Expertise and the perception of shape in information

Andrew Dillon*
School of Library and Information Science
Indiana University
Bloomington, IN 47405
email: adillon@indiana.edu
fax: (812) 855-6166
phone: (812) 855-6267

Dille Schaap
Dept. of Psychonomics
University of Amsterdam
The Netherlands

*to whom all correspondence should be addressed. Circulated by first author.
Accepted for publication in: *Journal of the American Society for Information Science*.

This item is not the definitive copy. Please use the following citation when referencing this material: Dillon, A. and Shaap, D. (1996) Expertise and the perception of structure in discourse. *Journal of the American Society for Information Science*, 47(10), 786-788.

Abstract

Ability to navigate an information space may be influenced by the presence or absence of certain embedded cues that users have learned to recognize. Experimental results are presented which indicate that experienced readers of certain academic journals are more capable than inexperienced readers in locating themselves in an information space in the absence of explicit structural cues.

Introduction

The transition from paper to electronic media is replete with difficulties despite the promise of liberation that seems to infuse even the technical literature on the subject (e.g., Nielsen, 1995). While hypermedia and digital libraries may afford the potential for rapid access to massive amounts of information, it remains an open question how the human user will ultimately respond to shifts in presentation that seemingly tax their capabilities to navigate and integrate information spaces (Dillon et al, 1993).

Research in domains such as discourse processing (van Dijk and Kintsch, 1983), reading Just and Carpenter, 1987) and hypertext usability (McKnight et al, 1990) all points to the conclusion that human information usage involves cognitive integration of material beyond the word, sentence or picture level. Part of this integrative process seems to involve the acquisition of structural representations of the layout and organization of the information (see e.g., Rothkopf 1971 for an early finding). Such work as led to the current conception of interaction with information as a form of navigation through a space, and the hope that diverse theoretical perspectives and empirical findings on such topics can be drawn together to inform the development of digital environments.

One crucial aspect of navigation, thus construed, is the individual's perception of structure in space. Drawing on work in physical environment navigation, Dillon (1991) argued that people may learn to perceive structure in information on the basis of repeated exposure, extracting or attuning to regularities in the material they read. This work demonstrated that when presented with isolated sections of text from academic journals conforming to the superstructure of introduction, method, results and discussion, expert users were capable of correctly allocating text to sections at 80% accuracy levels. Such accuracy did not seem to correlate with a user's recall of the material read. Dillon concluded from this that experts were capable of identifying their location in this information space on the basis of rapidly searching the information scene in front of them for embedded cues such as statistical formulae, references etc. and that such an activity occurred independently of, or perhaps in advance of, reading for comprehension.

One weakness of this interpretation was the absence of data for non-expert users. Indeed, original reviews of the Dillon (1991) study pointed out that failure to investigate the role of experience was a potential weakness of the navigational perspective based on structural knowledge development. The present paper explicitly tackles that issue by reporting sample results from an ongoing series of studies into how people learn to conceptualize the shape of information.

Method

Forty-eight subjects ranging in experience from novice to expert readers of the academic journal literature in cognitive science, participated in the present series of studies. These were respondents to a general call across students and faculty at Indiana University and each was paid \$10 for participating. Experience was assessed by a series of questions probing academic qualifications, reading experience, authoring experience, research and experimental experience and background. Resulting classification grouped subjects into novices ($n=19$), intermediaries ($n=16$) and experts ($n=13$), distinguished largely in terms of reading habits with respect to the literature being examined and disciplinary background.

All subjects performed three tasks, only one of which need concern us here. This involved viewing paragraphs of text from target journals presented one at a time on screen. Subjects were required to allocate the text to one of the four main categories of introduction, method, results or discussion, working as rapidly as possible. Automatic data-logging routines captured time per screen, target location and subject's allocation.

Paragraph contents were manipulated so that half of all material in each section contained the exact replica of the text from the published papers, the other half had certain embedded cues removed. These were: presence of references to other literature (e.g., Jenkins (1993)), statistical formulae (e.g., $t_5=3.14$, $p<.05$), acronyms of tests or research tools (e.g., EPI), and explicit references to terms introduction, method, results and discussion. No headings (explicit structural cues) were presented in any treatment.

The trial took place in an experimental room at SLIS, with the subject seated facing an Apple MAC IIci. The experimenter was present at all times and subjects indicated responses by selecting one of four buttons presented on screen using HyperCard. All subjects received practice screens and an opportunity to ask questions before starting the experiment. The hypotheses driving this specific aspect of the research was that experience would significantly determine accuracy.

Results

Accuracy and speed results from a 3-way ANOVA (category x cue x experience) are presented here in summary form. The results clearly demonstrate the role of experience in the formation of accurate models of the information structure.

<u>Source</u>	<u>df</u>	<u>SSQ</u>	<u>MSQ</u>	<u>F</u>	<u>p</u>
Category	3	233.55	77.85	14.6	<.001
Cue	1	1.83	1.83	1.07	.306
Expertise	2	241.94	120.97	17.93	<.001
Cat x Cue	3	365.03	121.68	52.37	<.001
Cat x Exp	6	113.59	18.93	3.55	<.01
Cue x Exp	2	9.35	4.68	2.73	.076
Cat x Cue x Exp	6	12.13	2.02	0.87	.519

Table 1. ANOVA partial summary table for errors per category and cue combination per subject group

The effect of experience was in the hypothesized direction for errors, with experts performing more accurately in all categories except the method text, an interesting finding that may reflect the presence of more intuitive cues for location in these paragraphs. The cue variable on its own is not a major predictor of accuracy, though it interacts significantly with category and approaches significance in its interaction with expertise. These data are summarized in figure 1 (the + and - signs refer to presence or absence of cues).

Insert Figure 1 here.

No main effect for speed and expertise was found ($F_{2,45} = 0.68$, $p > .5$) although there was a cue by expertise interaction ($F_{2,45} = 5.91$, $p < .01$), suggesting that the presence or absence of cues affected user groups differentially, as one would expect. Indeed, for novices, the absence of cues led to faster mean response (12.31 sec. v. 12.00 sec.) per card. For both intermediaries (12.81 sec. v. 14.00 sec.) and experts (11.58 sec. v. 12.40 sec.) the absence of cues led to longer decision times per card. Coupled with the accuracy results, these data are in line with a cue-based response strategy on the part of more knowledgeable users, and a more serial reading strategy on the part of novices.

Discussion

Ability to sense where one is in information space is a defining characteristic of expert information users. This ability seems to be reliant in part on the rapid identification of details present in the information space. In the case of the experimental article format, such details include, but are clearly not limited to, the presence of statistical or numerical data, references to other research, and acronyms. Novices or inexperienced users of an information type are likely to need help acquiring this sense of location, the precise form of which is the topic of on-going research. These results support Dillon's (1991) interpretation of the role of experience in the perception of structure.

References

- Nielsen, J. (1995) *Multimedia and Hypertext: The Internet and Beyond*. New York: Academic Press Professional
- Dillon A. (1991) Readers' models of text structures: the case of academic articles *International Journal of Man-Machine Studies* 35, 913-925.
- Dillon, A., Richardson, J. and McKnight, C. (1993) Space -the final chapter: or why physical representations are not semantic intentions. In C. McKnight, A. Dillon and J. Richardson (Eds.) *Hypertext: A Psychological Perspective*. Chichester: Ellis Horwood, 169-192.
- van Dijk, T. and Kintsch, W. (1983), *Strategies of Discourse Comprehension*, New York: Academic Press.
- Just, M., and Carpenter, P. (1987) *The Psychology of Reading and Language Comprehension*. Boston: Allyn and Bacon.
- McKnight, C., Dillon, A., and Richardson, J. (1990) A comparison of linear and hypertext formats in information retrieval. In R. McAleese and C. Green, *HYPERTEXT: State of the Art*, Oxford: Intellect, 10-19.
- Rothkopf, E. Z. (1971) Incidental memory for location of information in text. *Journal of Verbal Learning and Verbal Behavior*, 10, 608-613.

Acknowledgments

This research was funded in part by a grant from the Indiana University/University of Amsterdam Cognitive Science Exchange Program. Subject expenses were funded by a grant from the Indiana University School of Library and Information Science Faculty Research Fund. Data-logging facilities were developed by Geoff McKim. The authors also wish to acknowledge the useful comments of the anonymous reviewer.