interested in the transfer of learning achieved using the DOI Tree, and as a comparison browser we used the Internet Explorer. Participants in our study, using either browser, were constrained to navigate the Google Directory Web structure. For exposition purposes, *DOI* will refer to the DOI Tree and *Google* will refer to the standard Google Directory rendered via Internet Explorer.

Table 3 presents the experimental design. The design was aimed at measuring transfer of navigation knowledge learned with the DOI Tree to situations of varying degrees of similarity to the original learning conditions. The most specific level of transfer would be to identical navigation tasks carried out with the same interface.

Intermediate levels of transfer would be to different navigation tasks carried out in the same interface, or to the same tasks carried out in a different browser. The most general level of transfer would be to novel navigation tasks

Condition	Training	Transfer	
		Same tasks	Different tasks
DG Transfer	DOI (Tasks A)	Google (Tasks A)	Google (Tasks B)
GG Practice	Google (Tasks A)	Google (Tasks A)	Google (Tasks B)
GD Transfer	Google (Tasks A)	DOI (Tasks A)	DOI (Tasks B)
DD Practice	DOI (Tasks A)	DOI (Tasks A)	DOI (Tasks B)

Table 3. Experimental design. Training tasks were performed twice within the given interface.

carried out in a different browser. Users could be assigned to four different groups (Table 3). In the DG Transfer group, participants would work with the DOI Tree during a training phase in which they worked twice with navigation tasks A. Then they would participate in two kinds of transfer tasks: one in which they worked with the Internet Explorer (Google) on the same navigation tasks A, then on a different set of tasks B. In the GG Practice group, participants would work with Google throughout the experiment, but the transfer phases would test performance on either the same navigation tasks A or different tasks B. The GD Transfer and DD Practice groups mirror the first two groups.

Model of Transfer Effects

The design was aimed at measuring two sorts of transfer effects that may really be considered two ends of a continuum of general to specific transfer:

Information architecture transfer, which is meant to refer to a general sort of transfer that comes from browsing a Web site with a browser and coming to have some general knowledge of the information structure that transfers broadly to other navigation tasks and even other browsing interfaces. We can specify four kinds of information architecture transfer involved in this experiment:

- $T_{I:DG}$, which indicates transfer of navigation knowledge about the information architecture acquired from training on specific tasks using DOI to test situations involving novel tasks using Google on the same information architecture,
- $T_{I:GG}$, which indicates transfer of navigation knowledge from training on specific tasks involving Google to test situations involving novel tasks using Google,

- $T_{I:GD}$, which indicates transfer of navigation knowledge from training on specific tasks using Google to situations involving novel tasks using DOI,
- $T_{I:DD}$, which indicates transfer from training on tasks involving DOI to test situations involving novel tasks on DOI.

Task transfer, which refers to a more specific kind of transfer that is tied to specific navigation tasks. We can specify four kinds of task-specific transfer in this experiment:

- $T_{T:DG}$, which indicates transfer from tasks performed on DOI to the same tasks using Google,
- $T_{T:GG}$, which indicates task-specific transfer from training on Google to the same tasks performed later with Google,
- $T_{T:GD}$, which indicates transfer from tasks performed using Google to the same tasks using DOI,
- $T_{T:DD}$, which indicates task-specific transfer from training on DOI to the same tasks performed later with DOI.

Table 4 presents the model for the analysis of performance scores in this design. The performance time in the Training phase of the experiment is measured as Y_D for DOI and Y_G for Google. Performance scores in the Transfer phases are expected to include the transfer effects described above. Table 4 indicates how those transfer effects are expected to combine to yield the observed performance scores in the Transfer conditions. For instance, in the DG Transfer condition, the Transfer phase involving the same tasks as used in Training will exhibit task-specific transfer, $T_{T:DG}$, plus general information architecture transfer, $T_{I:DG}$. However, the DG Transfer phase involving different tasks will only exhibit the general information architecture transfer effect $T_{I:DG}$.

Condition	Training	Transfer	
		Same tasks	Different tasks
DG	Y_D	$Y_G - (T_{T:DG} + T_{I:DG})$	$Y_G - T_{I:DG}$
GG	Y_G	$Y_G - (T_{T:GG} + T_{I:GG})$	$Y_G - T_{I:GG}$
GD	Y_G	$Y_D - (T_{T:GD} + T_{I:GD})$	$Y_D - T_{I:GD}$
DD	Y_D	$Y_D - (T_{T:DD} + T_{I:DD})$	$Y_D - T_{I:DD}$

Table 4. Model for analysis of performance scores for the experimental design defined in Table 3.

Participants

Twenty participants (12 men and 8 women, average age = 22.5) were recruited by emailing three groups: summer interns within the company; psychology graduate students at a local university; and undergraduate and graduate students subscribed to a paid experiments list at that university. Company interns volunteered their time to participate. Outside participants' compensation ranged from \$15-\$40 for the session, depending on the success of eye tracking calibration.

Stimuli

We created the tasks by navigating through all of the major categories and their children nodes within the DOI tree and the Google Directory. Within the "same" and "different" task categories, participants were presented with two types of questions: "Retrieval" or "Comparison". We developed the Retrieval questions based on the webpage descriptions presented in the "Web Pages" section of the Google Directory. We determined the solutions by selecting the URL with the most unique webpage description, such that it was distinguishable from the set of returned results. We developed the Comparison questions by selecting a topic with many associated categories, of which one category had a discernable highest or lowest number of related links. The solutions were based on the parenthetical numbers located next to each category in the "Categories" section of the Google Directory. We verified that the solution paths for each task within each interface were identical. The word length of task questions in Phases 1-3 ranged from 21-25 words ($M = 23$), and Phase 4 tasks ranged from 19-32 words ($M = 24.2$). Task files were saved in .txt format in WordPad.

Procedure

Upon arrival to the lab, participants sat in front of a Dell Optiplex GX270 desktop computer equipped with two NEC MultiSync LCD 2080UX+ 20-inch monitors located side-by-side. After we obtained consent, participants wore an SMI Eyelink head-mounted eye-tracker, and we began the adjustment and calibration processes. We used the EyeLink Intelligent Eye Tracking System to handle eye-tracking functions.

After the calibration process was completed successfully, we instructed participants on the experiment procedure. We showed them how to use Weblogger, a program created in Visual Basic, which functioned as the "control center" for the experiment. The Weblogger box contained eye-tracker controls for connecting to the eye-tracking system, calibrating, and drift correcting; a digital timer with start and stop buttons; and the current task question. Weblogger also recorded all eye-tracking and web browser activity. Weblogger was placed on the left-hand monitor such that it would not interfere with the interface display area.

Participants were given 24 tasks, six of which were repeated randomly in three phases ("same" tasks A); the six tasks in the fourth phase had not been encountered

previously (“different” tasks B). The final six tasks also required participants to navigate through categories that they had not viewed before. Participants answered two types of questions within these tasks: “Retrieval” and “Comparison”. Retrieval questions required participants to find a specific URL described in the task. Comparison questions required them to find a category link that corresponded to the most or fewest items associated with the given task. We walked them through two practice tasks, one of each question type, while demonstrating how to navigate the DOI tree and/or Google Directory, depending on their assigned condition.

Regardless of condition, the interface was presented always on the right-hand monitor. Participants in a Practice condition (DD Practice or GG Practice) used the same interface to perform all 24 tasks. Participants in a Transfer condition (DG Transfer or GD Transfer) used each interface for 12 tasks. Weblogger was programmed to open the second interface automatically starting at Task 13. Participants were instructed to be certain that they read and understood the task before clicking on the Weblogger start button, such that only navigation time was measured. There was a five-minute time limit per task. When their allotted time expired, a message notified them to advance to the next task. Once they clicked on the start button, the assigned interface opened to its “home state”, which listed the 16 major parent categories, and then they began to navigate. The interface reset itself to the home state when they began the next task. Since participants in a DOI condition could not access directly the URL information needed for the task solutions, they were instructed to press the shift key on the keyboard and left-click on the node simultaneously when they believed it was the correct node pertaining to the task solution. This action launched Internet Explorer and brought them directly to the corresponding node within the Google Directory. Participants were allowed to use the Ctrl-F (“Find in Page”) command in Google to search through webpage descriptions, but they were prohibited from typing into the search query box.

Participants gave their answers by clicking on the URL or category that they believed correctly answered the question. Once the webpage loaded completely, they clicked on the stop button. They were not given any feedback on the accuracy of their answers. Since the tasks were repeated in Phases 1-3, we instructed them to clear their search history in Google after every sixth task. We performed a forced drift correction after every task to compensate for any eye tracking errors that occurred. An experimenter sat in the room with the participant for the duration of the experiment in case of any technical difficulties with the program.

At the conclusion of the experiment, participants completed an Internet-based survey on usability of and satisfaction with the interfaces used during the study.

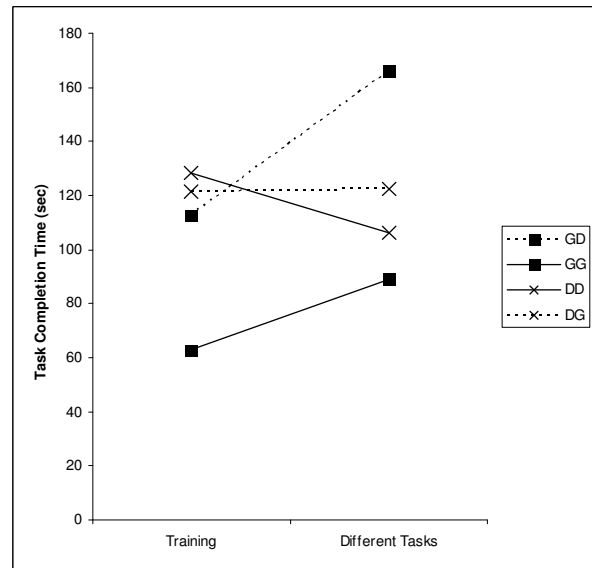


Figure 3. Mean observed task completion times for training and transfer to different tasks.

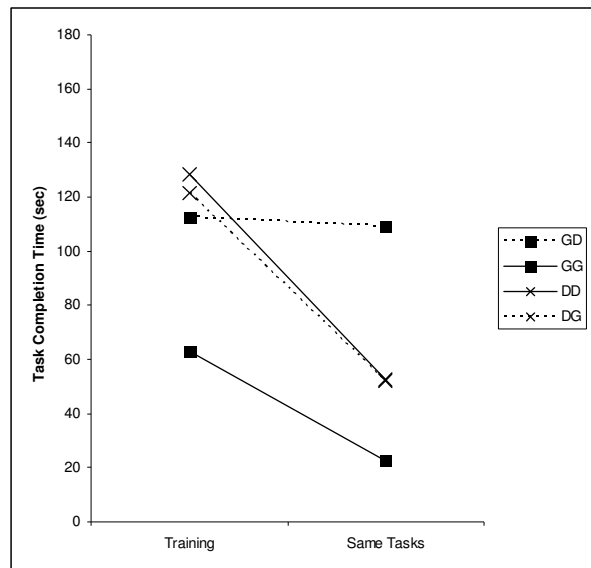


Figure 4. Mean observed task completion times for training and transfer to the same tasks.

RESULTS

The mean task completion times for each of the training and transfer conditions are presented in Figures 3 and 4. Figure 3 shows the mean task completion times for conditions in which participants were trained on one set of tasks using DOI or Google, and then performed novel tasks either with the same browser as in training (GG and DD conditions) or a different browser (GD or DG conditions). Each line corresponds to a different group of participants, so it is worth noting that in the training phase one should not expect differences between the GG and GD groups, as the tasks and browsers are the same. The apparent training

Effect Estimated		Estimate (sec)
Browser	Google (\hat{Y}_G)	125.87
	DOI (\hat{Y}_D)	169.81
Information Architecture Transfer	GG ($\hat{T}_{I:GG}$)	-2.40*
	GD ($\hat{T}_{I:GD}$)	23.21
	DG ($\hat{T}_{I:DG}$)	30.37
	DD ($\hat{T}_{I:DD}$)	-36.84
Task Transfer	GG ($\hat{T}_{T:GG}$)	-94.94**
	GD ($\hat{T}_{T:GD}$)	-43.04**
	DG ($\hat{T}_{T:DG}$)	-70.67**
	DD ($\hat{T}_{T:DD}$)	-56.94**

Table 5. Effect estimates.

(Notes: * indicates $p < .05$; ** indicates $p < .001$.)

phase differences in Figure 3 between the GG and GD groups are likely due to individual differences. Similarly, no difference is expected between the DD and DG groups in the training phase. Overall, Figure 3 suggests moderate negative or no transfer effects when participants train on tasks with one browser and then transfer to novel tasks on a different browser (DG and DG conditions). When training and transfer are done with same browser, there appears to be an improvement for the DOI, but not Google. Below, we discuss analyses that tease apart the effects due to differences among participants from the transfer effects. However, on the surface, Figure 3 suggests mixed results for general Information Architecture transfer.

Figure 4 presents the mean task completion times for conditions in which participants were trained on one set of tasks using DOI or Google and then performed the same tasks on the same (GG and DD) or different browser (GD or DG). Again note the same spurious differences in the training condition times as discussed around Figure 3. All of the transfer effects in Figure 4 appear to be moderately to

strongly positive. On the surface, Figure 4 suggests positive task-specific transfer in all conditions.

Analysis of Transfer Main Effects

Linear regressions were conducted using an effects coding scheme based on the model outlined in Table 4. The effect estimates yielded by this analysis are presented in Table 5. An analysis of variance was performed including Participants as a random variable. Table 5 shows a longer estimated task time for the DOI Tree as compared to Google, however this effect did not reach significance $F(1, 333) = 2.20$, $MSE = 12269.89$. This failure to find a general superiority for the DOI on navigation tasks is consistent with the general difficulty in finding such performance improvements with focus+context browsers over standard browsers.

The estimates in Table 5 for the Information Architecture Transfer show both positive and negative transfer effects when study participants were trained on one set of tasks and then encountered novel tasks in the transfer conditions. When the browser was the same in both training and transfer conditions (GG and DD conditions) the information architecture transfer effects suggest an improvement (decrease) in task time. When the browsers were different in training and transfer conditions (GD and DG conditions), the information architecture transfer effects suggest negative transfer (an increase in task time). The analysis of variance yielded a significant Information Architecture Transfer effect, $F(3, 333) = 2.75$, $p < .05$. Specific post-hoc planned comparisons indicated that the only statistically reliable Information Architecture transfer was in the GG condition: $F(1, 333) = 4.34$, $p < .05$.

Estimated effects of task-specific transfer in Table 5 suggest positive transfer for all four conditions. The analysis of variance showed an overall effect of task specific transfer $F(3, 333) = 76.27$, $p < .001$. Planned post-hoc tests indicated that each of the task-specific effects were significant: in the GG condition, $F(1, 333) = 60.96$, $p < .001$, in the GD condition, $F(1, 333) = 34.52$, $p < .001$, in the DG condition, $F(1, 333) = 47.68$, $p < .001$, and in the DD condition, $F(1, 333) = 40.87$, $p < .001$. None of these effects were pairwise different from one another. So, doing the same tasks in the training condition and transfer condition yielded improvements in task time, and these improvements occurred even when the same tasks were performed using different browsers (the DG and GD conditions). However, transfer was not statistically superior for DOI over Google (or vice versa).

Percent Transfer

We calculated percent transfer scores for the various conditions plotted in Figures 3 and 4 and these are presented in Table 6. To calculate these percent transfer scores we used the improvement in the same-task, same-browser conditions (GG and DD) as upper bounds on performance improvement. Consequently the same-task

percent transfer score in the same browser (GG and DD) conditions was defined as 100%. The same-task GG condition was used as the upper bound for the same-task DG, and different-task GG and DG conditions. The same-task DD condition was used as the upper bound for the same task GD, and different-task DD and GD conditions.

For conditions in which the same training tasks were presented in the transfer phase, the percent transfer was greater when going from DOI to Google than vice versa: In the DG condition, going from DOI to Google, there was 41.4% transfer, and in the GD condition, going from Google to DOI, there was 21.1% transfer. In general, the percent transfer on the same tasks as presented in training indicate superior transfer was achieved by users trained on the DOI Tree.

For conditions in which different tasks were presented in transfer, training on the DOI showed 39.3% transfer to the novel tasks performed with the DOI in the transfer phase (DD condition) and this was greater positive transfer than the similar condition in which users trained on Google and transferred to novel tasks on Google (GG condition; 2.5% transfer score). When training on one browser and one set of tasks and transferring to novel tasks on a different browser, DOI training showed less negative transfer than Google: DG yielded -31.2% transfer and GD yielded -71.6% transfer. In general, the percent transfer scores suggested more positive transfer or less negative transfer was achieved by participants trained on the DOI and then transferring to novel tasks.

Usability ratings

In the post-experiment survey, participants gave a series of ratings on comfort, comprehension of information, and satisfaction with the interfaces that they used during the study. Ratings were based on a Likert scale (1 to 5), with 5 representing the highest rating for any category.

Overall, participants were more comfortable with using Google ($M = 3.47$) than with using the DOI ($M = 2.80$), $t(14) = 2.320, p < .05$. For this analysis, we conducted a paired-samples t-test comparing all users who used Google for at least one phase of the experiment versus all users who used the DOI for at least one phase of the experiment. Given the familiarity of the Internet Explorer and Google to our participants, this result is not surprising. GG practice users ($M = 4.00$) rated their comfort with searching for information in Google significantly higher than DD practice users ($M = 3.00$) using the DOI, $t(4) = 3.162, p < .05$. GG practice users ($M = 4.20$) also were more satisfied in their ability to find and understand the content in the Google Directory than DD practice users ($M = 3.20$) using the DOI, $t(4) = 3.162, p < .05$.

Participants who used Google exclusively (GG practice, $M = 4.20$) were more satisfied in finding and understanding the content within the Google Directory than participants who used Google first (GD transfer, $M = 2.60$), $t(4) = -4.00,$

Transfer		Percent Transfer ($T\%$)
Same Tasks	GG	100*
	GD	21.1
	DG	41.4
	DD	100*
Different Tasks	GG	2.5
	GD	-71.6
	DG	-31.2
	DD	39.3

Table 6. Percent transfer scores for conditions presented in Figures 3 and 4 based on estimates in Table 5.

Note: * indicates upper bound on improvement (see text).

$p < .05$. Although this suggests some decrement in satisfaction with Google after exposure to the DOI Tree, there were no significant differences in ratings when we compared users who used Google second (DG transfer) versus GG practice users.

GENERAL DISCUSSION

The percent transfer scores (Table 6) and estimates of transfer effects (Table 5) suggest that the magnitude of transfer effects is ordered (from greatest improvements to least) into four different levels:

1. *Same-task, same-browser*, in which users perform navigation tasks through the information architecture and then perform the same tasks with the same browser later,
2. *Same-task, different-browser*, in which users perform navigation tasks through the information architecture in one browser and then do the same tasks using a different browser,

3. *Different-task, same-browser*, in which users perform one set of navigation tasks and then transfer that learning to novel navigation tasks in the same browser,
4. *Different-task, different-browser* (which yields negative transfer) in which learning to navigate in one browser on one set of tasks and then transfer to novel tasks on a different browser.

The percent transfer measures indicate that the DOI showed superior transfer to a standard Web (Explorer) interface across Levels 2-3 (Level 1 is used as the upper bound to calculate the percent transfer scores).

Overall we failed to find a significant performance difference between use of the DOI Tree versus a standard Web (Explorer) interface, and the general trend was that DOI navigation was slower. This could be a consequence of the unfamiliarity of the DOI Tree browser to the sample of users tested in this study. However, the lack of a robust (or even moderate) performance boost for the DOI browser is consistent with the general difficulty in finding boosts in navigation performance through use of focus+context visualizations [e.g., 11].

We had hypothesized that because more of the information structure is presented to DOI users that there would be more opportunity to learn about the information architecture, in comparison to users of a standard Web browser. The percent transfer scores tend to corroborate this claim, although a different analysis based on a model of the underlying transfer effects (Table 5) did not indicate statistically reliable superiority of the DOI Tree in producing transfer of navigation knowledge. In other words, the superior transfer effects produced by the DOI are less than robust and reliable.

More generally, one of the aims of this research has been to attempt to understand what people learn from different kinds of information browsers. Although we have presented an overall analysis of transfer effects, we believe that it is detailed models of learning, such as the kind developed previously in the ACT theory [1] that will provide a deep understanding of learning to navigate information structures.

ACKNOWLEDGMENTS

This work has been supported by the Advanced Research and Development Activity, Novel Intelligence from Massive Data Program Contract No. MDA904-03-C-0404 to S.K. Card and Peter Pirolli. We thank Christiaan Royer for comments on an earlier draft of this paper.

REFERENCES

1. Anderson, J.R. and Singley, M.K. The identical elements theory of transfer. in Anderson, J.R. ed. *Rules of*

- the mind*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1993, 183-204.
2. Bielaczyc, K., Pirolli, P. and Brown, A.L. Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, 13. 221-252.
3. Chen, C. and Yu, Y. Empirical studies of information visualization: a meta-analysis. *Int. J. Hum.-Comput. Stud.*, 53 (5). 851-866.
4. Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P. and Glaser, R. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13 (145-182).
5. Furnas, G. Generalized fisheye views. in *CHI '86, Proceedings of the ACM conference on human factors in software*, ACM, New York, 1986, 16-23.
6. Furnas, G.W. Small views of large structures, Unpublished manuscript, Bell Laboratories, Murray Hill, NJ, 1983.
7. Heer, J. and Card, S.K., DOI Trees revisited: scalable, space-constrained visualization of hierarchical data. in *Advanced Visual Interfaces*, (Gallipoli, Italy, 2004), ACM Press, 421-424.
8. Heer, J. and Card, S.K. Efficient user interest estimation in fisheye views *CHI '03 extended abstracts on Human factors in computing systems*, ACM Press, Ft. Lauderdale, Florida, USA, 2003.
9. Katona, G. *Organizing and memorizing*. Columbia University Press, New York, 1940.
10. Pirolli, P. and Card, S.K. Information foraging. *Psychological Review*, 106. 643-675.
11. Pirolli, P., Card, S.K. and Van Der Wege, M.M. The effects of information scent on visual search in the Hyperbolic Tree Browser. *ACM Transactions on Computer-Human Interaction*, 10 (1). 20-53.
12. Pirolli, P., Fu, W., Chi, E. and Farahat, A., Information scent and Web navigation; Theory, models, and automated usability evaluation. in *Human-Computer Interaction International Conference*, (Las Vegas, NV, in press), Lawrence Erlbaum Associates.
13. Polson, P.G., Bovair, S. and Kieras, D., Transfer between text editors. in *Proceedings of CHI '87 Human Factors in Computing Systems and Graphics Interface Conference*, (New York, 1987), ACM, 27-32.
14. Singley, M.K. and Anderson, J.R. *Transfer of cognitive skill*. Harvard University Press, Cambridge, MA, 1989.
15. Werner, K., A comparison of information visualization methods. in *CHI '02 extended abstracts on Human factors in computing systems*, (Minneapolis, Minnesota, USA, 2002), ACM Press, 872-873.