*Int. J. Human-Computer Studies* (2001) **55**, 761–786
doi:10.1006/ijhc.2001.0505
Available online at http://www.idealibrary.com on IDEAL®

# Graphic and numerical methods to assess navigation in hypertext

JOHN E. MCENEANEY
*Department of Reading and Language Arts, Oakland University, Rochester, MI 48309-4494, USA. email: mceneane@oakland.edu*

User navigation has been a central theme in both theoretical and empirical work since the earliest days of hypertext research and development. Studies exploring user navigation have, however, tended to rely on indirect navigational measures and have rarely tried to relate navigation to performance-solving problems or locating information. This paper proposes methods that lead to a more direct representation and analysis of user movement in hypertext and empirically explores the relationship of these measures to performance in a hypertext search task. Results of the study indicate that the proposed graphical and numerical methods have empirical significance and may be useful in assessing and modeling user navigation.


KEYWORDS: comprehension; empirical validation; navigational metrics; navigational patterns; path analysis; user paths; visualization.

## 1. Introduction

Although traditional print technologies are likely to be with us for many years to come, electronic print media are assuming major roles in the way we access and use information. Libraries are retiring printed card catalogs, software vendors ship documentation in CD-ROM format, and large public and private collections of documents are appearing on the web. There is ample evidence, however, that the transition from traditional print to electronic media has cognitive consequences and that simply making information electronically available will not assure it is effectively or efficiently used.

Research suggests, for instance, that hypertext materials are more cognitively demanding (Conklin, 1987; Egan, Remde, Gomez, Landauer, Eberhardt & Lochbaum, 1989) or require a greater degree of higher-level relational processing (Wenger & Payne, 1996) than traditional print. There are consistent findings even highly skilled readers of print experience orientation problems as they move around within hypertext networks (e.g. Van Dam, 1988; Neilsen, 1989; Edwards & Hardman, 1989). Moreover, with the dramatic increase in the use of electronic information fostered by the rapid development of the web, most of which is accounted for by less experienced computer users, the potential for wasted time and bandwidth is enormous (Nielsen, 1993; Nielsen & Sano, 1995).

In response to problems related to hypertext navigation, researchers and developers have created a variety of tools. Visual tools have been particularly popular among

1071-5819/01/110761 + 26 $35.00/0 

developers seeking to support user navigation. Site maps are now widely used and there is evidence that users find them helpful in navigating and in establishing a clearer idea of the organizational structure of a site (Utting & Yankelovich, 1989; Chen & Rada, 1996). In larger networks where complete site maps are impractical, fish-eye views (Furnas, 1986; Sarker & Brown, 1994; Bartram, Ho, Dill & Henigman, 1995; Pirolli, Card & Van Der Wege, 2001), clustering techniques that organize nodes into meaningful groups (Gloor, 1991; Mukherjea, Foley & Hudson, 1995), and a variety of other filtering and mapping techniques (Husemann, Petersen, Kanty, Kochs & Hase, 1997; Neves, 1997) have been developed to assist both in development and use of large-scale hypertext networks.

Numerical metrics defined as ratios of actual to possible links (Astleitner & Leutner, 1996) and links to nodes (Boyle & Hor Teh, 1992) have been proposed to help assess global properties of hypertext that may be related to difficulties users experience. More complex numerical metrics have been imported from social network theory, where algorithms designed to identify cliques in social groups can be applied to identify node clusters (Astleitner and Leutner, 1996), and novel metrics have been designed specifically to address the needs and perspectives of hypertext designers and researchers (Botafogo, Rivlin & Shneiderman, 1992).

There has also been interest in developing a better understanding of how users navigate hypertext under various task and environmental conditions. Some of these studies have employed "static" measures related to numbers of nodes or links accessed or measures of time and path length (Qiu, 1994; Schroeder & Grabowski, 1995; Melara, 1996). Others have analysed selected episodes of movement, tabulating navigation within or across sections of a network (Schroeder & Grabowski, 1995). Some investigators have applied statistical techniques to identify clusters of nodes and interpret user navigation in terms of these constructs (Lawless & Kulikowich, 1996). Other statistical approaches include collapsing data from large numbers of users into state transition probability tables that are used as a basis for analysis (Chen, 1997), and identification of statistical benchmarks that might be useful in developing theoretical models (Neilsen, 1989; Qiu, 1994).

Recently, there have also been studies based on models originally developed to account for spatial search tasks in physical environments, revealing surprisingly deep connections with search in hypertext. One recent study (Titus & Everett,1996) highlights striking parallels between the cognitive search strategies involved in "wayfinding" (Passini, 1984) in a complex architectural space and other forms of information gathering, including search tasks in hypertext. Even biological models that treat searching for information as "foraging" are now finding useful application in the study of hypertext (e.g. Pirolli, 1998; Pirolli & Card, 1999; Pirolli *et al*., 2001).

Perhaps the most consistent trend in user-navigation research, however, is the use of navigational paths as data. Although the concept of a path has been referred to using a variety of terms including "route" (Canter, Rivers & Storrs, 1985), "log file" (Barab, Bowdish, Young & Owen, 1996) and "audit trail" (Misanchuk & Schwier, 1992; Schroeder & Grabowski, 1995), these data sets all record the sequence of nodes visited by a reader in a hypertext session and often also include measures of time related to visits. That user paths are commonly employed in navigational studies should be no surprise. Although a time-stamped path misses deliberations that go into a users' decision-making (e.g. a pointer hovers momentarily over one link before moving on to another that is clicked), a path represents the single most complete measure of user navigation and thus affords an important window on the search process and the strategies readers apply in acquiring information (Lawless & Kulikowich, 1996). Moreover, since data can be recorded and formatted in an unobtrusive manner on-the-fly, this approach provides empirical investigators with a powerful data-collection tool that has the additional benefit of being seamlessly integrated with the delivery of experimental materials.

For all of the interest in navigational paths, however, there have been relatively few studies that have sought to examine the relationship between patterns of navigation and search outcome measures. Two exceptions are studies by Smith (1996) and Cardle (1994), who investigated the relationship of informal measures of search success (i.e. displays of frustration or confidence by subjects) with a path-efficiency measure. Both studies, however, are based on static measures that do not attempt to incorporate spatial or temporal features of user paths. In addition, the reliance on informal assessments of user success along with the incorporation of an outcome measure as a part of the proposed path measure undermines the general validity and objectivity of their metric.

Three other studies that sought to relate path data to outcome measures have been based on similarly indirect path measures. Chang and McDaniel (1995) relied on video transcripts of their subjects and accompanying think-aloud data to subjectively categorize navigational patterns. Pirolli, Pitkow and Rao (1996) and Chen (1997) adopted a more quantitative focus on user paths but employed large-scale aggregate data drawn from web-server statistics. By using these aggregated data sets, however, the "paths" studied had to be inferred from access logs and were thus subject to problems related to firewalls, proxy masking of user identity, intentional reloading of documents by users and missed hits as a result of local browser caches. Although discrete static measures, informal descriptive characterizations and aggregated summarizations of user navigation are useful starting points, it seems likely that there is something to be gained from methods that support a more direct analysis of user paths and consider the relationship between paths and outcome measures. It is the purpose of this paper to achieve these ends—to define methods that support the analysis of user paths and to empirically assess associations between these measures and hypertext search outcomes.

This paper analyses user paths in two different ways. One form of analysis is based on a graphical method intended to illustrate both individual and group paths in a way that makes navigational patterns visually distinct. The second form of analysis relies on path-specific structural metrics that are related to the visually distinct categories identified by graphic analysis. Following the definition of these path-specific structural metrics, the paper reports on two empirical studies that explore the relationship between the proposed analyses and quantitative outcome measures. The paper concludes by considering limitations of the proposed methods and measures and suggesting some potential applications.

## 2. Visualization techniques and associated metrics

The framework developed to assess and visualize user navigation is based on the traditional node-and-link model of hypertext. Although not without drawbacks (e.g. Stotts & Furuta, 1989; Turine, de Oliveira & Masiero, 1997), this model of hypertext has proven to be a useful conceptual framework from both theoretical and empirical perspectives. Central to this framework is the idea that a hypertext network can be formalized by nodes representing content and links representing structure. At least part of the popularity of the node-and-link model can be attributed to two simple but powerful formalisms that support analysis in this model: adjacency matrices that are well suited to computational analysis and labeled directed graphs (digraphs) that present structural information in a readily interpreted visual format. Briefly, an adjacency matrix is a table that records each unique link in a hypertext document. Typically, an adjacency matrix consists of a table of zeroes and ones with labeled rows and columns. A "1" in cell (a, b) indicates a link from node "a" to node "b". A "0" in a cell indicates that there is no *direct* link between the two nodes. In Figure 1(a), for instance, the two entries of "1" in the first row indicate that there is a direct link from node 0 to nodes 3 and 5. Zeroes appear in all other positions in this row because no other direct links are present. Figure 1(b) illustrates the digraph that corresponds to this adjacency matrix, demonstrating that the network consists of eight nodes (numbered 0–7) and 13 links (the double-headed arrow connecting node 0 and node 3 represents two links).

What makes these formalisms powerful is that they retain mathematical and visual simplicity yet capture important structural features that can be transformed to create other useful measures, including distance matrices, centrality metrics and other general measures of connectedness and linearity (Botafogo *et al*., 1992). Graphs and digraphs are widely used in hypertext research as a basis for visual analyses that can reveal structure that is difficult to discern in numerical formats (McDonald, Paap & McDonald, 1990; Mukherjea and Foley, 1995; Chen, 1997).

Two metrics of special interest in this investigation are compactness and stratum (Botafogo *et al*., 1992). The purpose of these metrics is to yield global network-based assessments of structure that are grounded in node-based centrality and status measures, with each metric ranging from zero to one. Compactness refers to the connectedness of a network, yielding values close to zero for sparsely linked networks and values close to one for densely connected networks. Stratum, on the other hand, refers to the degree of

(a)

| To / From | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 2 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 3 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 5 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 7 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |

(b)

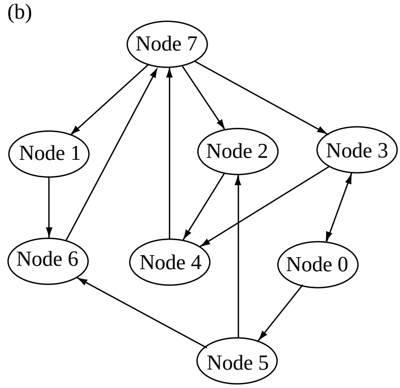



FIGURE 1. An adjacency matrix (a) for an 8-node hypertext network and its corresponding graphical representation (b).

(a)

| To<br>From | a | b | c | d | e | f | g |
|---|---|---|---|---|---|---|---|
| a | 0 | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ |
| b | 1 | 0 | 1 | ∞ | ∞ | 1 | 2 |
| c | ∞ | ∞ | 0 | ∞ | ∞ | ∞ | ∞ |
| d | 1 | ∞ | ∞ | 0 | ∞ | ∞ | ∞ |
| e | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| f | 2 | ∞ | ∞ | ∞ | ∞ | 0 | 1 |
| g | 1 | ∞ | ∞ | ∞ | ∞ | ∞ | 0 |

(b)

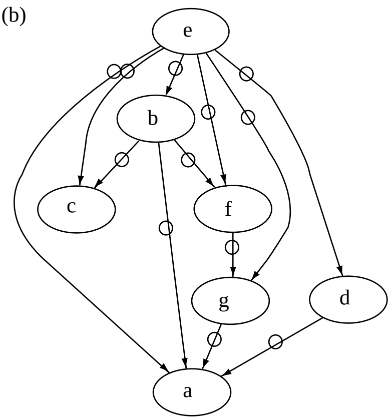



FIGURE 2. A distance matrix (a) and digraph (b) representing a network with nodes that are inaccessible from other nodes.

linearity of a network, as indicated by the extent to which a network is organized so that certain nodes must be read before others. More linear networks have stratum values closer to one, while those less linear are closer to zero.

Both compactness and stratum are based on transformations of an adjacency matrix. One of the important intermediate derivations in this transformation is the distance matrix where the entry in cell (a, b) indicates the minimum number of link traversals required to move from node "a" to node "b" (Floyd, 1962). The distance from a node to itself is considered to be zero. When a node cannot be reached from another, the distance between these nodes is taken to be infinite ($\infty$). Consider, for example, the network in Figure 2(b). Unlike the network in Figure 1, not every node can be reached from every other node. As a result of the more limited scope of movement in this network, some of the entries in the distance matrix have the value "$\infty$" [see Figure 2(a)].

The occurrence of infinite values in cell entries can, however, create problems in manipulating the matrix and, as a result, a converted distance matrix $C$ is defined by means of substitutions in the original distance matrix $D$ such that

$$C_{ij} = \begin{cases} D_{ij} & \text{if } D_{ij} \neq \infty, \\ K & \text{otherwise,} \end{cases} \tag{1}$$

where K is a finite conversion constant commonly assigned a value equal to the total number of nodes in the hypertext (Botafogo *et al*., 1992, 1994; De Bra, 1996), resulting in the converted distance matrix in Table 1 for the original distance matrix in Figure 2(a).

### 2.1. MEASURES OF NODE CENTRALITY

One important application of the converted distance matrix is in defining centrality measures that reflect how prominent any given node is within the larger structure of

TABLE 1

*A converted distance matrix for the network in Figure 2. This table also reports absolute (COD & CID) and relative (ROC & RIC) measures of node centrality in the right-marginal columns and bottom row*

| | To | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| From | a | b | c | d | e | f | g | COD | ROC |
| a | 0 | 7 | 7 | 7 | 7 | 7 | 7 | 42 | 5.0 |
| b | 1 | 0 | 1 | 7 | 7 | 1 | 2 | 19 | 11.2 |
| c | 7 | 7 | 0 | 7 | 7 | 7 | 7 | 42 | 5.0 |
| d | 1 | 7 | 7 | 0 | 7 | 7 | 7 | 36 | 5.9 |
| e | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 | 35.3 |
| f | 2 | 7 | 7 | 7 | 7 | 0 | 1 | 31 | 6.8 |
| g | 1 | 7 | 7 | 7 | 7 | 7 | 0 | 36 | 5.9 |
| *CID* | 13 | 36 | 30 | 36 | 42 | 30 | 25 | 212 | |
| *RIC* | 16.0 | 5.0 | 7.0 | 5.0 | 5.0 | 7.0 | 8.0 | | |

a network, either as a destination or as a point of departure. Prominent destinations (i.e. reference nodes) are those that have many paths leading to them. Prominent points of departure (i.e. index nodes) have many paths leading away from them. The concept of centrality, therefore, incorporates aspects of both connectedness and directionality and, not surprisingly, is related to how nodes are used in locating information (McKnight, Dillon & Richardson, 1990; Dillon, McKnight & Richardson, 1993).

Centrality measures include both node-based absolute metrics that are specific to a node in a network and more general measures intended to support comparisons across different networks. Absolute node-based measures of centrality that are important in the present context include converted in-distance (*CID*) and converted out-distance (*COD*), measures of the distance separating a node from other nodes in a network when viewing that node as either a destination or a point of departure. *COD* is calculated by summing across the row entries for a node in the converted distance matrix ($COD_i = \sum_j C_{ij}$) since row entries indicate distances *from* the chosen node *to* others. The *CID* of a node provides the conceptual complement of the *COD*, indicating the distance of a node from other nodes in a system when considering that node as a possible *destination*. *CID* is calculated by summing down the column for a given node in the converted distance matrix (i.e. $CID_j = \sum_i C_{ij}$).

Two important limitations of *COD* and *CID* as defined, however, are that these measures do not provide any network-wide assessment of structure, nor do they support comparisons to other hypertexts. As Botafogo *et al.* (1992) have noted, "a node with *COD* 200 in a hypertext with 1000 nodes might be much more central than a node with *COD* 50 in a hypertext with 100 nodes (p. 147)". When there is an interest in making comparisons across networks, it is necessary to employ *relative* rather than absolute measures in order to account for differences (e.g. size) between the networks compared. It is useful, therefore, to define a network-based metric that allows *COD* and *CID* values to be normalized and that is the purpose of defining the converted distance (*CD*) for a network. Unlike the node-based *COD* and *CID* metrics, *CD* is a global measure that is based on the network as a whole, rather than on specific nodes. In effect, *CD* is a measure of both the degree of linking in a network and its size since it is calculated by summing all entries in a converted distance matrix (i.e. $CD = \sum_i \sum_j C_{ij}$).

With $CD$, it is possible to define relative measures of centrality that support meaningful comparisons across different network structures. Relative out-centrality and relative in-centrality are, respectively, the normalized versions of $COD$ and $CID$ for a node and are defined as $ROC_i = CD/COD_i$, and $RIC_i = CD/CID_i$. Values for these normalized measures based on the digraph in Figure 2 are reported in the bottom rows and right-most columns of Table 1. Note that while $COD$ and $CID$ are inversely related to centrality, $ROC$ and $RIC$ are directly proportional, with higher values for $ROC$ and $RIC$ indicating greater degrees of centrality.

### 2.2. COMPACTNESS ($Cp$)

A network with $Cp = 1$ is fully connected—every node is connected to every other node. A network with $Cp = 0$ is completely disconnected; there are no links at all, only nodes. As the number of links in a network ranges between 0 and $n^2 - n$ (the maximum possible for a network of n nodes), $Cp$ varies between 0 and 1. More formally,

$$Cp = \frac{(CD_{\text{Max}} - \sum_i \sum_j C_{ij})}{(CD_{\text{Max}} - CD_{\text{Min}})}, \tag{2}$$

where $CD_{\text{Max}}$ and $CD_{\text{Min}}$ refer, respectively, to the maximum and minimum values a converted distance matrix can assume for completely connected ($CD_{\text{Min}}$) and completely disconnected ($CD_{\text{Max}}$) networks. $CD_{\text{Max}}$ and $CD_{\text{Min}}$ are, in turn, given by

$$CD_{\text{Max}} = K(n^2 - n) \quad \text{and} \quad CD_{\text{Min}} = (n^2 - n) \tag{3,4}$$

as a result of the substitution of the conversion constant $K$ for infinite values, and the fact that there are a maximum of $n^2 - n$ links.

### 2.3. STRATUM ($St$)

The stratum metric is designed to reflect the linear ordering implicit in the structure of a network. Recall that in defining $COD$ and $CID$ it was pointed out that these measures reflect the suitability of a node as either a point of departure or as a destination. The idea behind the stratum metric is to capture the extent to which point-of-departure nodes, destination nodes and those in between can be identified solely on the basis of network structure. As with the $Cp$ metric, stratum values for networks range from zero to one, with $St = 1$ indicating a strictly linear sequence of nodes (that allows one and only one path within the network), while $St = 0$ when a network is fully connected (every node is connected to every other node) and there is no structural basis for distinguishing points of departure and destinations. The relationship between $Cp$ and $St$ is not as simple as this example might suggest, however. While $Cp$ and $St$ are not independent, they do not vary in a simple fashion, as revealed by the two digraphs in Figure 3 that differ only by a single link from node d to node a. The addition of this one link creates a closed path or cycle that alters $St$ dramatically (from 1 to 0) while having only a small influence on $Cp$ (increasing it from 0.44 to 0.50). While the one additional link does not substantially

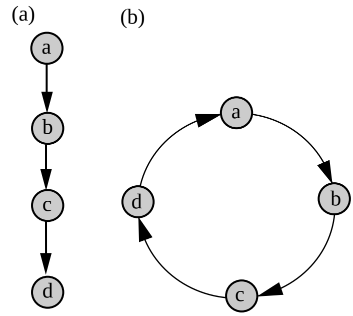



FIGURE 3. Two networks with similar compactness but radically different stratum values.

increase the connectedness of the network, it radically alters opportunities for navigation. In the original strictly linear network only one starting point (node a) assured that every node in the network could be reached. The structure of the network thus unambiguously identifies node a as the most appropriate starting point. The addition of the link from "d" to "a", however, now makes every node an equally suitable point of departure—every node can now be reached from every other and the overall distance required in each case is the same regardless of the starting node. Whereas *Cp* focuses on connectedness in a general way, *St* focuses on the consequences of structure for movement within the network.

Calculation of the *St* metric begins with the distance matrix and two concepts originally defined in social network theory: status and contrastatus (Harary, 1959). One common application of these terms is in social networks with an established system for seniority or "pecking order". In this kind of system, the status of an individual refers to the number of persons who are subordinate to that individual, while contratstatus refers to the amount of status weighing down on an individual from "above", and is calculated by summing numbers of *superordinates* (i.e. "bosses") for an individual. Prestige is defined as the status of an individual minus the contrastatus of that individual, and the absolute prestige for a network is calculated by summing the absolute values of finite prestige values for all the nodes in the network. Unlike the status and contrastatus measures, prestige ranges across both positive and negative values resulting in large positive values for individuals high in the pecking order, prestige at or near 0 for individuals in the middle of the pecking order, and "large" negative values for prestige among individuals low in the pecking order. Absolute prestige, as a sum of absolute values, is necessarily positive. Table 2 presents a distance matrix for the network illustrated in Figure 2 that includes these stratum-related measures. Note that status and contrastatus are calculated as $COD$ and $CID$ were, by summing across rows (status) and down columns (contrastatus), ignoring infinite values which indicate that nodes are unconnected. Moreover, in a distance matrix without infinite values, $COD = \text{status}$ and $CID = \text{contrastatus}$.

As was the case with $COD$ and $CID$, however, these stratum-related measures are absolute and therefore do not provide a suitable basis for comparisons across networks, which require a normalized measure. To that end, Botafogo *et al.* (1992) define the linear

TABLE 2

*A distance matrix for the network illustrated in Figure 2 with status (S), contrastatus (CS) and prestige (P) for each node and the absolute prestige (AP) for the network*

| | To | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| From | a | b | c | d | e | f | g | *S* | *P* |
| a | 0 | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ | 0 | − 6 |
| b | 1 | 0 | 1 | ∞ | ∞ | 1 | 2 | 5 | 4 |
| c | ∞ | ∞ | 0 | ∞ | ∞ | ∞ | ∞ | 0 | − 2 |
| d | 1 | ∞ | ∞ | 0 | ∞ | ∞ | ∞ | 1 | 0 |
| e | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 | 6 |
| f | 2 | ∞ | ∞ | ∞ | ∞ | 0 | 1 | 3 | 3 |
| g | 1 | ∞ | ∞ | ∞ | ∞ | ∞ | 0 | 1 | − 3 |
| *CS* | 6 | 1 | 2 | 1 | 0 | 2 | 4 | *AP* = 24 | |

absolute prestige (*LAP*) of a network with n nodes, showing that

$$LAP = \begin{cases} \dfrac{n^3}{4}, & \text{if } n \text{ is even,} \\ \dfrac{n^3 - n}{4} & \text{if } n \text{ is odd.} \end{cases} \quad (5)$$

and then go on to formally define the Stratum (*St*) of a network as

$$St = \frac{absolute\ \ prestige}{LAP}. \quad (6)$$

## 3. Visualizing and assessing path metrics

Although the compactness and stratum metrics were originally developed as tools to assist designers of hypertext avoid certain well-known problems, they have since been proposed as tools with potential application in assisting users to navigate hypertext systems (Rivlin, Botafogo & Shneiderman, 1994). In this section, adaptations of the compactness and stratum metrics are developed with the purpose of characterizing user movement (i.e. paths) in hypertext.

### 3.1. THE CONCEPT OF A PATH

Consider the set of nodes in a hypertext as represented by **P**, or by extension $\{p_1, p_2, p_3, \ldots, p_n\}$, where $n$ represents the total number of pages in the document. A path is defined as the sequence of nodes or pages loaded by a user in a reading session. Although this definition diverges somewhat from the graph-theoretic notion of a path (since it allows nodes to appear more than once in a path), this use of the term path is consistent with the way this term is commonly used in the literature (e.g. Pirolli *et al*., 1996; Eklund & Zeiliger, 1996; Cockburn & Jones, 1996), its divergence from the underlying graph theoretic model notwithstanding. More formally, a path is defined as a mapping of the set **P** onto the set of natural numbers, resulting in an ordered set or sequence $\langle p_i, p_j, p_k, \ldots, p_v \rangle$ consisting of as many elements as there were node visits.

A *path matrix* represents frequencies of transitions from each node to every other node within a given path (similar to that proposed by Pirolli *et al.* 1996). Moreover, if the hypertext under examination is closed, a path matrix can be normalized by representing every node in the hypertext, regardless of whether it appears in the path. As a result of this expansion, it is possible to sum individual user paths into group "paths". The normalizing expansion of the path matrix is achieved by inserting rows and columns filled with zeroes in the appropriate places in the path matrix so that every node in the hypertext is represented in the expanded path matrix [see Figure 4(d)]. The net result of this expansion is to embed the path within the larger structure of the hypertext. The structural features of the original path matrix are, however, preserved while establishing a normal form for all paths. As a result of this normalization, individual path matrices can be summed to yield group matrices.

Calculation of path metrics, like their structural counterparts, requires a distance matrix and suitable conversions. The user path in Figure 4(a), for example, consists of 17 transitions, beginning and ending at node 6 with a path matrix constructed by taking each from–to pair in the path (i.e. $6 \rightarrow 30$, $30 \rightarrow 6$, $6 \rightarrow 37$, etc.) and incrementing the appropriate cell in the path matrix. The resulting path matrix [Figure 4(b)] indicates the number of transitions from each node to every other node in the path. The distance matrix and converted distance matrix for a path are created by straight-forward adaptations of the procedures proposed by Botafogo *et al.* (1992), with one important change related to defining the conversion constant $K$. Creation of a distance matrix for a path begins by substituting a value of "1" for all entries in the original path matrix that exceed one (i.e. that represent multiple transitions across a link), in effect, creating a path "adjacency matrix". A path distance matrix can then be computed and a converted matrix created by replacing all zero cell entries (except those in the matrix diagonal) with the conversion constant $K$, where $K$ equals the number of nodes in the *path matrix*. The path diagram [Figure 4(c)] is constructed from the path adjacency matrix by creating a vertex for each node represented and an arc for each non-zero cell entry.

The simplicity of the path matrix is not without cost, however, as sequential information in the original path record may be lost. Nonetheless, it has the advantage of preserving the straight-forward graphic interpretation and the computational flexibility of the matrices defined for hypertext networks. Each node visited and each link traversed in a path is unambiguously represented in a path matrix, preserving the general structure of user movement. Moreover, as will be shown in what follows, this simplified view remains informative enough to make meaningful qualitative and quantitative distinctions between user paths.

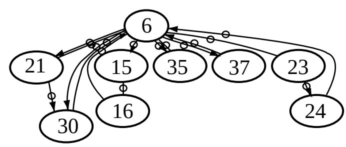
### 3.2. CALCULATING PATH METRICS

Although the path metrics developed in this paper are direct adaptations of the compactness and stratum metrics described by Botafogo *et al.* (1992, 1994), there is an important difference between path matrices and their structural counterparts. The difference has to do with whether metrics and diagrams are based on the *path matrix* or the *expanded path matrix* (which corresponds to the network matrix used in structural measures).

(a) Path = < 6, 30, 6, 37, 6, 15, 16, 6, 21, 6, 21, 30, 6, 23, 24, 6, 35, 6 >

| To / From | 6 | 15 | 16 | 21 | 23 | 24 | 30 | 35 | 37 |
|---|---|---|---|---|---|---|---|---|---|
| 6 | 0 | 1 | 0 | 2 | 1 | 0 | 1 | 1 | 1 |
| 15 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 23 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 24 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 30 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 35 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 37 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

(b)




(c)

| To / From | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 23 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 28 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 29 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

(d)

| To / From | 6 | 15 | 16 | 21 | 23 | 24 | 30 | 35 | 37 |
|---|---|---|---|---|---|---|---|---|---|
| 6 | 0 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 |
| 15 | 2 | 0 | 1 | 3 | 3 | 4 | 3 | 3 | 3 |
| 16 | 1 | 2 | 0 | 2 | 2 | 3 | 2 | 2 | 2 |
| 21 | 1 | 2 | 3 | 0 | 2 | 3 | 1 | 2 | 2 |
| 23 | 2 | 3 | 4 | 3 | 0 | 4 | 3 | 3 | 3 |
| 24 | 1 | 2 | 3 | 2 | 2 | 0 | 2 | 2 | 2 |
| 30 | 1 | 2 | 3 | 2 | 2 | 3 | 0 | 2 | 2 |
| 35 | 1 | 2 | 3 | 2 | 2 | 3 | 2 | 0 | 2 |
| 37 | 1 | 2 | 3 | 2 | 2 | 3 | 2 | 2 | 0 |

(e)

FIGURE 4. A path (a) and its corresponding path matrix (b), path diagram (c), expanded path matrix for a portion (nodes 21–30) of the path matrix [outlined within (b)] (d), and converted distance matrix for the path in (a) (e).

Moreover, since the order of a path matrix and its expanded path matrix can differ substantially, this choice has significant consequences for path visualization and the calculation of metrics.

With respect to visualization, the present study bases the order of a path diagram on the number of distinct nodes in a *path* rather than on the number of nodes in the network. This choice simplifies the visual display significantly, since the resulting diagram will not be crowded with unvisited nodes. It also has an important computational consequence, since it avoids flooding subsequent calculations with the conversion constant $K$ across all rows and columns for nodes not visited. Although a complete representation of nodes is clearly needed in assessing *network* structure, trying to distinguish user paths within a visual display of the much larger network seems ill-advised. The very nature of path analysis, particularly when it attempts to treat individual users, suggests that ignoring unvisited nodes is a better approach since any attempts to account for the influence of unvisited nodes is speculative at best. It, therefore, seems theoretically justifiable as well as computationally convenient to base path diagrams on the smaller path matrix, without the excess baggage of simultaneously accounting for the larger network.

The numerical side of this issue has to do with the basis for normalizing metrics. Recall that the definitions for $Cp$ and $St$ both involve normalizing measures. The formula for $Cp$ is normalized by establishing maximum and minimum values for $CD$ based on fully connected and fully disconnected networks of the same size as the network of interest. In St, linear absolute prestige ($LAP$) serves as the normalizing measure, again based on the number of nodes in the network of interest. The problem of normalization arises in path metrics because it is not immediately apparent how the size or order of the path network should be defined. If a reader adopts a path that fully explores a hypertext (i.e. every link is tried at least once), the order of the path network and the order of the hypertext are equivalent. In networks of any reasonable size, however, readers will typically traverse only a small portion of the links that are present, resulting in a path network that is smaller (perhaps a great deal smaller) than the hypertext. These two possibilities suggest two alternatives. One alternative is to continue to use the hypertext network as the basis for normalization, adopting the values used in calculating structural metrics. The alternative is to normalize the network according to characteristics of the path.

Although normalization on the basis of the hypertext structure, rather than the path, might be useful if there is interest in examining how different hypertext structures lead to different user paths, if the object of study is the user path within a given hypertext, it seems clear that normalization should be based on the path rather than the network since the idea behind such a metric is to distinguish different types of paths. For these reasons, normalization of individual path metrics is based on the number of distinct nodes represented in the path matrix, rather than on the nodes in the network. Given this, it is possible to define the specialized path metrics alluded to earlier. Specifically, path compactness ($P_{Cp}$) is defined as

$$P_{Cp} = \frac{(PCD_{\mathrm{Max}} - \sum_i \sum_j PC_{ij})}{(PCD_{\mathrm{Max}} - PCD_{\mathrm{Min}})}, \tag{7}$$

where $PC$ refers to the converted path matrix and $PCD_{\text{Max}}$ and $PCD_{\text{Min}}$ refer, respectively, to the maximum and minimum $CD$ values that the path matrix could assume for completely connected ($PCD_{\text{Min}}$) and completely disconnected ($PCD_{\text{Max}}$) path networks of a given size. $PCD_{\text{Max}}$ and $PCD_{\text{Min}}$ are defined in a manner analogous to $CD_{\text{Max}}$ and $CD_{\text{Min}}$ with

$$PCD_{\text{Max}} = K(n^2 - n) \quad \text{and} \quad PCD_{\text{Min}} = (n^2 - n), \tag{8,9}$$

where $n$ is the order of the path matrix and $K$ is the conversion constant.

The path stratum metric is defined as

$$P_{St} = \frac{path\ \ absolute\ \ prestige}{P_{LAP}}, \tag{10}$$

where the linear absolute prestige of the path ($P_{LAP}$) is defined analogously to Eq. (5):

$$P_{LAP} = \begin{cases} \dfrac{n^3}{4} & \text{if } n \text{ is even.} \\ \dfrac{n^3 - n}{4} & \text{if } n \text{ is odd.} \end{cases} \tag{11}$$

with $n$ representing the order of the path matrix.

## 4. Empirical validation of the proposed methods and metrics

This section describes the results of two independent empirical studies designed to assess whether the proposed path measures successfully distinguish navigational strategies adopted by users. More specifically, the focus of the studies is to determine whether the proposed graphic techniques and path metrics can be shown to be associated with success in a hypertext search task.

### 4.1. STUDY 1

Participants in Study 1 included 90 students at a medium-sized midwestern public university in the US. The experiment required subjects to respond to a set of academic advising questions using an electronic student advising handbook. Subjects answered as many questions as possible within a 15-min period. The handbook consisted of approximately 31 000 words in 78 text nodes structured in a hierarchical–linear fashion with major handbook divisions organized hierarchically and nodes within those divisions organized in a linear fashion. The handbook was intended to duplicate the content and overall structure of a print version that had been in use for a number of years.

The hypertext handbook was developed as an HTML document and presented using a JavaScript-based modification of the Netscape browser illustrated in Figure 5. The browser consists of three horizontally arrayed panels that eliminate all standard Netscape navigational aids so that subject movement through passages can be recorded and more effectively managed by the system. The narrow top panel (a) provides a title bar, "NEXT" and "BACK" buttons that support sequential movement through the document, a "Contents" button linked to a main table of contents, and a "Help" button that

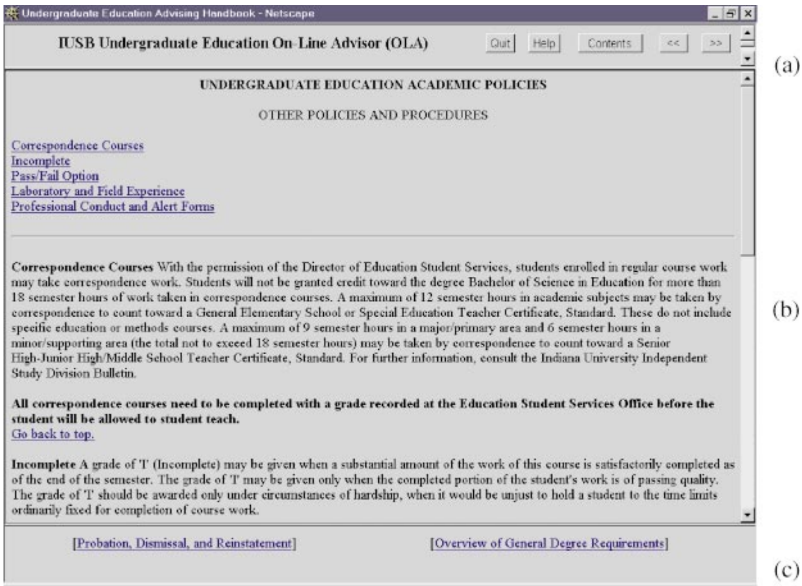



FIGURE 5. Illustration of the three-panel browser interface employed in the validation studies with the title bar/navigation panel at the top (a), a link panel at the bottom (c), and a scrollable passage viewer panel in the central position (b).

provides access to a page explaining navigational tools and features. Handbook content is presented in the large middle panel (b) with a right-side scroll bar appearing as needed. Links to other nodes are presented in the link panel (c) at the bottom. When subjects selected links, the viewer wrote records of the subject's path and time spent on each page to the browser cookie file.

Following completion of the experimental sessions, cookie files were retrieved and path data were extracted. Mathematica (Wolfram, 1999) routines were developed to create appropriately formatted graph files to display path diagrams using GraphViz 1.3 (Ellson, Koutsofios & North, 1998). Mathematica routines were also employed to calculate path compactness and path stratum metrics as described above. Subject responses to academic advising questions were scored on the basis of information provided in the handbook. Scoring procedures followed practices employed by McKnight *et al.* (1990) and Smith (1996), with zero points for incorrect and omitted responses, one half point for partially correct and correct-but-incomplete responses, and one point for complete and correct responses.

A preliminary graphical analysis was based on a subset ($n = 24$) of the total subject pool. In the first step of the graphical analysis, subjects were grouped according to their success in answering questions using the hypertext handbook. The "high" group consisted of students with the top three scores in each of four counterbalanced groups. Subjects had been divided into these groups to account for possible effects by question set (there were two question sets used) and presentation order (half of the subjects used a print version of the handbook first, while the other half used the hypertext version first) (McEneaney, 1998). The "low" group, on the other hand, consisted of those subjects with the lowest three scores in each of the counterbalanced groups, resulting in two groups of 12 that differed according to their success in carrying out the hypertext task. The next step in the analysis was to create and review path diagrams for the subjects in the two groups, with the intent of discerning visually distinctive patterns that might be related to success in the reading task.

*4.1.1. Results of graphical analyses in Study 1.* In reviewing path diagrams, it was apparent that there were visually salient features that seemed to be related to success in the experimental task. A number of examples of these distinctive path diagrams for individual high- and low-scoring subjects are presented in Figure 6. Review of the path diagrams suggested that subjects whose scores on the search task were low tended to assume a "passive" approach to locating answers in the handbook, similar to the "passive search mode" described by Titus and Everett (1996, p. 271) in characterizing information search. Specifically, it appeared that these passive readers relied much more heavily on a "page turning" strategy that followed the print version reading sequence embedded in the "next" and "previous" buttons, while subjects who did well on the search task adopted a more strategic approach. As a result, path diagrams for low-scoring subjects tended to display distinctively linear patterns of movement, while high-scoring subjects tended to produce more shallow hierarchical patterns of movement with the handbook table of contents serving as the root of their navigational tree, indicating repeated visits to the table of contents during the course of the browsing session.

Moreover, as indicated in Figures 7 and 8, similar navigational patterns resulted when group paths were generated by summing individual path matrices for high- and low-scoring subjects. In these diagrams, the path matrices for individual subjects were expanded, resulting in normalized path matrices that were summed. Since some link traversals probably represent unique idiosyncratic user decisions and navigational

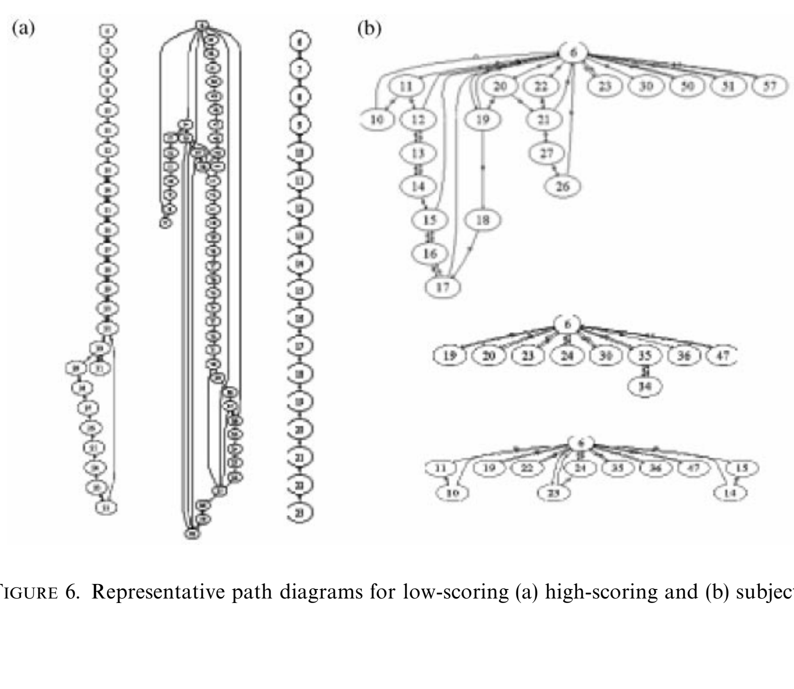



FIGURE 6. Representative path diagrams for low-scoring (a) high-scoring and (b) subjects.

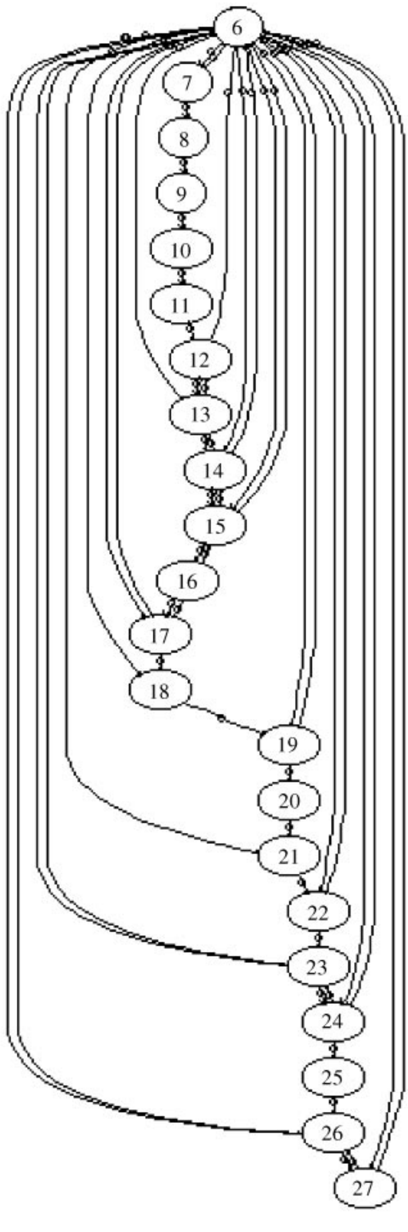



FIGURE 7. Group path diagram for low-scoring hypertext readers with the minimal path transition threshold set equal to 3.

errors, "noise" was eliminated by setting a threshold that had to be met in order for a group traversal to be displayed in the group path diagrams. In Figures 7 and 8, the threshold is set equal to 3, with the result that only those links are displayed that were traversed at least three times by the subjects in a group. While setting thresholds to other

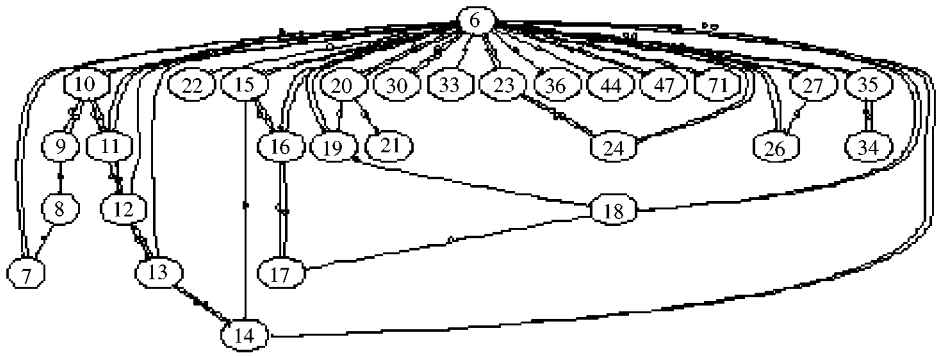



FIGURE 8. Group path diagram for high-scoring hypertext readers with the minimal path transition threshold set equal to 3.

values (5, 2, 1, etc.) altered group path diagrams in minor ways (mainly by increasing the number of nodes and links displayed), these alternative settings did not alter the characteristic linear and shallow hierarchical patterns associated with the low- and high-scoring groups.

The graphical analyses provide fairly compelling, if informal, evidence in support of distinctive navigational patterns associated with subjects' search strategies and resulting outcome measures. The more effective use of the hypertext handbook is associated with a shallow hierarchical path diagram that results from subjects making repeated trips back to the main table of contents to make decisions about how to locate information in the electronic handbook. Less effective use of the handbook is associated with a more passive linear path diagram that reflects users' reliance on sequential "page-turning" with users hoping to locate desired information by simply coming across it in their browsing.

These graphical analyses are also suggestive about what we might expect to find when we examine the association between the path metrics that have been defined and hypertext task scores. Specifically, the linear character of the low-scoring group suggests that path stratum will correlate negatively with subjects' hypertext scores, since low-scoring subjects seemed more likely to adopt linear navigational paths. Conversely, since the compactness of a bidirectional star pattern (topologically equivalent to what has been described as a shallow hierarchical pattern) tends to approach 1 as the number of nodes increases, while both bidirectional cycles and linear patterns approach values less than 1 (Botafogo *et al.* 1992), path compactness should correlate positively with hypertext scores. Accordingly, the research hypotheses explored in the quantitative analyses are as follows.

(1) The path compactness metric will correlate significantly, in a positive fashion, with subjects' hypertext search scores.
(2) The path stratum metric will correlate significantly, in an inverse fashion, with subjects' hypertext search scores.

*4.1.2. Results of Study 1 analyses.* Pearson correlation coefficients were determined for the experimental variables and hypertext search scores. Results of the analyses are

TABLE 3
*Correlations of experimental variables with hypertext task scores in Study 1 (with p values and numbers of subjects). All tests are two-tailed*

| | *r* | *p* | *n* |
|---|---|---|---|
| Print ability | 0.210 | 0.137 | 29 |
| Computer experience | 0.026 | 0.404 | 90 |
| Pages viewed | − 0.010 | 0.464 | 90 |
| Order of path matrix | 0.90 | 0.400 | 90 |
| Path compactness | 0.239† | 0.024 | 90 |
| Path stratum | − 2.05 | 0.052 | 90 |

† $p < 0.05$.

indicated in Table 3, with significant correlations flagged. All analyses employed two-tailed tests with $\alpha = 0.05$. Note that, although print reading ability, computer experience, and both total pages viewed and distinct pages viewed (i.e. order of the path matrix) failed to correlate with the hypertext reading score, path compactness correlated significantly and path stratum failed to achieve significance by only the slimmest of margins. Moreover, these correlations were as expected, with path compactness exhibiting a significant positive correlation and path stratum exhibiting an inverse (but marginally non-significant) correlation. These analyses suggest that the observed relationships are not likely to be the result of chance, and thus support the interpretation of path diagrams and the proposed metrics as empirically meaningful and potentially important measures of hypertext navigation.

*4.1.3. Discussion of results from Study 1.* Results of Study 1 suggest that informal visual analysis of navigational paths can provide important clues about readers' success in a hypertext search task. Less successful readers tended to adopt a more passive page-turning approach when locating information, resulting in more linear path diagrams. More successful readers, on the other hand, adopted a more strategic approach, resulting in more hierarchical patterns with one or more index or landmark nodes at the root. Subjects' search strategies appear to correspond to patterns noted in other recent studies that identify "passive vs. active" (Titus & Everett, 1996), "matching vs. exploration" (Thiel & Müller, 1996), and "browsing vs. search" (Agosti, 1996) approaches, depending on whether subjects rely on a built-in general search strategy (i.e. passive, matching and browsing approaches) or seek out a less obvious, but potentially more relevant, structural framework to support their search (i.e. active, exploration and search approaches.)

Statistical analyses provide support for the claim that the graphical and numerical methods described are related to success in the hypertext search task. Moreover, confirmation of specific directional predictions supports the proposed graphical interpretations of the metrics. The absence of significant correlations with print reading scores and prior computer experience highlights the significance of the observed correlations between the two path metrics and success in the hypertext task. It also, however, suggests that the power of the design (Cohen, 1988) may have been inadequate to discern associations. In an effort to address this potential problem, a second study was conducted using the same design, but with a larger subject pool.

### 4.2. STUDY 2

Participants in Study 2 included 133 students at a medium-sized midwestern public university in the US. Of these 133 students, print reading scores were available for 48. As in Study 1, subjects responded to a set of academic advising questions using an electronic student advising handbook, answering as many questions as possible within a 15-min period. Subject navigation in the handbook was stored to the browser cookie file and subsequently retrieved and reformatted for analysis. Path compactness and stratum metrics were computed for each subject and responses to academic advising questions were assigned full, partial, or no credit on the basis of the criteria used in Study 1.

Since the primary purpose of Study 2 was to enhance the statistical power of the design used in Study 1, analysis focused on quantitative correlational tests. As in Study 1, there were measures of print reading ability, computer experience, path length (i.e. total pages viewed) and order of the path matrix (i.e. unique pages viewed), along with the two path metrics (path compactness and path stratum.) In addition, a measure representing the number of times subjects loaded the handbook table of contents was included. As in Study 1, there was an expectation that the path compactness metric would correlate positively with hypertext scores, while the path stratum metric would correlate negatively with this score. All statistical tests were two-tailed with $\alpha = 0.05$.

*4.2.1. Results of analyses in Study 2.* Pearson correlation coefficients were computed for all variables, including path compactness and path stratum. Results are indicated in Table 4, with significant correlations noted. Results of the analyses suggest that statistical power was a factor in the earlier findings. Although Study 1 analyses indicated print reading ability and prior computing experience did not correlate significantly with the hypertext task, there were significant correlations in Study 2. Success in the search task also correlated significantly with use of the handbook table of contents, although this variable accounted for less variance than the print reading measure and the two path metrics. Correlations of path compactness and stratum with the outcome measure increased from Study 1, with computed $p$ values less than 0.001. There were no significant correlations between success in the task and either overall path length or the order of the path matrix (as was the case in Study 1).

TABLE 4

*Correlations of experimental variables with hypertext task scores in Study 2 (with p values and numbers of subjects indicated). All tests are two-tailed*

| | $r$ | $p$ | $n$ |
|---|---|---|---|
| Print ability | 0.477† | 0.001 | 48 |
| Computer experience | 0.294† | 0.001 | 133 |
| Pages viewed | 0.079 | 0.366 | 133 |
| Order of path matrix | 0.032 | 0.711 | 133 |
| Table of contents hits | 0.269 | 0.002 | 133 |
| Path compactness | 0.375† | 0.000 | 133 |
| Path stratum | − 0.348† | 0.000 | 133 |

† $p < 0.05$.

A follow-up stepwise multiple regression was applied to determine whether multiple variables might account for additional variance. As suggested by its large correlation with the task, the print reading measure was the first variable selected, and the resulting regression equation accounted for a significant amount of variance [$F(47) = 13.583$, $p = 0.001$]. None of the remaining variables contributed significant additional variance, however. In subsequent analyses, manually entering other variables first, followed by those remaining, resulted in the same outcome—there was no increase in the variance obtained beyond that resulting from regressing on a single variable.

*4.2.2. Discussion of results from Study 2.* Results from Study 2 suggest that the absence of significant correlations for print reading ability and prior computer experience with the hypertext reading task in Study 1 were likely due to inadequate statistical power. In Study 2, where statistical power was enhanced, these correlations were evident, with print reading ability accounting for the greatest degree of variance in predicting the outcome measure. In addition, Study 2 results reinforce the associations noted with path metrics. As in the first study, both metrics correlated significantly with the outcome measure in the expected manner. Although the table of contents variable added in Study 2 correlated significantly with the hypertext task, this correlation was weaker than either of the two path metrics and the follow-up multiple regression analyses suggest that the table of contents measure draws its predictive power from the same source as the path metrics, a predictable outcome given the role the table of contents page plays in navigational patterns.

In a more unexpected outcome, the multiple regression analyses showed that a single source of variance likely accounted for the predictive power of *all* the variables in the study. Apparently, these variables tap a common source of variance and, although somewhat speculative, the most promising candidate seems to be the orienting strategy subjects adopt when faced with a complex cognitive task. As noted earlier, numerous studies of information seeking have suggested that subjects tend to adopt one of two approaches in a variety of search tasks, with some subjects adopting a strategy of active engagement, while others rely on more passive strategies that require little effort, but may have only limited relevance to the task at hand. Moreover, similar active/passive strategy options have also been consistently noted in the literature on reading in print, with more effective readers typically adopting a more strategic, metacognitive approach to text, while less able readers tend to read with a poorly defined sense of purpose or little critical engagement (e.g. Barton, 1997; Gourgey, 1999; Lifford, Byron & Jean Ziemian, 2000; Taraban, Rynearson & Kerr, 2000).

While it is clear that there continues to be a pressing need for detailed empirically grounded theories of navigation that focus on particulars specific to hypertext (McKnight *et al.*, 1990), there may also be a need for a broader theory that can account for the ways people orient their problem solving when faced with complex tasks. The numerous studies cited above suggest a fundamental dispositional characteristic that may lead some subjects to more active engagement, while others are content to rely on more generic strategies related to the way the problem is presented (in this case, in the design of the interface.) There is also evidence that this fundamental orientation can interact with a subject's capacity to self-regulate or manage learning, with the somewhat surprising result that less effective self-regulators may actually benefit more from a passive orientation than they do from an active one (Beishuizen, Stoutjesdijk & Van Putten, 1994). Ultimately, understanding particulars of search in hypertext may require us to step back from navigational details far enough to see the broader dispositions and foundational metacognitive skills that people bring to complex learning tasks such as searching, reading or learning in hypertext.

## 5. General discussion and limitations

Results of the empirical validation suggest that the proposed methods and metrics can be productively applied in assessing user navigation. Moreover, the results reported suggest that navigational patterns and their associated metrics may be useful as indirect measures of user strategy and perhaps even of users' success in cognitively modeling the domain represented by a hypertext. If, as cognitive flexibility theory suggests, learning in hypertext materials involves the cognitive reconstruction of a domain space through repeated traversals across that space (Spiro & Jehng, 1990; Jacobsen & Spiro, 1995), the paths users choose are sure to have a powerful influence on learning outcomes in hypertext. In the present study, subjects who adopted shallow, hierarchical search strategies that more accurately modeled the organization of the hypertext materials were more successful in their search, while those who adopted more linear paths through the materials were less successful. In effect, more successful subjects recognized and took advantage of the structure of the domain space by returning to the "higher ground" of the table of contents and the broader cognitive view it afforded of the material presented.

That more successful hypertext users recognize and take advantage of the structure of the materials they are using is not very surprising. That, of course, is the purpose of graphic overviews, site maps and other techniques that have been shown to promote the more effective use of hypertext. One important aspect of the path record, however, is that this information is immediately and unobtrusively available *during* reading, something that is not generally true of user reading ability measures and impossible in the case of outcome measures. Given the demonstrated association of path information and outcome measures, it may be possible to apply real-time path data in support of user models that will lead to more effective adaptive hypertext systems. It may also be possible to apply these metrics in designing user paths to meet particular objectives or needs of users. Even in the absence of ready applications of the methods and metrics proposed, however, it seems worthwhile to explore these measures of user movement, given the wide use of other, less direct measures.

While the results of the studies are relatively clear-cut, three limitations suggest that these findings should be considered preliminary. One limitation has to do with the strength of the observed association between path metrics and hypertext task scores. A second limitation has to do with the choice that has been made with regard to normalization, and a third set of related limitations is associated with constraints imposed by the design of the study and experimental materials.

Although the observed associations between the proposed metrics and subjects' hypertext scores is not likely to be due to random variation, the strength of the association is not great. A weak-to-moderate association remains of significant interest, but it also suggests that this variable should be considered within a larger explanatory context. Regarding normalization, it is relevant to note that Botafogo *et al.* (1992)

recognize the general nature of the normalization problem in their work establishing structural metrics. In response to this problem, they suggest that alternative normalization procedures be considered, particularly for the stratum metric. They note, for instance, that since stratum depends on *LAP*, a measure that increases in the order of $O(n^3)$, it may be problematic to compare networks that have large differences in numbers of nodes. With the selection of the path matrix (rather than the expanded path matrix) as the basis for normalization, however, it is almost certain that there will be variation in the values of LAP used to normalize path stratum for subjects whose paths are shorter or longer. Given this, the question arises as to whether the variation in the order (i.e. the size) of path matrices is sufficient to call into questions the analyses that have been carried out.

Two circumstances of the present study suggest, however, that the problem of normalization has not compromised the specific results reported. One circumstance is that, although there was variation in the size of the path matrices used to calculate path stratum values, there was no significant correlation between subjects' search scores and the size of their associated path matrices either in Study 1 (Pearson $r = -0.011$, $p = 0.921$) or in Study 2 (Pearson $r = 0.032$, $p = 0.711$). Although path matrices varied, there was no systematic variation that might suggest an influence on the relationship noted between the outcome measure and path stratum.

The second circumstance is related to the observation of Botafogo *et al.* (1992, pp. 169–170) that differences in stratum values resulted when index and reference nodes were excluded from the stratum calculation, the implication being that the presence of prominent nodes could distort the stratum metric. In an effort to determine whether this should be a source of concern in the present studies, path stratum values were recalculated, excluding index and reference nodes (where these were identified as nodes with in- and out-degrees that differed from the network mean by more than two standard deviations.) Following the recalculation of path stratum values, paired *t*-tests were carried out to assess whether the original stratum values differed from those in which index and reference nodes had been excluded. As indicated in Table 5, there was no significant change in stratum means, supporting the conclusion that the path stratum metric had not been distorted significantly by the use of index and reference nodes.

The limitation having to do with the design of the validation studies arises because the investigations described here focus on the relationship between the proposed metrics and outcome measures using a correlational, rather than a causal design. Given the exploratory character of the studies, however, an experimental design that seeks to match subjects and manipulate treatments seems premature. Future work may well introduce genuinely experimental designs exploring these metrics. For the present, more naturalistic studies offer greater benefits.

Finally, it is important to note that the present investigation was limited to a test of the proposed methods and metrics utilizing a single hypertext accessed by a specific browser interface. At present, it cannot be known whether the proposed methods and metrics will work equally well in other hypertext structures or under different browsing conditions. It seems clear that the observed patterns of navigation reflect structural features of the hypertext as well as navigational dispositions of users and that what constitutes effective and efficient hypertext use is almost certainly related to the structure of a hypertext, which inevitably sets conditions within which readers operate. Earlier work has

TABLE 5

*Results of paired-samples t-tests comparing original stratum values (OS) to those computed with index and reference nodes excluded (RS). All tests are two-tailed*

| *Study* 1 | | | | | |
|---|---|---|---|---|---|
| Variable | | Mean | *N* | S.D. | |
| Original stratum (*OS*1) | | 0.2467297 | 90 | 0.2330525 | |
| Recalculated stratum (*RS*1) | | 0.2407829 | 90 | 0.2653631 | |
| Difference | Mean | S.D. | *t* | *df* | *p* |
| *OS*1 − RS1 | − 0.005947 | 0.2051451 | 0.275 | 89 | 0.784 |
| *Study* 2 | | | | | |
| Variable | | Mean | *N* | S.D. | |
| Original stratum (*OS*2) | | 0.2516914 | 133 | 0.2596466 | |
| Recalculated stratum (*RS*2) | | 0.2809128 | 133 | 0.2334345 | |
| Difference | Mean | S.D. | *t* | *df* | *p* |
| *OS*2 − *RS*2 | − 0.0292 | 0.2331489 | − 1.445 | 132 | 0.151 |

suggested that it may be important to understand how users select and apply strategies when faced with complex information-gathering tasks. The importance of the two studies described in this paper resides in the "window" the proposed metrics and graphic techniques provide on the ways strategies are selected and applied. While it might not be obvious how specific hypertext structures influence user movement, it seems fairly clear that the metrics and techniques proposed in this paper are well suited to the exploration of this and other related issues in user navigation.

Limitations notwithstanding, the proposed methods and metrics afford hypertext developers and researchers a number of important benefits. One benefit is that these methods and metrics support a more direct analysis of user movement in hypertext than has been possible before. A second benefit is that the concepts and computational framework these methods and metrics are based on are natural extensions of prior methods and metrics developed to analyse the structure of hypertext, and thus support a more general perspective whether one considers hypertext structure or movement within such structures. Finally, there are both informal and quantitative reasons for confidence since the metrics that have been proposed are clearly related to the graphical displays developed and these metrics have been shown to have significant empirical association with the successful use of hypertext materials.

## References

AGOSTI, M. (1996). An overview of hypertext. In M. AGOSTI & A. SMEATON, Eds. *Information Retrieval and Hypertext*. Boston: Kluwer Academic Publishers.

ASTLEITNER, H. & LEUTNER, D. (1996). Applying standard network analysis to hypermedia systems: Implications for learning. *Journal of Educational Computing Research*, **14,** 285–303.

BARAB, S. A., BOWDISH, B. E., YOUNG, M. F. & OWEN, S. V. (1996). Understanding kiosk navigation: using log files to capture hypermedia searches. *Instructional Science*, **24,** 377–395.

BARTRAM, L., HO, A., DILL, J. & HENIGMAN, F. (1995). The continuous zoom: a constrained fisheye technique for viewing and navigating large information spaces. *Proceedings, User Interface and Software Technology* 95, pp. 207–215. New York: ACM.

BARTON, M. L. (1997). Addressing the literacy crisis: teaching reading in the content areas. *NASSP Bulletin*, **81,** 22–30.

BEISHUIZEN, J., STOUTJESDIJK, E. & VAN PUTTEN, K. (1994). Studying textbooks: effects of learning styles, study task, and instruction. *Learning and Instruction*, **4,** 151–174.

BOTAFOGO, R. A., RIVLIN, E. & SHNEIDERMAN, B. (1992). Structural analysis of hypertexts: identifying hierarchies and useful metrics. *ACM Transactions on Information Systems*, **10,** 142–180.

BOYLE, C. & HOR THE, S. (1992). To link or not to link: An empirical comparison of hypertext linking strategies. *Proceedings of the 10th annual international conference on systems documentation, Ottowa, Canada*. New York: ACM.

CANTER, D., RIVERS, R. & STORRS, G. (1985). Characterizing user navigation through complex data structures. *Behaviour and Information Technology*, **4,** 93–102.

CARDLE, N. T. (1994). *A HyperCard on Celtic history to assess navigability measures for hypertext*. MSc. Dissertation, Institute of Information Technology, University of Nottingham, UK.

CHANG, C. & MCDANIEL, E. D. (1995). Informal search strategies in loosely structured settings. *Journal of Educational Computing Research*, **12,** 95–107.

CHEN, C. (1997). Structuralizing and visualizing the WWW by generalised similarity analysis. In M. BERNSTEIN, L. CARR, & K. ØSTERBYE, Eds. *Proceedings of the Eighth ACM Conference on Hypertext - Hypertext* 97, 177–186. New York: ACM.

CHEN, C. & RADA, R. (1996). Interacting with hypertext: a meta-analysis of experimental studies. *Human–Computer Interaction*, **11,** 125–156.

COCKBURN, A. & JONES, S. (1996). Which way now? Analysing and easing inadequacies in WWW navigation. *International Journal of Human-Computer Studies*, **45,** 105–129.

COHEN, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum.

CONKLIN, J. (1987). Hypertext: An introduction and survey. *IEEE Computer*, **20,** 17–41.

DE BRA, P. (1996). *Hypermedia structures and systems* [on-line]. Available: *http://wwwis.win.tue.nl/2L670/static/index.html.*

DILLON, A., MCKNIGHT, C. & RICHARDSON, J. (1993). Space, the final chapter or why physical representations are not semantic intentions. In C. MCKNIGHT, A. DILLON & J. RICHARDSON, J. Eds. *Hypertext: A Psychological Perspective*. Chichester: Ellis Horwood.

EDWARDS, D. M. & HARDMAN, L. (1989). Lost in hyperspace: cognitive mapping and navigation in a hypertext environment. In R. MCALEESE, Ed. *Hypertext: Theory into Practice*. Oxford: Intellect Limited.

EGAN, D. E., REMDE, J. R., GOMEZ, L. M., LANDAUER, T. K., EBERHARDT, J. & LOCHBAUM, C. C. (1989). Formative design-evaluation of SuperBook. *ACM Transactions on Information Systems*, **7,** 30–57.

EKLUND, J. & ZEILIGER, R. (1996). Navigating the web: possibilities and practicalities for adaptive navigational support. *Proceedings of AUSWEB'96, Gold Coast, Australia.*

ELLSON, J., KOUTSOFIOS, E. & NORTH, S. (1998). *GraphViz 1.3* [computer software]. Murray hill, NJ: Lucent Technologies. Available: http://www.research.att.com/sw/tools/graphviz/

FLOYD, R. W. (1962). Algorithm 97: shortest path. *Communications of the ACM*, **5,** 345.

FURNAS, G. (1986). Generalized fisheye views. *Proceedings of CHI 1986, Human Factors in Computing Systems*, pp. 16–23. Boston, April.

GLOOR, P. A. (1991). Cybermap, Yet another way of navigating in hyperspace. *Proceedings of the Third ACM Conference on Hypertext, Hypertext 91*, pp. 107–121. San Antonio, Texas, December.

GOURGEY, A. F. (1999). Teaching reading from a metacognitive perspective: theory and classroom experiences. *Journal of College Reading and Learning*, **30,** 85–93.

HARARY, F. (1959). Status and contrastatus. *Sociometry*, **22,** 23–43.

HUSEMANN, H., PETERSEN, J., KANTY, C., KOCHS, H. & HASE, P. (1997). A user adaptive navigation metaphor to connect and rate the coherence of terms and complex objects.

In M. Bernstein, L. Carr, & K. Østerbye, Eds. *Proceedings of the Eighth ACM Conference on Hypertext—Hypertext 1997*, 177–186. New York: ACM.

Jacobson, M. & Spiro, R. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: an empirical investigation. *Journal of Educational Computing Research*, **12,** 301–333.

Lawless, K. A. & Kulikowich, J. M. (1996). Understanding hypertext navigation through cluster analysis. *Journal of Educational Computing Research*, **14,** 385–399.

Lifford, J., Byron, B. E. & Jean Ziemian, C. (2000) Reading, responding, reflecting. *English Journal*, **89,** 46–57

McEneaney, J. (1998). Are less able readers disadvantaged by reading in electronic environments? *Proceedings of the IEEE* 1998 *International Symposium on Technology and Society*, pp. 28–32. South Bend, U.S.A.

McDonald, J. E., Paap, K. R. & McDonald, D. R. (1990). Hypertext perspectives: Using Pathfinder to build hypertext systems. In R. W. Schvaneveldt, Ed. *Pathfinder Associative Networks*: *Studies in Knowledge Organization*. Norwood, NJ: Ablex.

McKnight, C., Dillon, A. & Richardson, J. (1990). A comparison of linear and hypertext formats in information retrieval. In R. McAleese & C. Green, Eds. *Hypertext*: *State of the Art*, pp. 10–19. Oxford: Intellect.

Melara, G. E. (1996). Investigating learning styles on different hypertext environments: Hierarchical-like and network-like structures. *Journal of Educational Computing Research*, **14,** 313–328.

Misanchuk, E. R. & Schwier, R. A. (1992). Representing interactive multimedia and hypermedia audit trails. *Journal of Educational Multimedia and Hypermedia*, **1,** 355–372.

Mukherjea, S. & Foley, J. D. (1995). *Visualizing the world-wide web with the navigational view builder*. Paper presented at the Third International WWW Conference, 10–14 April. Darmstadt, Germany. Available: http://www.igd.fhg.de/www/ www95/ papers/44/mukh/ mukh.html.

Mukherjea, S., Foley, J. D. & Hudson, S. (1995). Visualizing complex hypermedia networks through multiple hierarchical views. *Conference Proceedings on Human Factors in Computing Systems*, *Denver*, *U.S.A*. New York: ACM.

Neves, F. D. (1997). The Aleph: a tool to spatially represent user knowledge about the WWW docuverse. In M. Bernstein, L. Carr, & K. Østerbye, Eds. *Proceedings of the Eighth ACM Conference on Hypertext – Hypertext 97*, 197–207. New York: ACM.

Nielsen, J. (1989). *The matters that really matter for hypertext usability*. *Proceedings of the ACM Conference on Hypertext—Hypertext 89*, 239–248. New York: ACM.

Nielsen, J. (1993). *Multimedia and Hypertext*. Boston: Academic Press.

Nielsen, J. & Sano, D. (1995). *Interface design for Sun's WWW site* [on-line]. *http://www.sun.com/sun-on-net/unidesign/*.

Passini, R. (1984). *Wayfinding in Architecture*. New York: Van Nostrand Reinhold Company.

Pirolli, P. (1998). Exploring browser design trade-offs using a dynamical model of optimal information foraging. *Proceedings of the SIG-CHI on Human Factors in Computing Systems* 1998, 33–40. New York, ACM.

Pirolli, P. & Card, S. K. (1999). Information foraging. *Psychological Review*, **106,** 643–675.

Pirolli, P., Card, S. K. & Van Der Wege, M. M. (2001). Visual information foraging in a focus + context visualization. In J. Jacko & A. Sears, Eds. *Proceedings of the SIG-CHI on Human Factors in Computing Systems 2001*, 506–513. New York, ACM.

Pirolli, P., Pitkow, J. & Rao, R. (1996). *Silk from a sow = s ear*: *extracting usable structures from the web* [on-line]. Available: http://www.acm.org/sigchi/chi96/proceedings/papers/Pirolli2/pp2.html.

Qiu, L. (1994). Frequency distributions of hypertext path patterns: a pragmatic approach. *Information Processing & Management*, **30,** 131–140.

Rivlin, E., Botafogo, R. & Shneiderman, B. (1994). Navigating in hyperspace: Designing a structure-based toolbox. *Communications of the ACM*, **37,** 87–96.

Sarker, M. & Brown, M. H. (1994). Graphical fisheye views. *Communications of the ACM*, **37,** 73–84.

SCHROEDER, E. E. & GRABOWSKI, B. L. (1995). Patterns of exploration and learning with hypermedia. *Journal of Educational Computing Research*, **13,** 313–335.

SMITH, P. A. (1996). Towards a practical measure of hypertext usability. *Interacting with Computers*, **8,** 365–381.

SPIRO, R. & JEHNG, J. (1990). Cognitive flexibility and hypertext: theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. NIX & R. SPIRO, Eds. *Cognition*, *Education*, *Multimedia*. Hillsdale, NJ: Lawrence Erlbaum.

STOTTS, P. D. & FURUTA, R. (1989). Petri-net-based hypertext: document structure with browsing semantics. *ACM Transactions on Information Systems*, **7,** 3–29.

TARABAN, R., RYNEARSON, K. & KERR, M. S. (2000). Metacognition and freshman academic performance. *Journal of Developmental Education*, **24,** 12–20.

THIEL, U. & MÜLLER, A. (1996). "Why was this item retrieved?": new ways to explore retrieval results. In M. AGOSTI & A. SMEATON, Eds. *Information Retrieval and Hypertext*. Boston: Kluwer Academic publishers.

TITUS, P. A. & EVERETT, P. B. (1996). Consumer wayfinding tasks, strategies, and errors: An exploratory field study. *Psychology and Marketing*, **13,** 265–290.

TURINE, M. A. S., OLIVEIRA, M. C. F. & MASIERO, P. C. (1997). A navigation-oriented hypertext model based on state charts. In M. BERNSTEIN, L. CARR, & K. ØSTERBYE, Eds. *Proceedings of the Eighth ACM Conference on Hypertext—Hypertext* 97, 177–186. New York: ACM.

UTTING, K. & YANKELOVICH, N. (1989). Context and orientation hypermedia networks. *ACM Transactions on Office Information Systems*, **7,** 58–84.

VAN DAM, A. (1988). Hypertext 87 keynote address. *Communications of the ACM*, **31,** 887–895.

WENGER, M. J. & PAYNE, D. G. (1996). Comprehension and retention of nonlinear text: Considerations of working memory and material-appropriate processing. *American Journal of Psychology*, **109,** 93–130.

WOLFRAM, S. (1999). *The Mathematica Book*, 4th edn. Cambridge: Cambridge University Press.

Paper accepted for publication by Associate Editor, Dr A. Dillon