

# Why structure and genre matter for users of digital information: A longitudinal experiment with readers of a web-based newspaper

Misha W. Vaughan<sup>a,\*</sup>, Andrew Dillon<sup>b</sup>

<sup>a</sup>*Applications User Experience Group, Oracle Corporation, 500 Oracle Parkway, Redwood Shores, CA 94065, USA*

<sup>b</sup>*School of Information, The University of Texas at Austin, USA*

Received 22 July 2002; received in revised form 28 October 2005; accepted 1 November 2005

Available online 27 December 2005

Communicated by K.S. Severinson Eklundh

## Abstract

In an effort to understand the impact of designing for digital genres on users' mental representations of structure, a two-phase study was conducted. In phase 1, six expert news readers and a panel of HCI experts were solicited for input regarding genre-conforming and genre-violating web news page design, navigation, and story categorization. In phase 2, a longitudinal experiment with a group of 25 novice web news readers who were exposed to one of the two designs over 5 sessions is reported. During these sessions a variety of user data were captured, including: comprehension (recall, recognition), usability (time on task, accuracy, user satisfaction), and navigation (path length, category node hits).

The between-group difference of web site design was significant for comprehension, usability, and navigation with the users of the genre-conforming design demonstrating better performance. The within-group difference of time was significant across these three measures as well, with performance improving over time. No interaction effect was found between web site design and time on comprehension or usability. However, a surprising interaction effect was found on navigation; specifically the breadth of navigation (i.e. the number of nodes visited for two classes of tasks) increased over time more dramatically for the genre-violating group than for the genre-conforming group. By examining the changes in these data over time and between the two designs, evidence for the development of users' mental representations of structure was captured.

© 2005 Elsevier Ltd. All rights reserved.

## 1. Introduction

Designers of digital information spaces face the constant challenge of designing systems that meet diverse users' needs. In an attempt to better serve users, a variety of design techniques have been employed, including: use of metaphors such as the 'desktop'; use of communication tools such as on-line chat; and use of navigation aids such as on-line indices and overview maps. In spite of the promise these tools offer, users continue to face wildly varying interface designs, navigation schemes, and metaphors across information spaces. This complexity in the interface increases the amount of work required of users to

navigate and manipulate digital resources, an activity that is largely secondary to their primary task (Dillon, 2000).

One way to improve our understanding of information space design is to approach the problem from a new angle. In this paper we consider the design problem as one of finding patterns in users' representations of information spaces and learning how to support the natural pattern-making tendencies of human cognition. Users' representations of information must encompass both the physical movement through (or navigation of) an information space and the comprehension necessary for decision-making about a space and its contents.

Users' representations are grounded in their personal experiences and evolve as a result of interactions with information spaces over time. If a user interacts with an information space repeatedly, it is generally assumed that he or she builds up a mental representation of that space with special emphasis on the orderly or consistent elements

\*Corresponding author. Tel.: +1 650 607 0652.

E-mail addresses: [mvaughan@acm.org](mailto:mvaughan@acm.org), [misha.vaughan@oracle.com](mailto:misha.vaughan@oracle.com) (M.W. Vaughan), [adillon@ischool.utexas.edu](mailto:adillon@ischool.utexas.edu) (A. Dillon).

within a space (Dillon and Vaughan, 1997). When an information space reflects an order that is shared by more than one person, there is a shared understanding of users' representations of the information space. Ideally, we wish to create spaces that maximize the shared perception of order across users. In seeking to understand individuals' shared conceptions, we can conceive of the designers task as an attempt to extract the design essentials for an information space to provide a consistent user experience for a community and/or a certain set of tasks. Of particular interest to the present researchers are these 'design essentials for an information space.' These essentials, we believe, facilitate the development of an accurate mental representation of a given information space by helping the user to order and structure the experience.

We are most interested in examining the structural design elements that facilitate the development of users' shared conceptions of order and structure for an information space. To explore these we borrow heavily from genre theory, cognitive psychology, and hypertext research to frame our question of interest as: what structural genre conventions facilitate the development of mental representations of information spaces?

The following paper first outlines relevant concepts from genre theory and cognitive psychology to explore order and structure and how we make meaning. Next we turn to the hypertext literature as a mechanism to link this theory with practice—specifically, to bridge information science with mental representations of structure and form. Then we present our three primary hypotheses, examining two competing web site structures over time. We then provide a detailed discussion of the methodology of a two-phase study. In phase one we collected detailed requirements about web-based news genre conventions from expert news readers. In phase two we conducted a longitudinal experiment with a group of 25 novice web news readers who were exposed over time to one of two web newspaper designs (i.e. a genre-conforming version and a genre-violating version, based on the data from phase one). This presentation of methods is followed by results and discussion. Our data offer insights into how users react to different versions of an information space, manipulated to conform or violate certain expectations of structure.

## 2. Literature review

### 2.1. Genre theory

The concept of genre has been defined, used, and theorized by a variety of fields including folklore, popular culture, linguistics, and rhetoric. Each approach has taken genre into a unique domain of inquiry such as folklore's study of oral traditions across cultures (cf. Ben-Amos, 1976; Bauman, 1992), popular culture's concern with literary formulas (cf. Cawelti, 1976; Berger, 1992), linguistics' examination of verbal utterances (Bahktin, 1986; Hasan, 1996), and rhetoric's teachings on academic and

disciplinary communication (Swales, 1990; Berkenkotter and Huckin, 1995). These fields each offer varying definitions of genre, implicitly or explicitly, and each raises important issues to consider when examining regularities in many forms of expression. However, a composite definition might be that a genre can be considered a class of communicative events, which share a set of conventions and rules to facilitate interaction by creating and maintaining expectations among a community of producers and consumers. This definition suggests that regularities in users' conceptions may be studied as both the physical manifestation as it is perceived by its users, and as the social, historical, and cultural meaning as interpreted by users.

Bauman (1992) states that genre's purpose is, at its heart, communicative production and reception; genres are an unavoidable product of living in a social context. Individuals create, use, and re-use genres in every act of communication. Tonkin (1992) argues that genres are part of oral history telling in the individual case and part of communicative practice as a whole. Berkenkotter and Huckin (1995) argue that a genre is 'owned' by a community and that a genre is shaped by its norms, epistemology, ideology, and social ontology, which are made present through communication. Finally, Swales (1990) argues that genres are products of discourse communities, maintaining and sustaining discourse over time via communicative events.

A second purpose of genres, suggested first by Cawelti (1976) and then later by Berger (1992), is emotional satisfaction. Cawelti argues in the case of the popular formula story, generic purpose is to induce an emotional state of pleasure and enjoyment. He argues that "audiences find satisfaction and a basic emotional security in a familiar form" (p. 9).

Such analyses of genre posit evolving forms, genres that are created and refined gradually over time, with the assumed lifetime of a genre being expressed in decades or even centuries. With the emergence of digital forms, Dillon and Gushrowski (2000) suggest this may not be so, and that new digital genres have already emerged. To identify a timeline however requires us to determine a starting point without any clear agreement about how genres begin, beyond the general assumption of genre theorists that they emerge via social interaction, cultural negotiation, and change.

The notion of expertise and the role of local or indigenous knowledge is important in this context. In a general sense, 'expertise' connotes one who is familiar with a generic form and its rules. Even early folklore scholars, who were more often concerned with taxonomies, argued for employing local definitions of genres because they were more meaningful, having come from the individuals who used and created the genres (Ben-Amos, 1976). Swales (1990), from his work on disciplinary communication, suggests that only an expert can best assess the degree to which a communicative event is an instance of a particular genre.

## 2.2. *Mental representations of structure*

Where genre theory provides a conceptual background and understanding in a socially grounded context, cognitive psychology, with its emphasis on individual thought processes, provides a natural extension to genre studies. In particular, the research on mental representations of text structure suggests how a user's mental representation of a text can arise, the varying forms it can take, and how it can affect a user's performance in information processing terms. Three theoretical concepts within cognitive psychology that most aptly shed light on genre concerns are: schema theory, mental models, and strategic discourse processing.

Mandler (1984) proposed one model of schematic processing which is most useful for our concern with information spaces—schemata for stories. She argues that story schemata appear to contain the concept of an episode with one or more nested constituents under an episode, suggesting a hierarchical organization of the mental representation for stories. Mandler proposes that story schemata are temporally ordered, with well-ordered and poorly ordered stories differentially affecting processing.

The mental model approach is perhaps most familiar due to Johnson-Laird (1983) who argued that a mental model “is analogous to the structure of the corresponding state of affairs in the world—as we perceive or conceive it” (p. 156). In textual terms, this ‘model’ is built up by a person hearing or reading a series of words which are then turned into mental representations. Patterns of repetition within texts would give rise to representational forms that are consistent, easing processing and lessening working memory load.

van Dijk and Kintsch's (1983) model of strategic discourse processing possesses components similar to Johnson-Laird's mental models, and it also integrates notions from schema theory. Perhaps most importantly, it posits the existence of mental representations for text genres in the form of ‘superstructures.’ They suggest a strategic reader response in the sense that the reader tries to create effective working hypotheses about the structure and meaning of a text as he or she reads. These hypotheses are then to be confirmed or disconfirmed with further reading. Discourse comprehension begins with the construction of a textbase directly from the discourse. With the successful development of a textbase, the reader is able to establish a situation model in memory. The situation model, much like Johnson-Laird's (1983) mental model, is “the cognitive representation of the events, actions, persons, and in general the situation, a text is about” (p. 11). The textbase is also built up through a set of macrostrategies that build up the global coherence, or gist, of a text. Macrostrategies may create several levels of macropropositions, forming one large macrostructure. van Dijk and Kintsch argue that the presence of macrostructures is directly expressed in a discourse in the form of headings, titles, and initial sentences (e.g. Evans, 1974; Kieras, 1981; Schwartz and

Flammer, 1981), in pronouns, connectives, and word order, and in the general organization of the text.

Finally, van Dijk and Kintsch argue that many discourse types may induce a higher-level mental representation in the form of superstructures. For those discourse types which exhibit a high degree of regularity, a superstructure organizes the macropropositions. A reader employs a schematic processing strategy when encountering such a discourse type with which he or she is familiar. This strategy states that a reader “will try to activate a relevant superstructure from semantic memory as soon as the context or the type of text suggests a first cue” (p. 16). From then on the superstructure may be used to more efficiently process the text in a top-down fashion.

These models of discourse comprehension are clearly related and all assume that in order for text processing to occur, knowledge (and large amounts of it) must be represented and organized in an efficient fashion. Regardless of exact terminology, there is broad agreement across theorists that humans seek order and patterns in trying to make sense of incoming textual information.

## 2.3. *Hypertext—bridging genre theory and mental representations of structure*

Outside of the mainstream theoretical literature, researchers in HCI and related fields have long examined user problems navigating hypertext information spaces (disorientation, digression, and cognitive overhead). Arguments for reducing the complexity of hypertext have often appealed to the use of genre. For example, Gygi (1991) makes an implicit appeal to genre for hypertext systems:

The major problem facing hypertext and hypermedia environments has to do with managing complexity—how not to overwhelm users with vast amounts of information. The authors of conventional documents spend considerable time organizing their presentations to that end. Hypertext documents sacrifice traditional discourse cues, both semantic and physical (p. 284).

‘Conventional documents,’ as suggested by Gygi, are organized in such a way as to reduce the cognitive overload of complex information. This is one essential function of genre, to facilitate communication of information. Gygi's suggestion that hypertext documents lack discourse cues can be interpreted as an appeal to designers to apply genre conventions to hypertext documents. Oren (1991) also makes a direct appeal to genre to facilitate the development of interfaces that reduce complexity for users:

As older, fixed media have found a necessity for genre, so we may suspect that the new flexible computing medium will have a greater need. With the range of design available, there must be a means of setting expectations and transferring knowledge from past exposures, if the content itself is to be appreciated (p. 471).

Implicit and explicit references to spatial cognition have been repeatedly made by researchers to improve users' mental representations of hypertext and analysis of spatial cognition has resulted in different navigation aids for hypertext. One class of these navigation aids are map-like views or overviews of a hypertext, e.g. map views and overview tools (Monk et al., 1988); map windows (Beard and Walker, 1990); map-like interfaces (Simpson, 1990); and hierarchical access tools such as a Home, or root, card, and Overview cards (Leventhal et al., 1993). Another class of navigation aids designed for hypertext might be termed 'linking structures,' e.g. hierarchical and combined network/hierarchical linking structures (Mohageg, 1992); linear and hierarchical structures (McDonald and Stevenson, 1996); and mixed hierarchical and network linking structures (McDonald and Stevenson, 1998). A variety of other design solutions have also been developed to improve the spatial, or navigational, cues for hypertext including: paths or guided tours, backtracking tools, history lists, bookmarks, landmarks, overview maps, fish-eye views, three-dimensional representations, multi-level overviews, and filters (cf. Vora and Helander, 1997).

Explicit appeals to spatial cognition have also indicated the link between hypertext systems and users mental representations of those systems. Edwards and Hardman (1993) argue that users' cognition of physical spaces can be extended to the hypertext domain. In a study of 27 users across three versions of a hypertext system, they found evidence of users creating survey-type, or map-like, representations of simpler hypertext structures. Dillon (1994) and Dillon et al. (1993) also extend this notion of spatial cognition to hypertext. They argue that three levels of mental representations for spatial cognition—landmark, route, and survey knowledge—can be used to understand users' mental representations of hypertext.

Another approach to improving the user's experience with hypertext has been to argue for improving the comprehensibility of hypertext itself. Thüring et al. (1995) and Vora and Helander (1997), drawing from mental representations of structure for reading, suggest that comprehension can be improved by (a) using appropriate metaphors, and (b) by improving the design of hypertext structures so that they are more coherent—such as providing link information between nodes (Vora et al., 1994).

A few studies have attempted to interrelate the benefits of navigation as well as comprehension aids for hypertext. Simpson and McKnight (1990) demonstrated that not only was an overview tool beneficial, but one that best represented the conceptual order of the information produced superior performance, i.e. a hierarchical view compared to an alphabetical view. Tripp and Roby (1990) examined the benefit of navigation and comprehension cues for a set of learning tasks using a hypertext environment. They used two types of cues, advance organizers and visual metaphors. Subjects given the advance organizer performed better than subjects without

the organizer. Dillon (1991) examined navigation and comprehension with expert users' of an information space. Subjects familiar with scientific journal articles were asked to identify the section (Introduction, Methods, Results, or Discussion) to which selected paragraphs belonged, using both print and computer-based media. Subjects were equally, and highly, accurate on each medium, with 80.6% accuracy for print and 82.5% for on-screen display. Participants made use of both surface cues in articles, as well as knowledge of the deep structure of journal articles to complete these tasks.

Dillon (2000), Dillon and Schaap (1996), and Dillon and Vaughan (1997) propose the notion of *shape* to explain the relationship between navigation and comprehension for users:

The concept of shape assumes that an information space of any size has both spatial and semantic characteristics. That is, as well as identifying placements and layout, users directly recognize and respond to content and meaning (Dillon, 2000, p. 523).

Shape combines two elements unique to information spaces: physical movement and semantic intent; not only must users understand the words, but they must move through the words. Shape is argued to be a user's mental representation which captures both factors. Dillon and Vaughan (1997) propose the following notion of shape:

Shape is a property of information that is conveyed both by physical form and information content. Separating these elements completely is perhaps impossible but one can talk of the distinction between the layout and sequencing of information as viewed by the consumer (user or reader) and the cognitive representation of meaning that employs (at least in theoretical terms) knowledge structures such as schemata, mental models and scripts (p. 99).

The concept of shape is a direct driver of the research study described here.

#### 2.4. In summary

In order to understand the general role of form and structure we drew from genre theory and cognitive psychology. First, from genre theory we understand that genre can be considered a class of communicative events, which share a set of conventions and rules to facilitate interaction by creating and maintaining expectations among a community of producers and consumers. At its heart, genre is communicative production and reception that provides a sense of emotional satisfaction. Second, we understand the importance of engaging with experts to understand the conventions and rules of any one genre. In cognitive psychology we focused on mental representations of text structures via the lenses of schema theory, mental models, and strategic discourse processing. We learned that research perspectives provide explanations for both

top-down and bottom-up processing of text, and generally that ordered text differentially affects processing. Patterns of repetition within texts give rise to representational forms that are consistent, easing processing and lessening working memory load. Perhaps most importantly, we can posit the existence of mental representations for text genres in the form of ‘superstructures.’

We chose hypertext to bridge these two theoretical frames with information science. As noted above, researchers and user interface designers have an instinctive sense that genres can facilitate the development of interfaces that reduce complexity for users. Researchers in this space tended to approach the problem as one of comprehension of hypertext or navigation of hypertext, and studied a variety of user interface tools to ameliorate user confusion in information spaces. The concept of ‘shape’ is proposed as an evolution in thinking that combines comprehension and navigation. Shape is argued to be a user’s mental representation that captures both factors.

### 3. Hypotheses

If genre form can drive comprehension, an information space that is designed to facilitate the development of users’ mental representations of that space’s shape should produce demonstrable improvements in user performance. In addition, an information space that violates expectations, although not unusable per se, should produce measurable negative effects. A comparison can then be made of users’ performance over time between two such information spaces.

Two web news sites (genre-conforming and genre-violating) were created based on a study in which web-based news genre conventions were collected from expert web news readers and design guidance was provided by a panel of HCI experts. In the experiment that followed, we allowed users to interact with each web site on a series of information tasks over multiple sessions. A more detailed discussion of the two web sites follows. Our aim was to determine how the genre cues manipulated in the interfaces affected users initially, and over time.

As we were interested in comparing two differently structured web-based newspapers, a genre-conforming design (that met expectations) and a genre-violating design (that violated expectations), we proposed the following hypotheses:

H1. *Structure*. A web-based newspaper designed in accordance with structural genre conventions (i.e. a genre-conforming design) will enhance user performance significantly more so than a web-based newspaper violating these conventions (i.e. a genre-violating design).

‘User performance’ in this hypothesis refers to Comprehension, Usability, and Navigation. Comprehension was measured both at the macrostructural level (i.e. recall and recognition of story headlines) as well as at the superstructural level, (i.e. recall and recognition of news categories). Usability was measured via time on task (for a set of information-seeking tasks), accuracy (for a set of

information-seeking tasks), and satisfaction. Navigation was measured via navigation paths for a reading task, category node visits for a reading task, navigation paths for a set of information-seeking tasks, and category node visits for a set of information-seeking tasks. The specifics of these measures will be explicated in further detail in the methods section of the experiment.

H2. *Time*. Users who are repeatedly exposed to a web-based newspaper will manifest a significant improvement in performance scores over time.

H3. *Structure* × *Time*. Repeated exposure to a web-based newspaper designed in accordance with structural genre conventions (i.e. a genre-conforming design) will significantly improve user performance more so than repeated exposure to a web-based newspaper violating these conventions (i.e. a genre-violating design) over time.

## 4. The web sites

### 4.1. Step 1—eliciting information from experts

#### 4.1.1. Background

In order to create genre-conforming and genre-violating web site designs for news, we first sought information about the design of web-based newspapers, specifically, about the *conventions* which made them web-based news. At this time, web-based newspapers were an emerging medium, and no clear standards had yet developed. We were not so interested in the content of an individual article, as we were interested in how a group of articles were assembled into a web news site. A review of genre theory and cognitive psychology literature suggested that the best place to start for information about genre conventions, was with “experts”—where expertise was defined as both: a high degree of experience participating or interacting with a genre, and experience producing for the genre. For this study, this meant recruiting people both with a high degree of experience reading web-based news, as well as experiencing with web design.

Our second task was to determine how to elicit the information about genre conventions. Based on the notion of shape, we choose to use several lines of attack that would elicit information about comprehension as well as navigation. Specifically we asked the participants about the following issues:

1. Genre-conforming web news categories.
2. Genre-conforming web news page design and navigation.
3. Genre-violating web news categories.
4. Genre-violating web news page design and navigation.

#### 4.1.2. Participants

Six expert readers of web-based news (4 women, 2 men) participated in a group brainstorm session designed to elicit this information. Participants all read web-based

news five days a week or more (where news was defined as general interest news rather than specialized news such as technology) and all had previous experience designing web pages with HTML.

#### 4.1.3. Method

A large open room with a wall-to-wall white board and a bank of Internet connected computers was used to host the group. Participants were asked about each of the issues, one issue at a time. They were given 10–15 min to respond to each question individually in writing. Their responses were then collected and written anonymously on a white board, and discussed as a group. Overlaps in content were pointed out by the facilitator. Each participant was asked for his or her agreement or disagreement with each of the listed responses. If at least half of the participants agreed upon an item, this constituted consensus in favor of including an item in a master list of ‘agreed upon’ elements. If an item garnered fewer than half the votes, this item was explored more fully to ensure that it was equally understood by all members of the group and a second vote was taken. If agreement by at least half of the participants could not be reached, then the item was documented, but not included in the master list. This discussion process took between 15 and 20 min for each issue. Where participants were required to draw or sketch their responses, these were drawn on the white board and discussed in a fashion similar to the written responses. Before beginning each question, an assistant wrote or sketched the ‘agreed upon’ lists or drawings on a flip chart. Each page of the chart was then mounted on a nearby wall so that it could be easily viewed by all participants. Total session running time was approximately 3 h with two breaks. The sessions were videotaped with a tripod-mounted camcorder to record the consensus-building process and as a back-up source for documenting the designs.

#### 4.1.4. Results

Participants were able to respond with ease to the first two issues, genre-conforming web news categories and genre-conforming web news page design and navigation. As will be shown, they had a more difficult time with the genre-violating case. The results of the group brainstorm session are reported below.

*Genre-conforming web news categories.* Participants’ individual responses to the inquiry about ‘genre-conforming web news categories’ produced three broad groupings of categories: information resources, design features, and news categories. Participants defined information resources as, “Things you might go to a newspaper for, or would expect to find there, that are not news.” For example, classifieds, real estate, consumer information, education, and learning were listed as information resources. Because information resources were defined by the participants as ‘not news,’ this set of items was scoped out of the rest of the discussion. Participants also listed many ‘features’ which cut across all of the categories. For example, photos,

archives, keyword search, time stamp on stories, readers’ forum, customization by geographic scope, and customization of categories, were defined as features. Regarding the features, customizing news categories and customizing geographic scope were the only features agreed upon by three or more of the participants.

This left the items which the participants defined as news categories. Participants identified 15 agreed upon general news categories (see Table 1). After the agreed upon news categories were identified, subjects were asked to indicate the categories’ position, or rank, in a linear list (to determine how to order this information in a web site). However, there was little agreement among participants’ rankings. The only ranking all participants agreed upon was placing a category called ‘Headlines/Front Page Stories’ at the top of the list. In an effort to elicit some sense of order, the facilitator asked participants to indicate the *general* location of categories in a linear list, i.e. top, middle, or bottom. Table 1 displays the agreed upon (i.e. by three or more of the participants) news categories and their groupings.

Participants also proposed some alternative orderings for news categories: alphabetical order, a personalized order, and an order modeled after print newspapers. Table 1 indicates participants’ expectations for web-based news categories were very similar to the categories that currently appeared in print newspapers.

*Genre-conforming web news page design and navigation.* The next issue participants responded to was their notion of ‘genre-conforming web news page design and navigation.’ Participants first drew their own responses on paper, these were then transferred to the white board for discussion. Two participants submitted sketches only for a home page; three submitted sketches for a home and category page; and one submitted sketches for a home, category, and story page. In some cases, participants

Table 1  
Participants’ agreed upon categories and groupings for news

Position in list	News categories
Top	Headlines/Front Page Stories World/International National State Region/Local
Middle	Sports Business/Finance/Economy Technology Science/Research Health Editorial/Letters/Opinion/Columns
Middle/bottom	Food and Travel Politics
Bottom	Arts/Entertainment/Diversion Weather

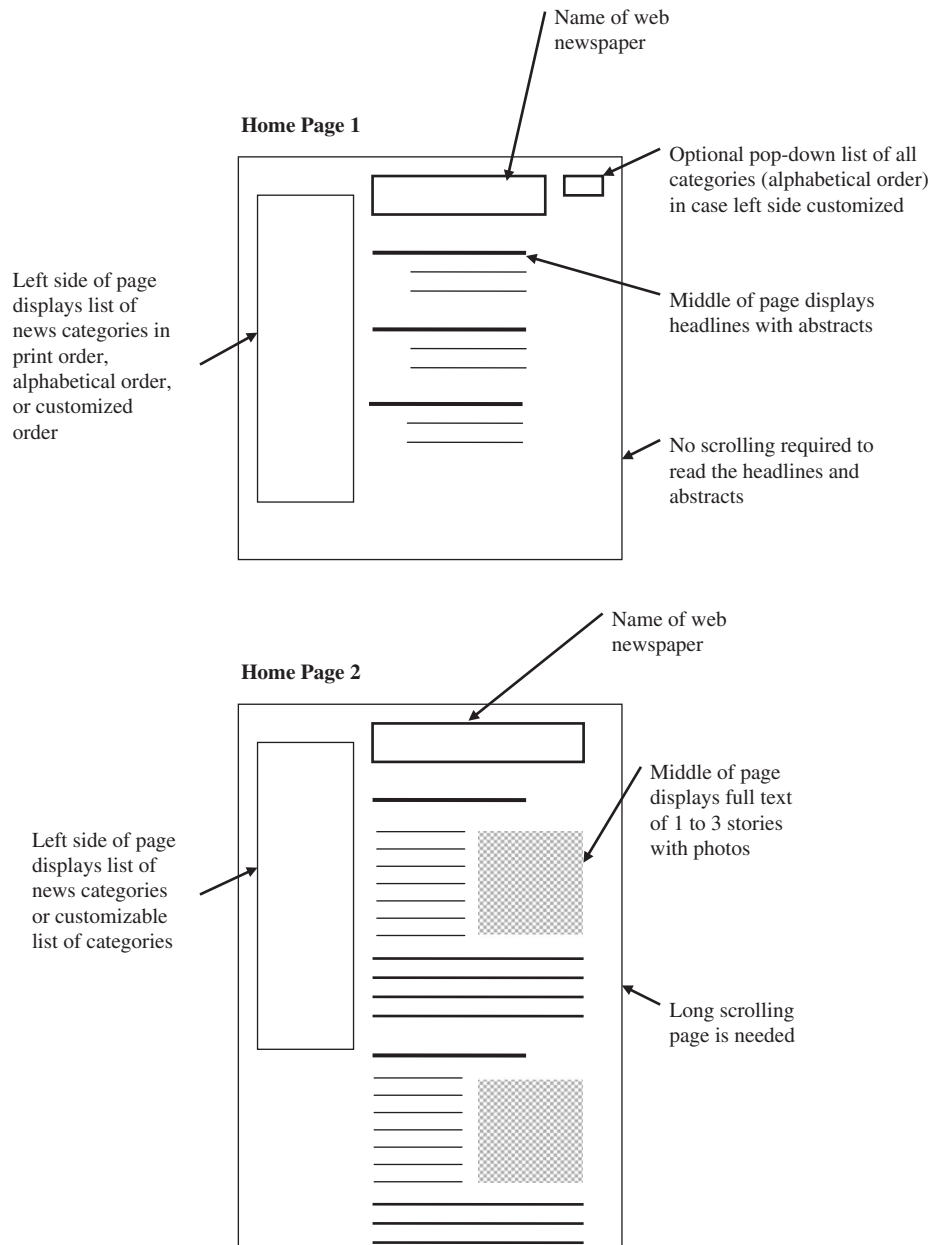


Fig. 1. Participants' agreed upon design components for a genre-conforming news home page.

brought back into the discussion design features mentioned previously. As a group these sketches were reduced to a smaller, distilled set of 'agreed upon' sketches shown in Figs. 1–3.

Fig. 1 indicates the key elements of a web news home page for these participants. The critical components for web page design and navigation were: a left-side navigation bar, a page name, story titles, and story abstracts/text. Participants were equally divided on whether the home page should have a list of headlines and abstracts (Home Page 1), or feature 1–3 full text stories with photos (Home Page 2); therefore both designs were reported.

Fig. 2 indicates there was relatively good agreement about the design of a genre-conforming web news

'category' page. Participants indicated the importance of the following: the name of the category being viewed, left-side navigation, story titles, story abstracts, time and date stamp for each story, a chronological order for the stories, and an indication of stories already read.

In Fig. 3, the agreed upon designs for a 'story' page are shown. Participants were equally divided on the non-use (Story Page 1) or use (Story Page 2) of a left-side list navigation bar, so both options are documented. Between both groups the key commonalities for a story page were: the name of the current category, the story title, and the story text.

*Genre-violating news categories.* The third discussion question attempted to identify a set of conventions for a

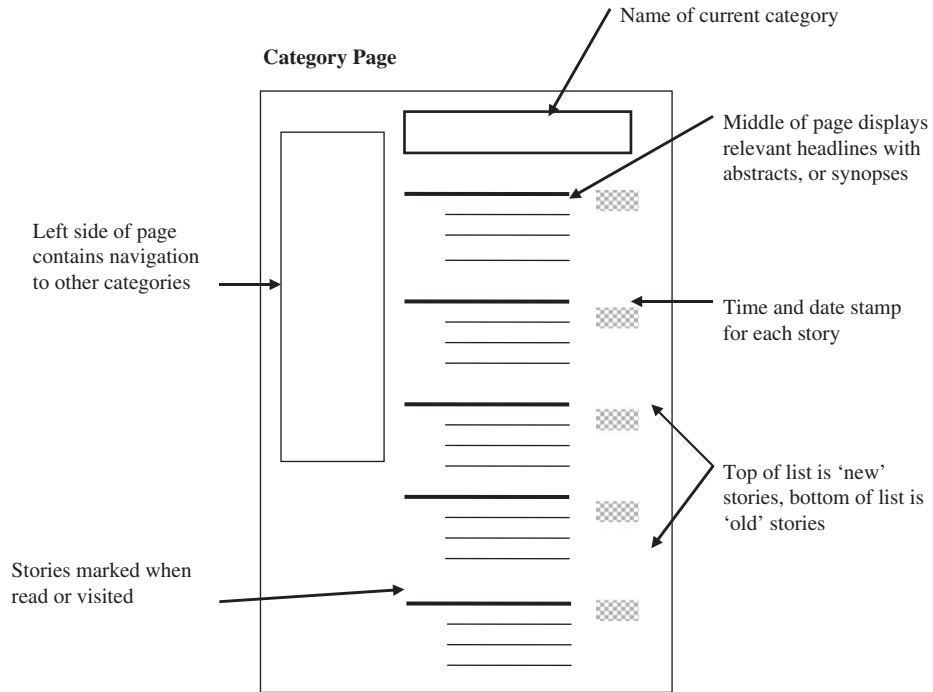


Fig. 2. Participants' agreed upon design components for a genre-conforming news category page.

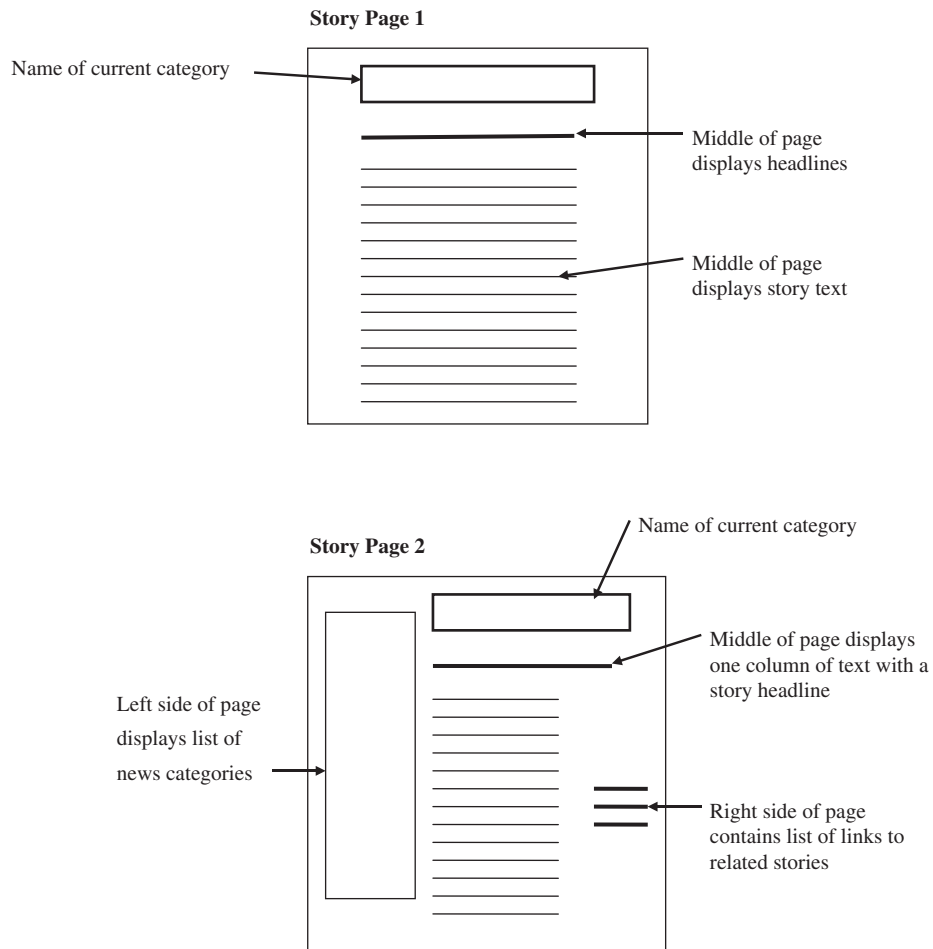


Fig. 3. Participants' agreed upon design components for a genre-conforming news story page.



Table 2  
Participants' agreed upon genre-violating news category characteristics

---

Clever or jingoistic names, e.g. Things, Stuff, Water, Dirt
Tabloid news, e.g. celebrity news, unverified information, subjective news, rumors, graphic or explicit news
Categories that are too low in specificity, e.g. News
Categories that are too high in specificity, e.g. News for Middle Aged Seniors
Unclear category labels by topic or geography, e.g. Important News, Other News
No categories provided at all
Category names that change in scope over time
Category names that are too long
'Feel good' or human interest categories
Large conceptual overlap in category names

---

Table 3  
Participants' brainstormed list of genre-violating news design suggestions

---

No column or width control, i.e. no resizable pages
Providing no user choice or control in general
Elimination of categories altogether
Placing stories in more than one category
Making paragraphs too long
Using blinking or animation such as with ads
Having visual clutter or lack of white space
Having content that is out of date, i.e. very old news
Having links that do not function
Putting stories into the wrong categories
Having no way to get to a Home page
Overuse of the newest technologies
Having pages that take too much time to download
Using pages that scroll right and left
Not using thumbnail photos or interleaved photos

---

genre-violating news web site. Participants were asked to indicate what categories and characteristics of categories they would consider 'bad' for a web newspaper. Their responses indicated that they were more comfortable thinking about the *characteristics* of bad categories, than about bad category names themselves. For this reason, the group discussion diverged from the original topic and instead centered around genre-violating category characteristics. Table 2 displays the agreed upon genre-violating category characteristics.

*Genre-violating web news page design and navigation.* The fourth discussion question probed into what participants would consider genre-violating page design and navigation for web newspapers. After observing how difficult it was for participants to generate responses to the previous question individually (the genre-violating category characteristics), the facilitator decided to re-structure the responses to this question. The facilitator asked the participants to brainstorm as a group, rather than respond individually, about design ideas that would violate their sense of a good news web site (see Table 3). Unfortunately, the session ran short on time and subjects began displaying signs of fatigue. For this reason, the brainstormed list was not reviewed for inter-subject agreement.

The brainstormed list of 'bad' design characteristics certainly produced suggestions for a genre-violating web site. However, all of these suggestions would produce a design that was actually unusable. Given that the goal of this question was to produce suggestions for a news web site that violated design expectations, but was still usable, none of these recommendations were implemented. Instead, a different strategy would have to be used to elicit design suggestions for a genre-violating web news site.

## 4.2. Step 2—prototype the designs

### 4.2.1. Background

The next step was to produce web-based prototypes that contained real world content and that were based on the

suggestions of the expert user group. Web news sites for the genre-conforming and genre-violating designs were prototyped in HTML. An ACCESS database was used to store the news content and Coldfusion was used to serve up three different types of pages populated with content.

### 4.2.2. Genre-conforming web news page design and navigation

For the genre-conforming web news site three types of pages were identified via the group discussion: a home page, a category page, and a story page. Each of these page types and their inter-page navigation was prototyped in HTML (see Figs. 4–6).

Based on the group discussion of a home page (i.e. Fig. 1), a genre-conforming home page design was created (Fig. 4). A left-side navigation bar was created to display the available news categories. A space across the top of the home page was used to indicate the name of the web site, i.e. Web Newspaper. The left-side navigation bar and the page name area were also visually separated with background color blocks. Participants had made two different recommendations regarding the layout of the home page news content. Half the participants suggested that the front page incorporate 1–3 key stories with photographs. The other half argued for a one-column, text only design where a title and abstract were displayed together. The '1–3 stories with photographs' design proved unfeasible using the ACCESS database; the program did not allow for easy inclusion of photographs into the page layout. For this reason, the one-column, text-only format was used. In addition, the pop-down list of categories was not incorporated into this design because all categories would be shown at all times.

Based on the group discussion of a category page (i.e. Fig. 2), a genre-conforming category page design was created (Fig. 5). The category page was given a left-side navigation bar, identical to the home page, and also displayed the current category across the top of the page. This page made use of the same one-column, text only, headline and abstract layout as the home page. The content

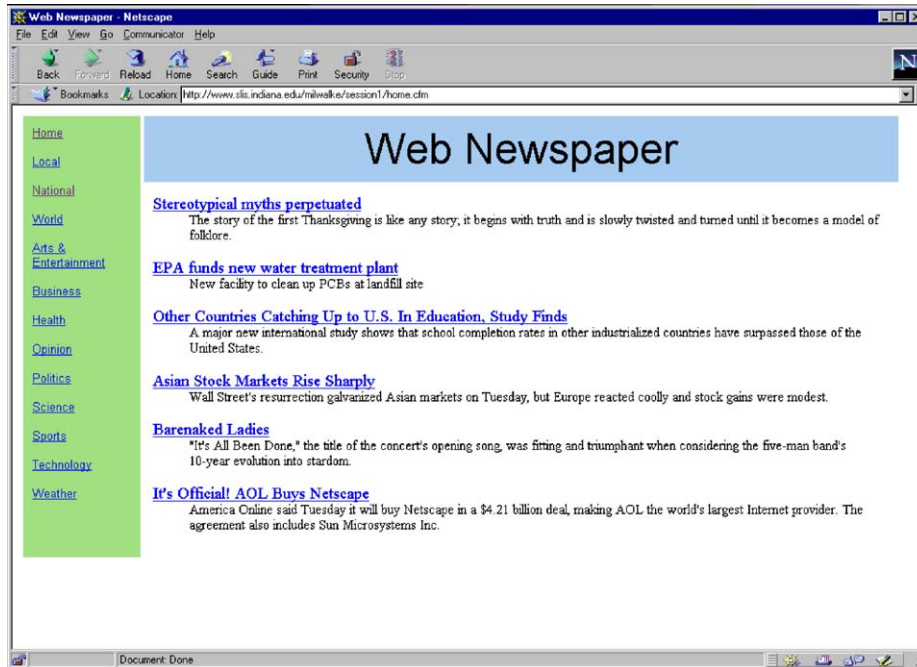


Fig. 4. Home page for genre-conforming web newspaper.

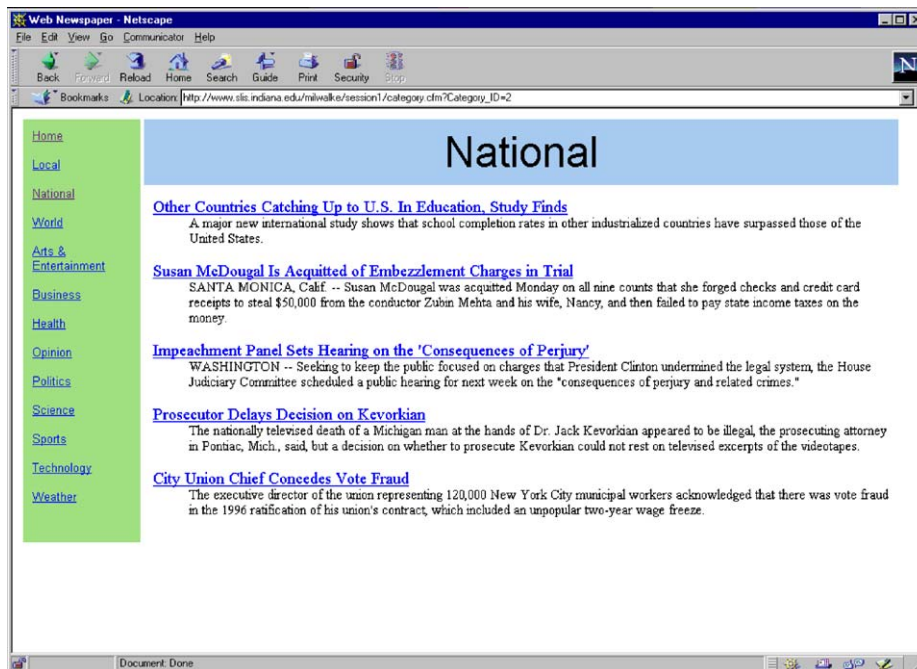


Fig. 5. Category page for genre-conforming web newspaper.

scope would be narrowed to a given category. Stories that had been perused, that is, visited, were marked by a change in link color. No time or date stamps were displayed in this prototype because participants in the experiment would be using the stories one or two days after they were actually pulled from the news wire. This would make the stories appear out of date, and thus 'not news.'

Based on the group discussion of a story page (i.e. Fig. 3), a genre-conforming story page design was created

(Fig. 6). Regarding the story page design, participants were evenly divided on the presence or absence of a left-side navigation bar. Hypertext literature suggests that designs that provide this kind of information facilitate a user's navigation, and so a left-side navigation bar was included at the story page level as well. Consistent with the home page and category pages, this page displayed a banner across the top of the page indicating the current category heading. The story headline and

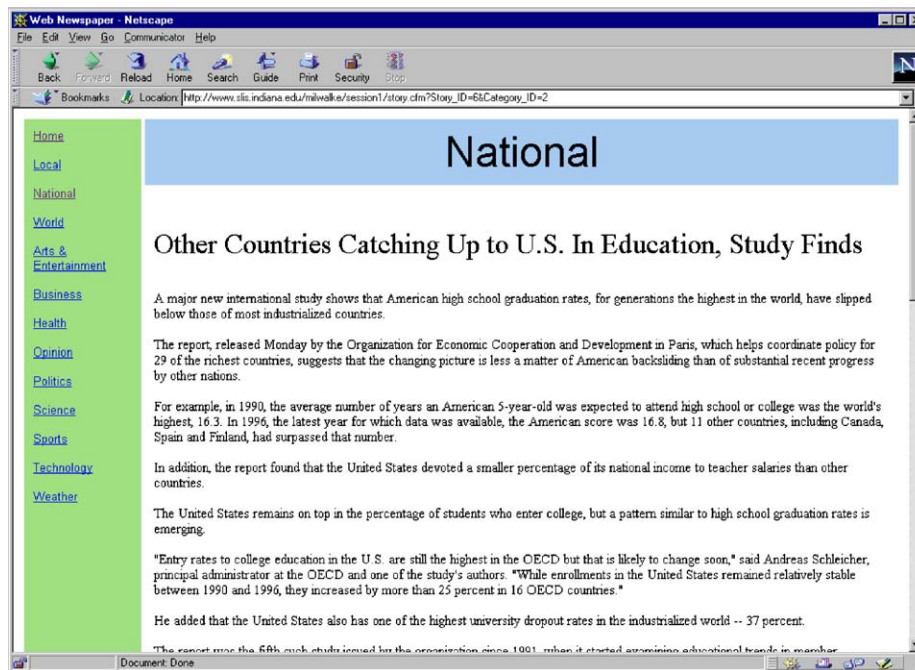


Fig. 6. Story page for genre-conforming web newspaper.

the full text of the story were displayed in the one-column, text only format. Related story links were not implemented at this time due to a lack of a large enough pool of content.

#### 4.2.3. Genre-conforming web news categories and content

The next challenge was to populate this prototype with actual content. News content was collected first from a local on-line college web newspaper (the intended test population was college undergraduates), and then buttressed with content from *CNN* on-line (cnn.com) and the *New York Times* on-line (nytimes.com).<sup>1</sup>

The availability of content was assessed for each of the stated genre-conforming news categories (i.e. Table 1). Headlines was the first category of news and was straightforward to implement. It involved taking stories from other categories and having them also serve as headlines. The model used by the college campus web newspaper was duplicated. The top two ‘campus’ stories and the top three ‘region’ stories were used as headline stories, and then were assigned to an additional news category. Finally, the category name ‘Headlines’ was changed to ‘Home’ in order to reflect existing web navigation models, i.e. NetScape Navigator’s and Internet Explorer’s use of the Home button for a starting point.

The next set of headings were the geographic categories, i.e. World, National, State, and Local. The college web site

offered content only for Local, so World and National content were taken from *The New York Times* on-line. State was ultimately excluded as a category because there was no similar content area on the web news sites chosen for content. Because the design would be tested with participants from the local university, their web newspaper’s order of these categories was borrowed: Local, National, World.

It was difficult to determine a meaningful order for the remaining news categories, therefore they were listed in alphabetical order: Arts and Entertainment, Business, Health, Opinion, Politics, Science, Sports, Technology, Travel, and Weather. The college web site provided content for three of these categories: Arts and Entertainment, Opinion, and Sports. Content was then drawn from *The New York Times* on-line to fill four of the remaining categories: Business, Politics, Science, and Technology. Lastly, content was drawn from *CNN* on-line to fill the final two categories: Health and Weather. Although a Travel category appeared on nytimes.com and cnn.com, a Food and Travel category did not. Due to this conceptual lack of clarity, this category was excluded from the final list.

A final list of twelve unique news categories was produced: Local, National, World, Arts and Entertainment, Business, Health, Opinion, Politics, Science, Sports, Technology, and Weather. Headlines/Home was a redundant category because it borrowed five stories from these other 12 categories. Five news stories were then drawn for each category from the web news sites, creating a database of 60 news stories.

<sup>1</sup>No copyright was violated by using this content; both The New York Times and CNN online allow use of a downloaded copy of their content for personal, non-commercial use.

4.2.4. The genre-violating web news page design and navigation

The next step was to determine the design of the genre-violating web newspaper. As mentioned previously, the suggestions for the genre-violating design provided only unusable design alternatives. In an effort to produce an alternative, genre-violating web news design, a panel of four experts from computer science, cognitive psychology, human–computer interaction, and information science was called together to provide input. The panel suggested manipulating design components based on using the shape framework mentioned previously. They argued for manipulating both the content and form of web news, that is, both spatial and semantic components. Specifically, they suggested modifying the look-and-feel, or form (i.e. navigation bar, use of color, and page layout) as well as modifying the story categorization, or content. Table 4

summarizes the suggestions made for designing the genre-violating web newspaper, contrasted with the genre-conforming design.

A genre-violating design was created using these suggestions. Figs. 7–9 show the implementation of a Home Page, a Category Page, and a Story Page. The alternative story categorization scheme will be discussed later.

In the Home Page design (Fig. 7), the title of the web newspaper is right justified and is displayed with a smaller point size. The left-side navigation has been moved to the right side, and the color blocks denoting regions of the page have been removed. Finally, the one-column title and abstract layout was changed to a two-column layout. The story titles are displayed to the left of the abstracts.

In Fig. 8, the genre-violating category page, a similar layout is used. The title of the category appears at the top right and in a smaller point size. The left-side navigation is

Table 4  
Summary of suggestions for a genre-violating web newspaper design

Design features	Genre-conforming design	Genre-violating design
Navigation bar	Left-side navigation bar	Right-side navigation bar
Background color	Blue to distinguish the 'heading' region Green to distinguish the navigation bar region	No color
Page headers	Approx. 36 point Helvetica; Center justified	Approx. 24 point Helvetica; Right justified
Page layout (Home page and Category page)	One column, with headlines placed directly above abstracts	Two columns, with headlines placed directly to the left of abstracts
Page layout (Story page)	One column, with headlines placed directly above stories	Two columns, with headlines placed directly to the left of stories
Story categorization	Traditional news categories	Time by geography matrix

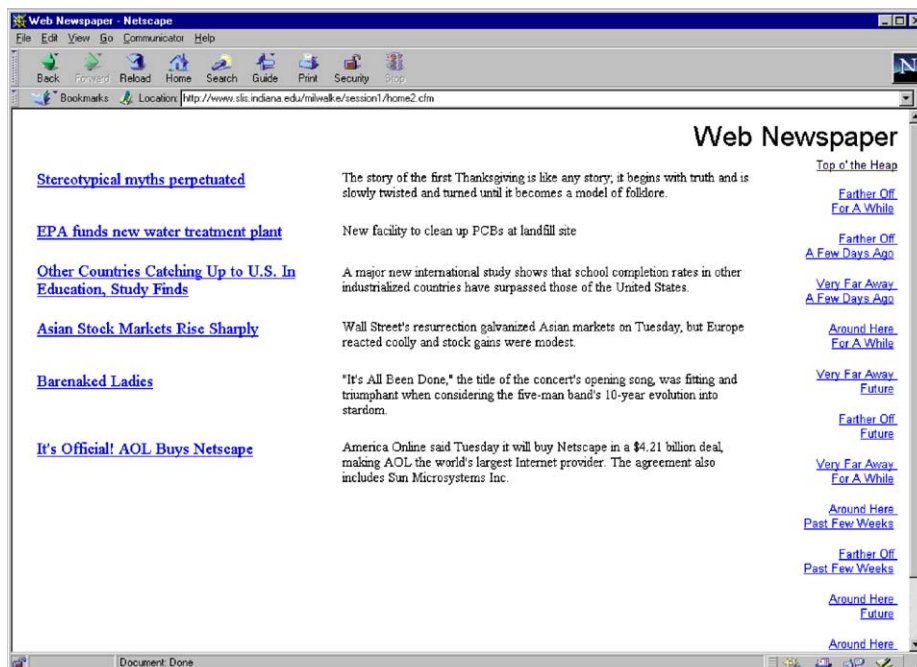


Fig. 7. Home page for genre-violating web newspaper.

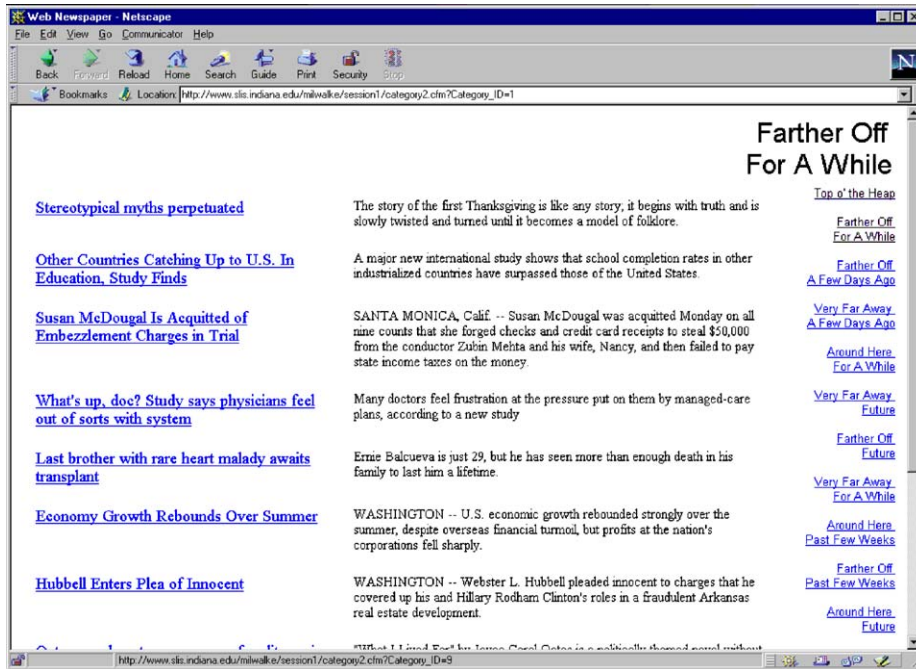


Fig. 8. Category page for genre-violating web newspaper.

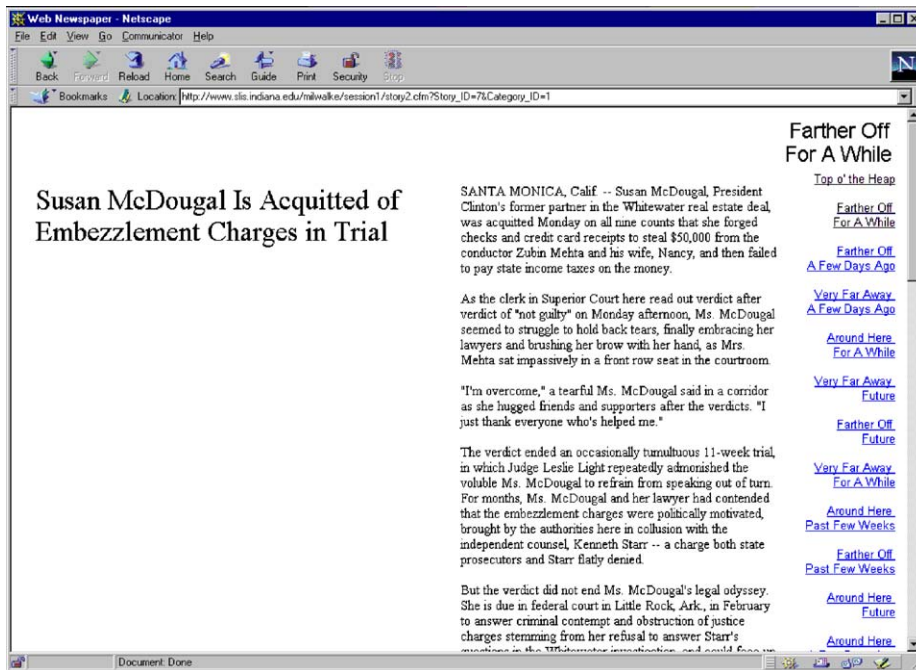


Fig. 9. Story page for genre-violating web newspaper.

again moved to the right, and the same two-column layout is duplicated.

In Fig. 9, a genre-violating story page can be seen. The category title is right justified and in a smaller point size. The story title is to the right of the full text of the story itself. Again, the left-side navigation has been moved to the right.

#### 4.2.5. Genre-violating web news categories

The panel also suggested a method to manipulate the story categorization scheme—moving from general news categories to using time and geography categories. This would allow the use of the same overall textbase, but with a different structure. Using a new categorization scheme would require the user to completely re-think the overall,

Table 5  
Time-by-geography categorization matrix

Geography categories	Time categories			
	A few days ago	Past few weeks	For a while	Future
Around here				
Farther off				
Very far away				

or global, structure of the information space, from a semantic perspective. This re-categorization was accomplished by using a matrix of three Geography categories by four Time categories (see Table 5).

This matrix was presented as a 12-item list in the right-side navigation bar, e.g. item one was labeled Very Far Away—A Few Days Ago. There was no clear means by which to order the matrix vertically. It was decided that the items in the list would be given one initial random order; this order was then maintained across all sessions and across all subjects in the genre-violating group for the experiment.

The same story database was used between both designs. Placement of a story into one of the 12 cells of the matrix was determined by examining the text of the story and categorizing it according to a pre-specified set of definitions. A story was placed into one of the three geographic areas by using the following definitions:

- Around here—events or stories in or about anything within the state.
- Farther off—events or stories in or about anything within the continental United States and outside of the state.
- Very far away—events or stories in or about anything outside the continental United States.

Certain types of stories, such as sports and politics, had ambiguous geographic possibilities. Sports stories could contain information about teams from within the state but whose matches actually occurred outside of the region. The decision was made to classify sports stories as Around Here if they were about a sporting event that occurred within the state, or if they were about local teams. Stories covering politics also carried geographic ambiguities; a story could cover policy made in Washington, DC but regarding a far off country. The decision was made to categorize political stories according to the geographic region to which the policy related.

Each story was also categorized by time. A story was placed into one of the four time categories by using the following definitions:

- A few days ago—events or stories that occurred within the past week.

- Past few weeks—events or stories that occurred or were ongoing for 2–12 weeks (3 months).
- For a while—events or stories that occurred or were ongoing for more than 12 weeks (3 months).
- Future—information or announcements about an event or story that would occur in the future.

Some types of stories, such as science and sports, had ambiguous time possibilities. Science stories were categorized based on an indication of how long the research had been on-going. If there was no indication of the duration of the research, then it was treated as an announcement within the past few days (i.e. A Few Days Ago). Sports stories, if they were almost entirely about the coverage of one particular game, were categorized as occurring within the past week. However, if the story provided extensive commentary about how a particular game fit within an overall season of play, then it was categorized as occurring over the Past Few Weeks or For A While, depending on the duration mentioned.

Each story in the genre-conforming design was thus given an alternate categorization in the genre-violating design. To categorize the stories, a first reading was done using the pre-specified sets of definitions to determine which of the twelve cells best suited a story. After this first pass, a second pass was made in an effort to fill cells in the matrix which were still empty. If a cell was still empty, the news database was re-examined and a new story was selected that fit the final criteria needed for an empty cell. This story then also replaced a story in the genre-conforming database, to ensure the content was consistent. In this way, no cell was left empty. However, because this alternate categorization scheme was used, the number of stories that fell under each category heading for the genre-violating design was not consistent. In contrast, the number of stories under each category in the genre-conforming design was consistent.

It is important to note that we did not attempt to match every variable at every level, i.e. our two categorization schemes used the same textbase, but the number of stories in the alternative categorization varied across categories (see Table 6). Practically, it would have been impossible to create two different versions of the database of stories that were equally matched on structure. More importantly, our emphasis was on creating two qualitatively different experiences—with genre as the motivating factor—and seeing how users reacted both initially and over time.

#### 4.3. Step 4—follow-up

To validate the two designs, the original participants of the brainstorm group were asked to review the genre-conforming and genre-violating web news designs. The participants were emailed the URL of the two designs and asked to comment on the degree of fit between their understanding of the group's consensus and the actual designs. All participants agreed that the designs matched

Table 6  
Number of stories by category in genre-violating website

Genre-violating news categories	Web site version				
	1	2	3	4	5
Farther off—For a while	11	13	11	10	13
Farther off—A few days ago	8	6	14	9	15
Very far away—A few days ago	3	2	2	4	3
Around here—For a while	1	3	3	1	2
Very far away—Future	2	2	1	2	1
Farther off—Future	7	9	6	5	3
Very far away—For a while	3	6	2	3	5
Around here—Past few weeks	6	4	1	2	2
Farther off—Past few weeks	6	6	6	10	4
Around here—Future	6	2	5	8	2
Around here—A few days ago	6	6	8	4	9
Very far away—Past few weeks	1	1	1	2	1

their experience of the group's consensus for genre-conforming web news conventions and for genre-violating web news conventions.

## 5. The experiment

### 5.1. Participants

A total of 25 participants (68% female) completed the study; 13 in the genre-conforming condition and 12 in the genre-violating condition. Overall, these participants read a print newspaper an average of 3.28 days a week, demonstrating a general familiarity with news form and layout. Participants were recruited from two undergraduate information technology classes with web research components, at a large Midwestern university. This helped ensure at least a basic level of familiarity with web browsing. Subjects were screened for prior use of web-based newspapers; only subjects who never made use of a web-based newspaper were admitted.

During their participation in the study; subjects were asked to refrain from keeping up with the news in any manner, print, radio, TV, or on-line form. This was deemed important in reducing possible contamination effects of external news knowledge on the comprehension scores. Contamination was also controlled for by: (a) limiting respondents' recall of macrostructural and superstructural knowledge to headlines, keywords, or parts of headlines seen in the web site, and (b) incorporating false headlines and categories into the recognition test of macrostructural and superstructural knowledge.

### 5.2. Tasks

Two types of tasks were used to assess performance. First, users were asked to perform a general news reading task for 20 min. Subjects were instructed to, "Scan all the news until you feel you have a good, overall sense of the news." Subjects were then asked to perform a set of six

information-seeking tasks. Tasks ranged from easy to complex. Easy tasks required using a keyword from the question to match a keyword in the title of a story; slightly more difficult questions used a keyword in the question to match a keyword in the text of a story; and more complex questions required reading the text of a story and providing an opinion. These questions were modeled after Leventhal et al.'s (1993) tasks involving textual content. The following are examples of each type of question:

- *Keyword-in-title*: Who just arranged a peace accord between Netanyahu and Arafat? ('Netanyahu' and 'Arafat' appear in the title of the story.)
- *Keyword-in-story*: What company recently released new security software? (None of these words appears in the title, but 'security software' appears in one story in the technology section.)
- *Explanatory*: If a person has cardiac apoptosis what are his or her chances for survival? (There was no one answer provided in the text of the story, but rather, several options that the reader had to think through and then offer an opinion about.)

Two questions were created for each level of difficulty, producing a set of six tasks for each session. A new set of information-seeking tasks (see Tasks section below) were generated for each session and presentation of the tasks was randomized.

### 5.3. News content

To better match readers' interest in the news content of this web newspaper, content was drawn primarily from the on-line campus newspaper. The news was collected three days before each wave of subjects was scheduled to begin. Five news stories were drawn from each category, creating a database of 60 news stories. The previous section provides additional details regarding how this content was derived.

### 5.4. Procedures

Participants were randomly assigned to one of the two designs. Each session lasted approximately 1 h, and interactions with the software were videotaped. Participants were first asked to scan all the news items on-line in order to gain a general sense of the news content. They were limited to a 20-min reading session, during which reading task navigation scores were collected. Following reading, participants were given a series of paper-based questionnaires and asked to recall as many headlines, parts of headlines, or keywords from headlines as possible, and as many categories of news as possible. Participants were then given a 5-item Likert-scale satisfaction questionnaire. After completing the questionnaire, they were given a 30-item recognition test. At this point, participants received a 5-min break if needed. After the break, they were given a

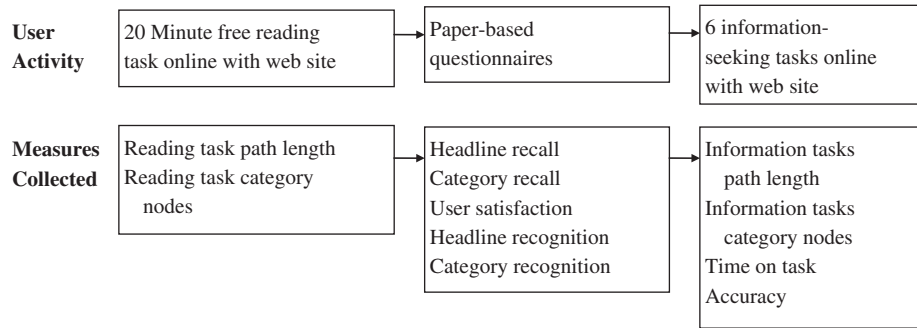


Fig. 10. Flow chart of experimental procedures.

set of 6 information-seeking tasks to complete on-line. Speed, accuracy, and a second set of navigation paths were recorded for the tasks. Fig. 10 provides a flow chart of these procedures along with the measures collected.

Participants returned for four more sessions. At each session, they were given an updated version of the news and again asked to scan it, recall the news headlines and categories, answer the satisfaction questionnaire, complete the recognition test, and then perform a set of information-seeking tasks. Participants were paid a total of \$55.00 for their participation in graduated increments.

### 5.5. Facilities

Participants were run individually through each session using the university's usability laboratory. The laboratory is composed of two rooms, a testing room and an observation room. Participants could be viewed via a one-way mirror, and communication was facilitated by speakers and microphones. The hardware setup included a PC with a 100 MHz processor, 16 MB of RAM, running Windows NT, a keyboard, a mouse, and a 17" monitor. The monitor resolution was set at 600 × 800 to allow proper function with the scan converter. The software was Netscape Communicator 4.0.

### 5.6. Independent variables

**Structure.** As indicated previously, the first independent variable identified was structure. By structure we mean, information space designs that systematically manipulate both navigation and semantic cues. Two web-based newspapers were designed according to expert news readers' conceptions and a set of expert reviewers recommendations for genre-conforming and genre-violating web newspaper designs. One design, the genre-conforming version, incorporated the optimal representations of structure for a web-based newspaper. The second, or genre-violating, version incorporated the violations of users' expectations for what good web news design should be, but was not unusable per se.

**Time.** The second independent variable was time. Participants were measured over five different sessions

with approximately a 3-to-5 day lag between sessions. The lag was necessary to accommodate running all of the subjects before generating each succeeding version of the web newspaper. Each test session involved a new 'edition' of the newspaper, so that contents for each were identical for both interfaces but unique for the test session (1–5).

### 5.7. Dependent variables

Three dependent variables were identified to track changes in user performance between groups and over time.

**Comprehension.** Comprehension of macrostructural knowledge was assessed by asking subjects to write down as many major news headlines, parts of headlines, or keywords from headlines they could recall. Headlines were chosen, rather than the news text itself, for two reasons: (a) to increase the probability of subjects responding strictly to what they were presented with on screen and not to 'general knowledge of news', e.g. news overheard from conversations; and (b) headlines are typically written to capture the gist of a news story and hence make for a fair test of macrostructural knowledge. Superstructural knowledge was assessed by asking subjects to write down as many categories of news as they could recall. Recall for headlines and categories was measured as the total number of correct responses.

Recognition was also used as a measure of comprehension and was assessed with a 30-item questionnaire. Ten true headlines, ten false headlines, five true categories and five false categories were randomly intermixed. Subjects were asked to indicate whether the items did or did not appear in the web site. Recognition for headlines and for categories was measured as the number of correctly identified trues and correctly identified falses. A new recognition test was created for each of the five sessions.

**Usability.** The amount of time needed to complete a task was measured for the six information-seeking tasks in each session. Timing began when subjects finished reading the task out loud and ceased when subjects stated "OK," "Done," or their answer. If subjects worked for 5 min without finding an answer, they were stopped and asked to move on to the next task. This time was deemed more than



sufficient to locate target information in pilot tests. A time-stamped videotape was used to capture task times to the second.

Accuracy was also measured on these tasks. Each task was presented on a 8.5" × 5.5" sheet of paper, after reading the task aloud, subjects completed the task and recorded their answers on the task sheet. Responses were scored with a 1 for a correct answer and a 0 for a wrong answer. The final score represented the total number of correct responses.

Finally, a 5-item Likert-scale questionnaire captured subjects' satisfaction with the web newspaper at the end of each session. On a scale from 1 to 7, where 1 was Strongly Disagree and 7 was Strongly Agree, subjects were asked to indicate their responses to the following questions:

- I liked reading the news using this web newspaper.
- I had fun reading the news on-line using this web newspaper.
- I felt comfortable using this web newspaper to read the news.
- I felt competent reading the news on-line with this web newspaper.
- I think this is a good example of what a web newspaper should look like.

These questions sought to capture the subject's general level of satisfaction with the web newspaper by addressing: the degree to which they liked it, their level of comfort and competence, whether it was a fun experience, whether it reflected their model of a 'web newspaper,' and the extent to which satisfaction shifted over time.

*Navigation.* The navigation paths followed by subjects during the news reading task and during the information-seeking tasks were recorded for each session. The coding captured the sequence in which each page (or node) was accessed. The length of the navigation path was calculated as the total number of pages, or nodes, visited and revisited, including category nodes and the home page. The breadth of the navigation path was calculated as the number of category nodes visited at least once, not to exceed the total number of category nodes available. In order to maintain comparability of navigation paths between tasks, subjects, groups, and sessions, subjects were

asked to begin each task from the Home page of the web site.

### 5.8. Experimental design

The design for this experiment was a between-subjects design for structure (i.e. the genre-conforming versus genre-violating designs) and a within-subjects design for time (i.e. over five sessions). The three dependent variables were assessed via eleven measures during each session, with each group (see Table 7).

The design required a doubly multivariate repeated measures ANOVA with eleven, five-dimensional dependent variables. In addition, a set of univariate trends were run on time.

## 6. Results

### 6.1. The analyses for H1–H3

We first assessed the between group effects with three separate, doubly multivariate analyses on the three dependent variables. Specifically, the first group contained the Comprehension measures, i.e. headline recall, headline recognition, category recall, and category recognition; the second group contained the measures associated with Usability, i.e. speed, accuracy, and satisfaction; and the third group contained the Navigation measures, i.e. reading task path length, reading task category nodes, information tasks path length, and information tasks category nodes.

Across the three analyses there was only one significant interaction effect between time and web news reading condition, Comprehension, while Usability, and Navigation demonstrated no interaction. However, closer inspection of the univariate analyses for Comprehension indicated no significant interactions. For these reasons, the results for each set of hypotheses will be reported in order, H1–H3.

A note about prior newspaper reading experience: there were significant but sporadic correlations between prior newspaper reading experience and the dependent measures at each time. Only 16% of the data points significantly correlated with prior newspaper reading experience. In

Table 7  
Relationship of variables and measures

Variables	Measures
Superstructural comprehension	Recall of news categories, recognition of news categories
Macrostructural comprehension	Recall of news headlines or parts of headlines, recognition of news headlines or parts of headlines
Time on task	Time on task for 6 information-seeking tasks
Accuracy	Correct responses for 6 information-seeking tasks
Satisfaction	5-item Likert-scale questionnaire
Reading navigation	Path length; i.e. total number of nodes accessed; breadth, i.e. number of category nodes accessed
Information-seeking navigation	Path length; i.e. total number of nodes accessed; breadth, i.e. number of category nodes accessed

addition, some of these correlations were not conceptually clear, e.g. prior newspaper reading experience was correlated with category recall at time 1, but not for time 2–5. For these reasons, prior newspaper reading experience was not used as a covariate in the models.

## 6.2. User satisfaction questionnaire

The 5-item user satisfaction questionnaire was tested for internal reliability at each time point. Cronbach's  $\alpha$  on each time was satisfactory (Nunally, 1978): Time 1  $\alpha = .78$ , Time 2  $\alpha = .84$ , Time 3  $\alpha = .82$ , Time 4  $\alpha = .81$ , Time 5  $\alpha = .86$ . The five items were then summed to produce one satisfaction score at each time for each subject.

## 6.3. Hypothesis 1—structure

H1. A web-based newspaper designed in accordance with structural genre conventions (i.e. a genre-conforming design) will enhance user performance significantly more so than a web-based newspaper violating these conventions (i.e. a genre-violating design).

A significant main effect for the between-group difference, i.e. web news reading condition, was found on each of the three sub-groups: Comprehension  $F(4, 20) = 17.07$ ,  $p < .001$ ; Usability  $F(3, 21) = 5.89$ ,  $p < .01$ ; and Navigation  $F(4, 20) = 8.31$ ,  $p < .001$ . The overall pattern of differences on the individual measures (see Table 8) demonstrates superior performance for the genre-conforming condition.

### 6.3.1. Comprehension

On macrostructure comprehension, there were two measures of comprehension, headline recall and headline recognition. For headline recall and headline recognition there were no significant between-groups differences.

On superstructure comprehension, there were two measures of comprehension, news category recall and news category recognition. For category recall there was a significant difference between-groups,  $F(1, 23) = 70.89$ ,  $p < .001$ , with the genre-conforming group recalling more categories. For category recognition there was no significant difference between groups.

### 6.3.2. Usability

On time on task for the set of information-seeking tasks, there was a significant between-groups difference,  $F(1, 23) = 18.61$ ,  $p < .001$ , with the genre-conforming group performing information-seeking tasks more quickly.

On accuracy for a set of information-seeking tasks, there was no significant difference between groups, suggesting that users in both information spaces could locate information accurately.

On satisfaction, there was no significant difference between groups.

### 6.3.3. Navigation

On navigation paths for a reading task, there was a significant difference between groups on path length, with the genre-conforming group having longer paths,  $F(1, 23) = 4.40$ ,  $p < .05$ , suggesting that the genre-conforming space encouraged users to explore further while reading.

On category node visits for a reading task, there was no significant difference between groups at the  $p < .05$  level, however, the two groups were significantly different at the  $p = .069$  level,  $F(1, 23) = 3.62$ . Users of the genre-conforming web site visited more category nodes.

On navigation paths for a set of information-seeking tasks, there was a significant difference between groups, with the genre-conforming group having shorter navigation

Table 8  
Dependent measure means by type of web news reading condition

Dependent measure	Genre-conforming condition	S.E.	Genre-violating condition	S.E.
<i>Comprehension</i>				
Headline recall	10.2	1.3	8.2	1.4
Headline recognition	15.3	0.5	14.2	0.5
Category recall***	10.0	0.6	3.1	0.6
Category recognition	9.2	0.3	8.6	0.3
<i>Usability</i>				
Speed***	6.7	0.6	10.8	0.7
Accuracy	4.9	0.1	4.9	0.1
Satisfaction	25.6	1.6	22.8	1.6
<i>Navigation</i>				
Reading task path length**	39.0	4.2	26.4	4.3
Reading task category node visits*	10.5	0.8	8.2	0.9
Information tasks path length***	22.5	2.8	40.0	2.9
Information tasks category node visits***	12.0	1.7	24.6	1.7

\*\*\*Significant at the .01 level.

\*\*Significant at the .05 level.

\*Marginally significant (i.e.  $\sim .06-.08$ ).

paths,  $F(1, 23) = 18.96, p < .001$ , meaning they had a better sense of where to find information.

On category node visits for a set of information-seeking tasks, there was a significant difference between groups,  $F(1, 23) = 27.31, p < .001$ , with the genre-conforming group visiting fewer category nodes, meaning they went more directly to target information.

6.4. Hypothesis 2—time

H2. Users who are repeatedly exposed to a web-based newspaper will manifest a significant improvement in performance scores over time.

There are multiple answers to this question of the effects over time. The first is the main effect for time, that is, whether the scores increased or decreased over time. The second is the nature of the trend over time, that is, whether the effect over time is a significant linear trend.

There was a significant main effect found for the within-subjects factor of time on one of the three analyses, Usability,  $F(12, 12) = 4.78, p < .01$ , and a marginally significant effect found on the other two analyses, Comprehension,  $F(16, 8) = 3.12, p = .053$ , and Navigation,  $F(16, 8) = 2.623, p = .084$ . The overall pattern of differences on the individual measures (see Figs. 11–14) demonstrates improved performance over time.

*Comprehension.* For headline recall there was a significant effect over time,  $F(2.50, 57.53) = 10.55, p < .001$ , with recall increasing linearly,  $F(1, 23) = 12.93, p < .01$ . For headline recognition there was a significant effect over time,  $F(4, 92) = 7.65, p < .001$ , with recognition increasing over time somewhat linearly,  $F(1, 23) = 3.68, p = .068$ . For news category recall there was a significant effect over time,  $F(2.71, 62.39) = 20.37, p < .001$ , with recall increasing

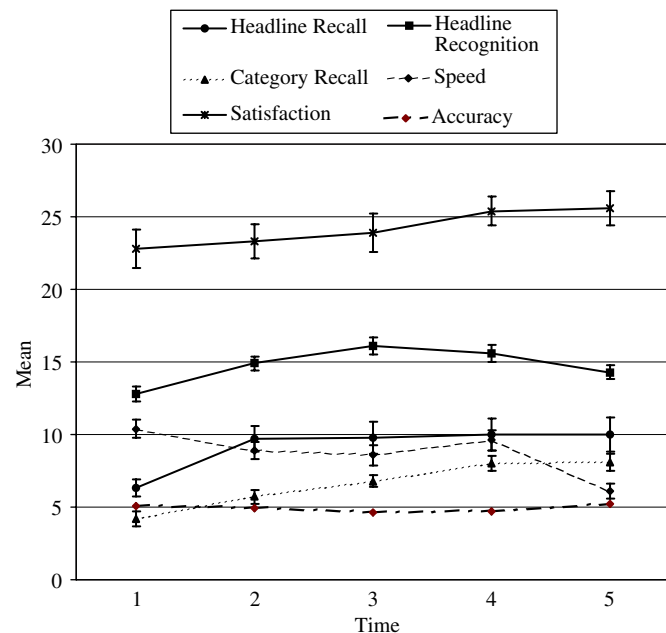


Fig. 11. Significant dependent measures over time.

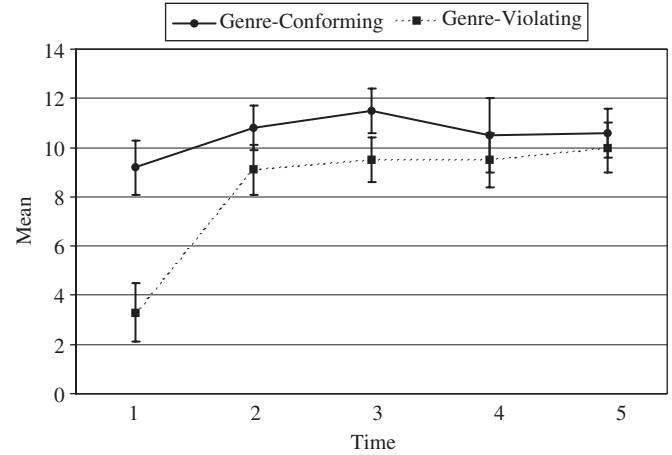


Fig. 12. Web sites by reading task category nodes over time.

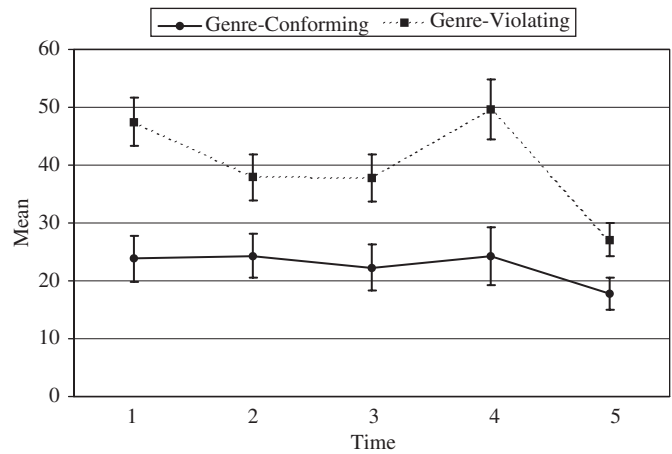


Fig. 13. Web sites by information task path length over time.

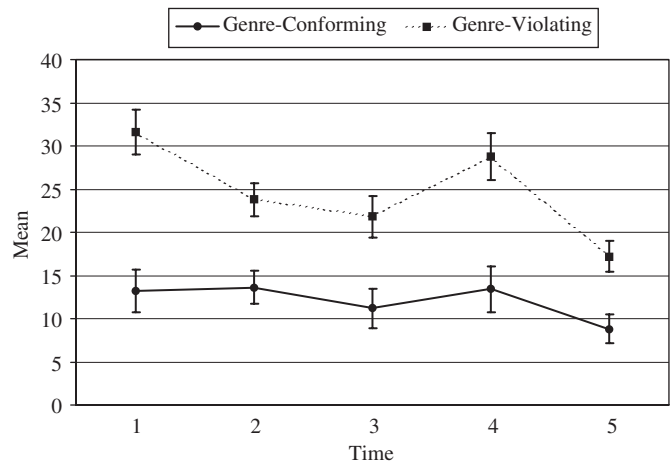


Fig. 14. Web sites by information task category nodes over time.

linearly  $F(1, 23) = 45.49, p < .001$ . For news category recognition there was no significant effect over time.

*Usability.* There was a significant effect for speed over time,  $F(4, 92) = 11.98, p < .001$ , with speed, or time on task, decreasing linearly,  $F(1, 23) = 25.61, p < .001$ . There was no significant effect for accuracy at the  $p < .05$  level.

However, accuracy was significant at the  $p = .057$  level,  $F(4, 92) = 2.38$ , increasing over time. There was a significant effect for satisfaction scores over time,  $F(4, 92) = 6.75$ ,  $p < .001$ , with satisfaction increasing linearly,  $F(1, 23) = 15.77$ ,  $p < .001$ .

*Navigation.* There was no significant effect over time for the path length on the reading task. There was a significant effect for the number of visits to category nodes on the reading task over time (see Fig. 12),  $F(1.91, 43.85) = 14.19$ ,  $p < .001$ , with the number of visits increasing linearly,  $F(1, 23) = 16.05$ ,  $p < .001$ . There was a significant effect in the length of navigation path for the information-seeking tasks over time (see Fig. 13),  $F(4, 92) = 6.21$ ,  $p < .001$ , with path length decreasing linearly,  $F(1, 23) = 10.91$ ,  $p < .01$ . There was a significant effect in the number of visits to category nodes for the information-seeking tasks over time (see Fig. 14),  $F(3.14, 72.29) = 9.38$ ,  $p < .001$ , with the number of visits decreasing linearly,  $F(1, 23) = 18.84$ ,  $p < .001$ .

### 6.5. Hypothesis 3—structure $\times$ time

H3. Repeated exposure to a web-based newspaper designed in accordance with structural genre conventions (i.e. a ‘genre-conforming’ design) will significantly improve user performance more so than repeated exposure to a web-based newspaper violating these conventions (i.e. a genre-violating design), over time.

Across the three analyses there was only one significant interaction effect between time (the within-subjects effect) and web news reading condition (the between-subjects effect). Comprehension demonstrated the only significant interaction effect,  $F(16, 8) = 3.24$ ,  $p < .05$ , while Usability,  $F(12, 12) = .89$ ,  $p = .578$ , and Navigation,  $F(16, 8) = .66$ ,  $p = .772$ , did not. However, none of the univariate analyses for Comprehension or Usability displayed significant interaction effects. Within Navigation, two measures demonstrated a significant interaction effect.

There was a significant interaction effect between time and group type on the number of category nodes visited for the reading task,  $F(1.91, 43.85) = 5.21$ ,  $p < .01$ . However, the interaction was opposite to that predicted; the number of nodes visited increased over time more dramatically for the genre-violating group than for the genre-conforming group.

The number of category nodes visited for the information-seeking tasks demonstrated a significant interaction between time and group type,  $F(3.14, 72.29) = 2.89$ ,  $p < .05$ . As with the above measure, the biggest change over time occurred with the genre-violating group.

## 7. Discussion

### 7.1. Overview

Both genre and cognitive psychology theoretic perspectives were applied to the problem of structuring users’

conceptions of an information space. Data were initially collected from expert users of web-based news to support the creation of two web news sites—one that violated and one that conformed to users’ expectations. From a user performance perspective, the navigation data provide interesting insight to the impact of these designs. Users who worked with a system that met their expectations (i.e. an interface that provided a ‘good’ news structure), appeared to have a performance advantage over those who worked with a system that violated their expectations (i.e. an interface that provided a non-traditional news structure). Specifically, users of the genre-conforming design viewed more of the web site during the reading tasks, and performed more efficiently on the information-seeking tasks.

However, it is also clear that users were able to learn a system that possessed a wholly new structure. Users of the genre-violating system clearly struggled with the unfamiliar system initially, however, they were able to acquire adequate knowledge of the site’s structure over time. Over repeated exposures, users of the genre-violating system increased their exploration of the web site during the reading task, and increased their searching efficiency during the information-seeking tasks.

One caveat we should address up front is the variability of the amount of content under each category in the genre-violating condition (Table 6). While the total number of stories remained stable over web site versions (i.e. 60) in the genre-violating condition, there is not a fixed number of stories for each category per version. However, there are relatively equivalent numbers per category, per version of website. For example, “Farther Off—For A While” has more stories overall than other categories, but the number of stories per web site is around  $11 \pm 2$ . To better demonstrate this, we have included a table of the categories ranked by number of stories per version (see Table 9). The concern is that the degree of variability in the number of stories per category will differentially affect user’s performance, thus confounding potential differences due to

Table 9  
Rank order of categories by number of stories in genre-violating websites

Genre-violating news categories rank	Web site version				
	1	2	3	4	5
Farther off—For a while	12	12	11	12	11
Farther off—A few days ago	11	10	12	10	12
Very far away—A few days ago	5	4	5	7	7
Around here—For a while	2	5	6	1	5
Very far away—Future	3	4	3	4	2
Farther off—Future	10	11	9	8	7
Very far away—For a while	5	10	5	5	9
Around here—Past few weeks	9	6	3	4	5
Farther off—Past few weeks	9	10	9	12	8
Around here—Future	9	4	7	9	5
Around here—A few days ago	9	10	10	7	10
Very far away—Past few weeks	2	1	3	4	2

genre. We would argue that the degree of stability provides some scope for rapid strategic learning, and hence is evidence for the importance of genre. At the same time, we recognize the need for future research on this topic and would recommend “stable story structure” as a variable to consider.

A second caveat we should comment on is the issue that the genre-violating condition presented its news categories in an initial randomized order that was then maintained over the duration of the five sessions. Initial performance measures of participants in the genre violating condition may have been artificially reduced by this added complexity. However, the authors are comfortable with the conclusion that participants were able to rapidly learn the structure over time, based on the improvements in their scores over time.

Taken together, these findings highlight the importance of understanding users’ existing genre knowledge when designing digital information spaces as well as the importance of testing user responses over time. Leveraging knowledge about users’ conceptions of information spaces can clearly benefit non-expert users, even when applying this knowledge to a new medium, to create useful and usable information spaces for immediate use. For regular use, it is clear that initial user responses are not fully indicative of responses after several sessions.

## 7.2. *Mental representations of structure, time, and web newspapers*

### 7.2.1. *Comprehension*

Comprehension was assessed at the macrostructural and superstructural levels. Macrostructural comprehension reflects a reader’s sense of the news in a given instance, and superstructural comprehension reflects a reader’s overall sense of the structure of news. For both types of comprehension, two measures were used—recognition and recall. Recognition memory is generally considered a more sensitive measure of memory than recall (Zechmeister and Nyberg, 1982).

Between the two groups, subjects were equally able to recognize the gist of the news (macrostructure) and the structure of the news (superstructure). Subjects in the genre-conforming group, however, demonstrated significantly superior ability to freely recall the structure of news, as well as (although not statistically significant) a greater recall of the gist of the news. The significant difference in superstructural recall suggests that subjects in the genre-conforming group learned the structure of ‘news’ and were able to retrieve it better than subjects in the genre-violating group.

One possible explanation for the difference between subjects’ ability to recognize rather than recall the gist and structure of the news lies in the sensitivity of the measures. Recognition is a more sensitive measure of memory; it is a cognitively less demanding task than recall because it does not require the user to actively search his or her memory

for information. The user is presented with a cue that is an exact match for something that may be encoded, but not necessarily stored in memory (e.g. Shepard, 1967; Zechmeister et al., 1978). In contrast, recall is a measure of what has actually been passed into long term storage by the user and requires the user to search and retrieve information in storage with no extra cues provided (e.g. Tulving and Osler, 1968; Watkins and Tulving, 1975). Hence, the amount of cognitive work required for recall is greater than that required for recognition.

The greater task demands of recall thus allowed for better discrimination of between group differences. Overall performance on superstructural recognition was 91% for the genre-conforming condition and 86% for the genre-violating condition. On macrostructural recognition, performance was 77% for the genre-conforming condition and 71% for the genre-violating condition. Overall, performance on superstructural recall was 17% for the genre-conforming condition and 5% for the genre-violating condition, and for macrostructural recall was 17% for the genre-conforming condition and 14% for the genre-violating condition.

The effects of time were apparent on both macrostructural and superstructural comprehension. Macrostructural knowledge improved over time on both recognition and recall, and superstructural knowledge improved over time on recall. Subjects were increasingly able to recognize the content of the news, recall the content of the news, and recall the structure of the news, over time. In fact, all three measures demonstrated significant linear trends. One possible explanation for the effect of time on recognition and recall for the gist of the news (macrostructure), is that subjects were increasingly able to create mental representations of the news content itself; i.e. they were becoming more efficient at building up macrostructural representations of the news due to repeated exposure and use. As argued by van Dijk and Kintsch (1983), the consistencies across one’s macrostructural representations can facilitate the development of superstructural representations—in this case of news. The significant improvement over time in subjects’ recall of superstructural information suggests this was the case. The lack of difference over time on superstructural recognition can potentially be explained by two factors, sensitivity of the measure or insufficient time. Recognition may not have registered much of an improvement over time due to its high sensitivity, or subjects may not have had enough exposures to the web sites for a representation strong enough to influence superstructural recognition to develop.

Across the five sessions there were no significant interaction effects between time and structure on comprehension. Subjects’ comprehension in the genre-conforming group did not benefit significantly more over time than subjects’ comprehension in the genre-violating group; both groups improved approximately equally over time. This was surprising because the hypothesized interaction was that the genre-conforming group would benefit

significantly more over time. One possible explanation for this lack of interaction effect is that both web sites possessed a baseline level of structure; a group of structural components was held stable between the designs to increase their comparability (i.e. the use of category headings, the presence of an index, and the three types of pages). Perhaps the presence of these structural elements, even in the genre-violating form, was enough to assist users in effectively developing mental representations of structure. If the performance of the genre-violating group was facilitated by this level of structure, enabling them to perform at a level closer to the genre-conforming group than expected, the opportunity for a significant interaction effect was reduced.

### 7.2.2. Usability

Usability was assessed via three measures: speed, accuracy, and satisfaction. Speed was examined over time by looking at the time required for users to complete a set of information-seeking tasks. Accuracy was examined by looking for the correct number of users' responses on a given set of information-seeking tasks. Satisfaction was examined via a questionnaire that assessed users' general level of comfort, ease of use, and competence with the web newspaper.

The between-group differences on usability were limited to one measure, speed. Subjects working with the genre-conforming web design were able to perform the set of information-seeking tasks significantly faster than subjects working with the genre-violating design. Subjects' ability to perform significantly faster on the genre-conforming design may be explained by the presence of a mental representation for news. If, as proposed in the discussion of Comprehension, subjects in the genre-conforming group were able to develop a superstructural representation for news, their knowledge of the structure would allow them to more accurately predict where answers might be found in the web site and to more quickly retrieve the answers.

In contrast to the differences in speed, both groups submitted equally accurate answers to the information-seeking tasks and were equally satisfied with the designs. The lack of difference between groups on accuracy was not wholly unexpected; other comparisons of accuracy have also failed to detect between-group differences (e.g. Monk et al., 1988; Dillon, 1991). This may have occurred because the tasks were not complex or challenging enough to capture significant between-group differences. The lack of difference between groups on satisfaction, however, was surprising given the differences in the interface and the significant impacts on speed and recall. Yet, this is not the first time that satisfaction data has failed to parallel behavioral data in the evaluation of software systems (cf. Nielsen, 1993; Vora et al., 1994).

The effects of time on usability were much more apparent. Speed, accuracy, and satisfaction all significantly increased over time. Users were able to perform more quickly, with more accuracy, and with greater satisfaction

over time. Similar to comprehension, these scores demonstrated significant linear trends, suggesting possible improvements even after Time 5. As argued previously, improvements in macrostructural and superstructural comprehension over time point to the development of mental representations of structure. Extending this argument to usability suggests that by accessing these increasingly developed mental representations, users were able to improve their actual use of the web sites. Speed and accuracy increased and thus users were more satisfied with their experience.

The lack of an interaction effect is visible in the means of the usability measures, particularly for speed and satisfaction. As has been suggested previously, this lack of an interaction between time and structure may well be due to the presence of some structure in both interfaces. Even though the genre-violating design presented information in a non-conventional structure, users were still able to discern some structure and could, in turn, build on this information to develop their mental representations.

### 7.2.3. Navigation

Navigation was measured on two tasks: a reading task and a set of information-seeking tasks, using two different measures: path length and path breadth. It was argued that the genre-conforming design would encourage greater exploration of the web site during the reading task, resulting in more efficient navigation during the information-seeking tasks.

The between-group differences on these measures of actual system use are apparent. The genre-conforming web design appears to have enhanced users' exploration of the web site during the reading task, as well as performance during the information-seeking tasks. One possible explanation for this is that the genre-conforming design facilitated the development of users' mental representations of structure (both macrostructural and superstructural). Specifically, the presentation of information in the genre-conforming design encouraged more active exploration of the web site. Increased active exploration helped users develop more accurate macrostructural and superstructural representations; and more accurate representations made users' search strategies more efficient during the information-seeking tasks.

Navigation seemed to be the more sensitive to the between-group differences than Comprehension and Usability. One explanation may be that navigation is an excellent process measure for hypertext; it is sensitive to real-time use of the system in a way that comprehension and usability are not. Comprehension and usability do not capture process so much as output from interacting with the system (Dillon, 1994).

The effect of time on navigation is also readily apparent. Over time, users increased their navigational exploration (i.e. increased path length and increased number of category nodes visited) in the reading tasks and improved their navigational efficiency (i.e. decreased path length and

decreased number of category nodes visited) in the information-seeking tasks. In addition, these effects demonstrated a significant linear trend, suggesting potential increases in exploration and efficiency even after Time 5. These improvements in actual use of the system—more exploration in the reading tasks and more efficient searching in the information-seeking tasks—could easily reflect the development of mental representations of structure. Repeated exposure helped users develop these representations, which in turn helped improve their actual use of the system.

Unlike comprehension and usability, navigation was impacted by the interaction of structure and time. The interaction effect, however, was not in the predicted direction. It was hypothesized that users in the genre-conforming condition would benefit most by exposure over time. Instead, users in the genre-violating design made the most gains over time in terms of increasing navigational exploration and efficiency. One explanation for this difference between groups may be that the opportunities for gains over time were not equivalent; the genre-conforming design was sufficiently usable from the outset that the possibility for improvement was less than was available in the other condition.

### 7.3. Implications

#### 7.3.1. For on-line news and information spaces

The impacts of genre-conformance and repeated exposure on user comprehension, usability, and navigation were positive. Overall, these results indicate that both structure and time make a difference in terms of user performance with web-based information spaces. When considering the design of on-line news in particular, and information spaces in general, these data have two compelling implications.

First, building an information space based on the structural genre conventions and the expectations of users will provide a performance and comprehension advantage for users in both the short and medium terms. For a web designer, this means being aware of the audiences' prior knowledge and experiences with a given genre, and finding ways to incorporate that structure into the web site design. Second, and perhaps more importantly, providing a consistent structure over time will allow users to build mental representations of an information space—even if the structure violates prior conventions and expectations. The participants in this study demonstrated an ability to improve their performance over time with both the genre-conforming and genre-violating interfaces. For web designers, this suggests the importance of providing a stable structure on web sites intended to draw repeat users.

An additional interesting aspect of time was that while users of the genre-violating interface were able to improve their performance, they were unable to equal the performance of users in the genre-conforming group. When a designer develops a 'new' web genre, he or she must expect

performance losses for users who come to the site for the first time, and should perhaps target users who are willing to do the cognitive work of learning a new genre.

The reader is reminded that generic conventions evolve to serve a community of readers and producers. This suggests that the generic structural conventions of on-line news will evolve as well. Thus, attention should be paid to emerging news conventions, and they should be incorporated into web site design. Idiosyncratic or completely original designs come with a performance cost.

#### 7.3.2. For hypertext and disorientation

This study is one of the few examinations of repeated use of alternate hypermedia designs. The effects of the genre-conforming web news design on comprehension, usability, and navigation were significant. The superior comprehension scores suggest an improved comprehensibility of the hypertext structure for the genre-conforming design. It has already been argued that this improved comprehensibility led to the superior navigational exploration scores on the reading tasks, the superior speed scores on the information-seeking tasks, and the more efficient navigation scores on the information-seeking tasks. These improvements in performance suggest that, indeed, providing the user with both semantic (i.e. comprehension) and navigational (i.e. spatial) aids reduces user confusion and disorientation with a hypertext. Such findings offer support for the construct of information shape and the spatial-semantic model of information space posited in Dillon (2000).

#### 7.3.3. For cognitive psychology

This study extended van Dijk and Kintsch's (1983) notion of macrostructures and superstructures to *collections* of news stories. van Dijk and Kintsch developed their theory based on individual news stories, not on the concept of a collection of stories and they suggest that a reader employs a set of strategic processes to make sense of a news story. Macrostructures arise from features in the text, such as headings, as well as from the overall meaning or gist of the text and ultimately help a reader organize a text. van Dijk and Kintsch suggest that stories which exhibit a high degree of regularity (i.e. a user perceives regularity between multiple macrostructures) will help a reader build a mental representation known as a superstructure. A superstructural representation organizes a person's macropropositions about a story more efficiently because it allows a reader to exploit knowledge of the discourse type (a.k.a. a genre).

For this study, van Dijk and Kintsch's notions of macrostructure and superstructure were extended to a collection of news stories. This study proposed that in reading a collection of stories a reader would develop a macrostructure for the news presented in a given session, i.e. the news for a given day. It further proposed that repeated exposure to a collection of stories would help the reader develop a superstructural representation for the collection, e.g. a mental representation for web news. The gain in users' performance over time on recognition for

both macrostructural and superstructural information suggests that subjects were at least encoding and temporarily maintaining both macrostructural and superstructural information. The increase in recall scores over time for macrostructural and superstructural information suggests that subjects were able to store and retrieve macrostructural and superstructural knowledge. In addition, the genre-conforming group's superior recall scores on superstructure demonstrates the significant performance advantage afforded to readers who are working with a familiar discourse type. These data are congruent with previous work that suggests it is possible for a user to develop a mental representation, for a story's discourse type, such as a folktale (Mandler, 1984), a story (Johnson-Laird, 1983), a news story (van Dijk and Kintsch, 1983), or an academic article (Dillon and Schaap, 1996), but also for the organization of a collection of stories, such as an academic journal (Dillon et al., 1989).

#### 7.3.4. *For genre theory*

From a genre theoretic perspective, these data provide experimental support for the importance of generic conventions and the role of time. The working definition of genre stated that: genre is a class of communicative events which shares a set of conventions and rules and facilitates interaction by creating and maintaining expectations in producers, listeners, and readers. The features identified by the focus group reflected the shared set of conventions and rules for producers and readers of news. The superior performance of the genre-conforming group demonstrates that the creation and maintenance of expectations has a basis in one's cognitive processes, particularly in facilitating the development of mental representations of structure. The speed and navigation scores indicate that these representations do indeed provide for more efficient communication. Finally, the effect of time on the genre-violating group demonstrates that a mental representation for a particular generic form may take some time to develop, but that it can indeed develop.

#### 7.3.5. *For usability evaluation*

For usability evaluations, the present data suggest that exposing users to new designs for single trials or sessions may fail to reveal important dynamics in the user response to technology that occur with repeated exposure. Since most users will be expected to repeat interactions with a given design (except for certain cases), the usability of designs cannot be easily determined by watching first-time use, which is more a measure of learnability than usability. These data support the view that users can adapt to an interface that behaves consistently, even if the design is less than optimal in many other respects. However, even after 5 sessions, the performance of users in the genre-violating space was not as good as users in the better design. Future research could usefully extend the trial period to determine if and when performance differences between the designs disappear. Such work would help establish a firmer basis

for cost-benefit analysis in usability work and help us better understand how long-term are the effects of various design attributes.

#### Acknowledgements

This research was funded in part by the Special Libraries Association Goldspiel Award to Andrew Dillon, and Ph.D. support funds from the U.S. Department of Education Higher Education Act program for Misha Vaughan. This paper is based on the first author's doctoral dissertation work and was conducted at the Usability Laboratory in SLIS at Indiana University, Bloomington.

#### References

- Bakhtin, M.M., 1986. In: Holquist, M., Emerson, C. (Eds.), *Speech Genres and Other Late Essays* (Trans. by Vern W. McGee). University of Texas Press, Austin, TX.
- Bauman, R. (Ed.), 1992. *Folklore, Cultural Performances, and Popular Entertainments*. Oxford University Press, New York.
- Beard, D.V., Walker II, J.Q., 1990. Navigational techniques to improve the display of large two-dimensional spaces. *Behaviour & Information Technology* 9, 451–466.
- Ben-Amos, D., 1976. *Folklore Genres*. University of Texas Press, Austin, TX.
- Berger, A.A., 1992. *Popular Culture Genres: Theories and Texts*, vol. 2. Sage, Newbury Park, CA.
- Berkenkotter, C., Huckin, T.N., 1995. *Genre Knowledge in Disciplinary Communication: Cognition/Culture/Power*. Lawrence Erlbaum, Hillsdale, NJ.
- Cawelti, J.G., 1976. *Adventure, Mystery, and Romance: Formula Stories as Art and Popular Culture*. The University of Chicago Press, Chicago.
- Dillon, A., 1991. Readers' models of text structures: the case of academic articles. *International Journal of Man-Machine Studies* 35, 913–925.
- Dillon, A., 1994. Information as a structured space. *Designing Usable Electronic Text: Ergonomic Aspects of Human Information Usage*. Taylor & Francis, Bristol, PA, pp. 105–120.
- Dillon, A., 2000. Spatial-semantics: how users derive shape from information space. *Journal of the American Society of Information Science* 51, 521–528.
- Dillon, A., Gushrowski, B., 2000. Genres and the web: is the home page the first digital genre? *Journal of the American Society for Information Science* 51, 202–205.
- Dillon, A., Schaap, D., 1996. Expertise and the perception of shape in information. *Journal of the American Society for Information Science* 47, 786–788.
- Dillon, A., Vaughan, M., 1997. 'It's the journey and the destination': shape and the emergent property of genre in evaluating digital documents. *The New Review of Hypermedia and Multimedia* 3, 91–105.
- Dillon, A., Richardson, J., McKnight, C., 1989. The human factors of journal usage and the design of electronic text. *Interacting with Computers* 1, 183–189.
- Dillon, A., McKnight, C., Richardson, J., 1993. Space—the final chapter or why physical representations are not semantic intentions. In: McKnight, C., Dillon, A., Richardson, J. (Eds.), *Hypertext: A Psychological Perspective*. Ellis Horwood, New York, pp. 169–191.
- Edwards, D.M., Hardman, L., 1993. 'Lost in hyperspace': cognitive mapping and navigation in a hypertext environment. In: McAleese, R. (Ed.), *Hypertext: Theory into Practice*. Intellect, Oxford, UK, pp. 90–105.
- Evans, H. (Ed.), 1974. *Handling Newspaper Text*. Heineman, London.



- Gygi, K., 1991. Recognizing the symptoms of hypertext...and what to do about it. In: Laurel, B. (Ed.), *The Art of Human-Computer Interface Design*. Addison-Wesley, New York, pp. 279–287.
- Hasan, R., 1996. In: Cloran, C., Butt, D., Williams, G. (Eds.), *Ways of Saying: Ways of Meaning: Selected Papers of Ruqaiya Hasan*. Cassell, New York.
- Johnson-Laird, P.N., 1983. *Mental Models*. Cambridge University Press, New York.
- Kieras, D.E., 1981. Topicalization effects in cued recall of technical prose. *Memory and Cognition* 9, 541–549.
- Leventhal, L.M., Teasley, B.M., Instone, K., Schertler, D.S., Farhat, J., 1993. Sleuthing in HyperHolmes™: an evaluation of using hypertext vs. a book to answer questions. *Behaviour & Information Technology* 12, 149–164.
- Mandler, J.M., 1984. *Stories, Scripts and Scenes: Aspects of Schema Theory*. Lawrence Erlbaum, Hillsdale, NJ.
- McDonald, S., Stevenson, R.J., 1996. Disorientation in hypertext: the effects of three text structures on navigation performance. *Applied Ergonomics* 27, 61–68.
- McDonald, S., Stevenson, R.J., 1998. Effects of text structures and prior knowledge of the learner on navigation in hypertext. *Human Factors* 40, 18–27.
- Mohageg, M.F., 1992. The influence of hypertext linking structures on the efficiency of information retrieval. *Human Factors* 34, 351–367.
- Monk, A.F., Walsh, P., Dix, A.J., 1988. A comparison of hypertext, scrolling, and folding as mechanisms for program browsing. In: *Proceedings of the HCI '88 Conference on People and Computers IV*, pp. 421–435.
- Nielsen, J., 1993. *Usability Engineering*. AP Professional, New York.
- Nunally, J.C., 1978. *Psychometric Theory*, second ed. McGraw-Hill, New York.
- Oren, T., 1991. Designing a new medium. In: Laurel, B. (Ed.), *The Art of Human-Computer Interface Design*. Addison-Wesley, New York, pp. 467–479.
- Schwartz, M.N.K., Flammer, A., 1981. Text structure and title-effects on comprehension and recall. *Journal of Verbal Learning and Verbal Behavior* 21, 61–66.
- Shepard, R.N., 1967. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior* 6, 156–163.
- Simpson, A., 1990. The interface to a hypertext journal. In: *Human-Computer Interaction INTERACT '90 Proceedings of the IFIP Third International Conference on Human-Computer Interaction*. Elsevier, Cambridge, UK, pp. 869–874.
- Simpson, A., McKnight, C., 1990. Navigation in hypertext: structural cues and mental maps. In: McAleese, R., Green, C. (Eds.), *Hypertext: State of the Art*. Intellect, Oxford, UK, pp. 73–83.
- Swales, J.M., 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press, New York.
- Thüring, M., Hannemann, J., Haake, J.M., 1995. Hypermedia and cognition: designing for comprehension. *Communications of the ACM* 38 (8), 57–66.
- Tonkin, E., 1992. *Narrating Our Pasts: The Social Construction of Oral History*. Cambridge University Press, Cambridge, UK.
- Tripp, S.D., Roby, W., 1990. Orientation and disorientation in a hypertext lexicon. *Journal of Computer-Based Instruction* 17, 120–124.
- Tulving, E., Osler, S., 1968. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology* 77, 593–601.
- van Dijk, T.A., Kintsch, W., 1983. *Strategies of Discourse Comprehension*. Academic Press, New York.
- Vora, P.R., Helander, M.G., 1997. Hypertext and its implications for the Internet. In: Helander, M., Landauer, T.K., Prabhu, P. (Eds.), *Handbook of Human-Computer Interaction*, 2nd, completely revised ed. Elsevier Science, New York, pp. 877–914.
- Vora, P.R., Helander, M.G., Shalin, V.L., 1994. Evaluating the influence of interface styles and multiple access paths in hypertext. In: Adelson, B., Dumais, S., Olson, J. (Eds.), *Human Factors in Computing Systems CHI '94 Conference Proceedings*. Association of Computing Machinery, New York, pp. 323–329.
- Watkins, M.J., Tulving, E., 1975. Episodic memory: when recognition fails. *Journal of Experimental Psychology: General* 104, 5–29.
- Zechmeister, E.B., Nyberg, S.E., 1982. *Human Memory, An Introduction to Research and Theory*. Brooks/Cole Publishers, Monterey, CA.
- Zechmeister, E.B., Curt, C., Sebastian, J.A., 1978. Errors in a recognition memory task are a U-shaped function of word frequency. *Bulletin of the Psychonomic Society* 11, 371–373.