

FINAL DRAFT

Perceiving and using Genre by Form— An Eye-tracking Study

MALCOLM CLARK^a, IAN RUTHVEN^b, PATRIK O'BRIAN HOLT^a

^a IDEAS Institute, The Robert Gordon University, Aberdeen, Scotland

^b The Department of Computer and Information Sciences, University of Strathclyde, Glasgow, Scotland

This paper reports on our approach to the analysis of genre recognition using eye-tracking. We focused on eight different types of e-mail, such as calls for papers, newsletters and spam, which were chosen to represent different genres. Our study involved the collection of oculographic behaviour data metrics, such as fixations and saccades to highlight the ways in which people view the features of genres. We found that genre analysis based on purpose and form (layout features, etc) was an effective means of identifying the characteristics of these e-mails.

Our research, carried out on a group of 24 participants, highlighted their interaction with the e-mail texts and the visual cues or features perceived and also the strategies they employed for the processing of the texts. The results showed that readers can determine the purpose and form of genres, that occasionally form and content can be separable, that some features cause fixations and that some readers are prompted to respond by using saccadic behaviour (e.g. regressive saccades) over the shape of the e-mails (form).

Introduction

This research focuses on overlaps of information retrieval (IR), cognitive science and genre studies, merging and utilizing these to help understand how structured e-mail texts are perceived and classified. We are interested in how the context of a community of practice gives rise to standardized information forms produced by members of the community, and how these forms can be exploited by IR technology to improve retrieval effectiveness.

The large international IR communities, which include a variety of international academic organizations, commercial companies (Google, Yahoo et al.), the text retrieval evaluation conference (TREC), and, the initiative for the evaluation of XML retrieval (INEX), have recently started to understand the importance of (technologically) structured text.

However, until now, two important approaches have been largely overlooked: the naturally occurring structures called genres and the human perceptual are used to identify and employ them. Search engines typically only take into consideration genres that are based on

topical categories, but there is a huge potential for genre-sensitive classification and retrieval tools based on techniques used by members of a community to perceive and utilize genre in their work environments.

This research is based on an examination of the visual layout of genres and, in particular, how they are used within an e-mail community. Within most e-mail accounts, there tend to be socially-constructed communicative behaviours, that is, genres, which emerge to improve the efficiency of the activities in a “community of practice” (Wenger 1999). In our work we use a variety of e-mail genres to investigate how such genres are recognised, and we use eye-tracking techniques to identify salient features of e-mail genres that lead to their classification.

Based on an earlier pilot study, (Clark, Ruthven & Holt 2008), we investigated the ways in which the participants interacted with and used the genres of e-mails, the features or attributes they perceived, and whether their perceptive processes could be modelled and understood. An important aim was to find out

whether the genres could be said to have been 'recognised' from a constructivist (Gregory 1966) viewpoint or whether the texts 'afforded' (ecological (Gibson, J 1986; Gibson, JJ 1986)) the ways in which they could be used, that is, their purposes.

Eye-tracking equipment facilitates the capture of the eye movement behaviour which is involved at variable levels of cognition, for example, in semantic and oculomotor processes (Poole & Ball 2006). We were able to assess the direction in which the participants were looking at any given time, the ways in which they looked and the sequences involved. Our aim was to discover how a user gazes at the stimuli presented, and to identify, in particular, the fixation patterns and saccadic data, that are aligned with perceptual processes. Information processing is said to be suppressed during saccades, and fixations are linked with intensive cognitive processes that could provide clues to the type of perceptual organisation involved.

The *Background* section outlines previous work in this area, including a brief review of recent work on eye-tracking in IR and information search (IS), with an explanation and description of the types of *ocular measurements*.

Next, we present our *eye-tracking study*, including our *research questions*, the *methods* utilized, the references of the *corpus* used, the *variables*, *task* and *procedure*, the *types of data* recorded and the *participants*, concluding with the details and a *summary of the results*. The final section gives the *conclusions* derived from this research.

Background

The principal aim behind this work was to test, empirically, the ways in which people may exploit genres by examining any perceptual strategies and useful features that are used to identify the *genre* of e-mails.

Genre is often treated as the classification of movies or books such as, for example, "westerns", "short stories" or "detective novels". Although, of course, this definition has some relevance, the term "genre" embodies a much wider range of contexts. For the purpose of this experimental eye-tracking study, genre was defined by two principal components: its communicative purpose and form (or layout)

(Yates, J. A. & Orlikowski 2002, p. 15). We exclude other features such as style although we will consider these in future work.

Yates and Orlikowski (1992), in their pioneering work on the concept of genre, suggested that: "Genres (for example, the memo, the proposal, and the meeting) are typified communicative actions characterized by similar substance and form and taken in response to recurrent situations", which can be used to identify types of organizational communication. The form is the set of structures and layout that show the user the document's form through its structure, such as lists and headings, regardless of the topical nature of the writing. The communicative purpose represents many attributes such as arguments and discourse structure.

The experiment described in this paper examines how the textual formatting features of e-mails categorize the purposes for which the text has been written and identify the form elements that are common to various e-mail genres. The methodology we used was to examine whether these categories of purpose and form were actually *perceivable* and measurable.

There are two prominent visual perception processes through which human beings are thought to perceive: the *Ecological* (including the theory of Affordances) and *Constructivist* processes.

The constructivists defend a top-down approach according to which perception begins with recognition, that is, the animal uses sensory information, and builds or constructs this incomplete information to make sense of it (Braisby & Gellatly 2005). Toms and Campbell, for example, argue that perception is a top-down process where the readers recognize the genres through the attributes of the layout which forms the basis of document recognition, and, although Toms and Campbell, like Lakoff (1987), refer to the bottom-up process and suggest that genres may "act as a single gestalt" (Toms & Campbell 1999, p. 2015). However, they fail to explore other possibilities, such as how a genre is perceived when the document is displayed to a reader. In their conclusions, however, Toms and Campbell (1999, p. 2015) query how the form of the document affects a user in the first few seconds and this begs the question: how does the structure of a genre aid

in text comprehension and use? These are two of the questions which will form a central part of this research.

Ecologists (Gibson and others) argue for an alternative and direct (bottom-up) framework for perception and Gibson not only challenged the stages but also introduced the notion of 'affordance' as a centre piece to his theories (Gibson, J 1986). Gibson's theory of affordances is intended to define how meaning and perception are related: he argues that instead of perceiving objects (such as texts) and then adding meaning later, there are visual combinations of invariant properties of objects which provide cues on how to act in relation to these objects.

Watt(2009), for example, argues that genre is present for an intentional purpose, that is, to give guidance to readers about what to do with the text by assisting the action-making process, and that genre should thus be employed more widely in IR. He argues that perception and genre work together. For example, the headline of a news story is intended to grasp the reader's attention or as Michaels(2007) says: "attention is guided by genre information", and the abstract of an academic article allows a filtering decision to be made on whether the article is relevant or not; thus the reduction in the reader's "cognitive load" allows the reader to decide that he need not read a whole document because the genre provides the invariant cues to its relevance in its structure. So in essence the text structure has also "afforded" its purpose and form.

To summarize, the ecological school believes that the goal of perception is to perceive in order to act, in this context, it could be the act of directing the attention of the reader to the salient properties of the text, and the constructivists assert that the final goal is that we perceive for recognition which would need a level of cognitive processing. Of course, locating evidence of the use of these perceptual processes is no easy task but this research project aims to provide an insight to form the basis for further research: this study attempts to show how people use structured texts in relation to genre and perception and will offer some useful empirical data and indicators as to whether capturing these types of interaction with texts is feasible.

Eye-tracking

Eye movement work began in the very early 20th century. Before eye-tracking equipment was developed, several methodologies were implemented which on the surface may sound somewhat barbaric and highly invasive: for example, the electro-oculographic method required electrodes to be mounted on the skin around the eye of a participant so that the experimenters could record and measure any changes in 'electric' potential to detect eye movements. Eye-tracking¹ has also been used extensively in many fields of research, some examples being interface development (1999), understanding how users read webpages (Nielsen 2006), understanding sign language (Muir & Richardson 2005), understanding how users view images (Underwood, Humphrey & Foulsham 2008), image search & retrieval (Buscher, Dengel & Elst 2008), and IR, in particular, relevance feedback (Joachims et al. 2005).

The main metrics in eye-tracking generally come under the category of fixations, saccades, gaze data and scanpath measurements. In this paper we concentrate on fixations and saccades.

Fixations are moments when the eyes are relatively stationary, cognitively trying to decode the information, whilst saccades are rapid eye-movements which occur between fixations (Rayner 1998).

For both fixations and saccades there are several measurements that can be chosen in studies, (Poole & Ball 2006; Rayner 1998):

1. Number of total fixations – more fixations may mean less skimming and more scanning for certain features or information. More cognitive processing.
2. Fixation duration – more duration means attention getting information or lack of understanding of certain information.
3. Number of saccades – may indicate more skimming with cognitive action suppressed.
4. Regressive saccades - may indicate less meaningful cues (in this case text formatting, keywords etc) and more scanning over areas already searched.

¹ Rayner (1998) gives a fuller description of methods used over the last 100 years.

Eye-tracking Study

Our study aimed to understand the relative contribution of form and communicative purpose in understanding how people identify genres. In our case the genres are genres of e-mail texts. In this section we outline our experimental methodology, discussing the *eye-tracking system*, the *corpus* used, the *participants* and the *types of data* recorded and the *methodology* employed.

Our study was designed to answer the following research questions:

1. What are the relative contributions of form (structure) and purpose (content) to identifying genre?
2. What oculographic methods do humans use when viewing and utilizing the invariant layout cues, such as white-space patterns or other formatting features that constitute genres? Are such features fixated upon and/or merely viewed with saccadic behaviour such as regressions?
3. Do/can participants 'skim' the shape of e-mail texts by possibly using saccadic behaviour?

Apparatus

We used the Arrington PC-60 Viewpoint Eye-tracker which is a desk-mounted infra-red camera system that allows the experimenter to choose one of three methods: pupil only, corneal reflection only, or both together at the same time. For the purpose of this experiment, the latter seemed to be most suitable. The Viewpoint software computes pupil height and width to better than 0.03 mm instantaneously with no averaging and has blink detection and suppression. The software recorded eye data: X, Y position of gaze, pupil height and width, ocular torsion, delta time, total time, and regions of interest (ROI) in which fixations, gaze times and saccades for each stimuli can be computed and recorded in an ASCII file.

The monitor viewed by the participant was recorded by the Freeware application Wink²

which records using Flash. This allowed playback of the session for each participant.

The experiment was run using one high specification dual core PC running Windows XP with two monitors connected to the same PC. The 'desktop' was split into two spaces, one for the participant's stimuli and the rest for recording. The stimuli (figure 2) were shown randomly on a 15-inch monitor of 1024 x 760 pixel resolution whilst recording was performed on a monitor of the same type and resolution.

Corpus

The e-mails collected for this task came from two domains: e-mails from sources internal to the institution of the first author and external e-mails. Although many e-mail types are genres in their own right which have evolved from memos, all e-mails today contain sub-genres with their own individual purpose and form such as, calls for papers, newsletters, orders and spam. Along with external e-mails (Table 2) there are also internal e-mails (Table 1) relative to particular organizations.

We defined these e-mail types by a prior investigation of what e-mails commonly occurred in e-mail accounts of several colleagues in our institution so the selected e-mail types are ones that may be familiar to our experimental participants. None of these colleagues participated in the study, and the experimental participants did not contribute any e-mails to the study. The typical e-mails that were found to emerge, especially in this academic environment, were personal library account updates, e-mails from information technology services regarding downtimes in the institution, calls for papers and the other types detailed in Tables 1 and 2.

These e-mails are normally composed of several layers or sections, organized in a certain form (observable features such as use of uppercase, centring, salutations) that defines the type of genre and a distinct communicative purpose. We were interested in whether these attributes of e-mail genres are perceived by readers (are they observed?) and used by readers (do they help in classifying e-mails?). The study tested eight genres of e-mails and the invariant features which were used to make decisions during the genre identification.

²Available at: <http://www.debugmode.com/wink/>

Table 1: Internal e-mails

Type	E-mail Purpose	E-mail Form
ITS Outage(ITS)	Announces downtimes of servers and systems	Structural features: title uppercase, emboldened text items listing outage information.
Seminar(Sem)	Similar to call for paper but internal announcement of invited talk.	Structural features: uppercase titles centered, block of text about speaker, abstract, and block of text about organizer.
Library(Lib)	Message from library; reminder that a book is ready for collection/ return.	Structural features: block of centered text recipient details in uppercase. Opening salutation. Block of text (two paragraphs) terms and conditions, list of renewal item(s) being referred to

Table 2: external e-mails

Type	E-mail Purpose	E-mail Form
Call for papers(Cfp)	Calls for submissions for conferences and workshops by announcing the requirements and important dates.	Structural features: large title, block of centred text (sometimes uppercased). Block of text explaining the event. Bullet lists explaining scope of subjects for conference. Important dates – titles and dates in list format
Cinema(Cin)	Announces cinema listings, dates and times.	Structural features: uppercased cinema name/title

		rectangular block of text with name of film, rating, length, times per day of show.
Spam(Spm)	Scam letters with the motive of deceiving people to send money for a false cause.	Structural features: spam uses 'letter' variation format. Top lines indicate type of spam i.e. Nigerian letter, Lottery scam etc
Newsletter(NL)	Summarizes all the weekly news from an organization, i.e. Aberdeen Football club.	Structural features: lists of items emboldened. Opening salutation to the recipient. Emboldened title with small summary paragraph and URL below each for the e-mail. URL at end to un-subscribe.
Orders(Ord)	Confirmation from a business of an order for item(s) online i.e. Next, Tesco etc	Structural features: Order number and 'thank you for the order details'. Table created with format using lines consisting of symbols (- * /) with details of the order: quantity, item ordered unit cost and at very bottom total cost. Delivery address uppercased and date order being delivered.

The e-mails were transformed into images, a process which was necessary for the eye-tracker as the system only allowed stimuli in bitmap format. In total 48 images of each genre of e-

mail were collected and used in the study. A 'pool' of very similarly sized e-mails of each genre (6 of each) were collected. The e-mails were converted into landscape format with font size 14 points. Each e-mail was transformed into four representations making 192 images in total.

Participants

Participants were 24 unpaid volunteers. All 24 participants were between 20-48 years old. 12 were aged between 20-29, 6 between 30-38 years old and 6 between 42-48 years old; 18 were male.

20 participants came from academic backgrounds: 5 lecturers, 1 senior lecturer, 4 undergrad students, one masters student and 9, PhD students. The students came from a variety of academic backgrounds. 4 non-academics also took part: one clerical assistant, one personal assistant and two IT system administrators.

Participants came from a variety of nationalities with the majority from the UK who had English as a first language. All participants were fluent in written and oral English. The mix of nationalities gives us another direction for analysis in the future regarding the influence of participants having English as a first language or not. Regarding e-mails, circa 5 had had almost no experience with e-mail; 1 were not very experienced; 8 were not experienced or experienced; 8 were quite experienced; 2 were very experienced.

Participants were asked about their prior familiarity with the eight types of genres used for this study. The results were as follows (table 3). For nearly all categories the majority, at least 70% of participants, were familiar with the category (score of at least 3). The exceptions were Cinema and seminar announcements which only pertained to some participants who signed up for these announcements and e-mails from the library which were only relevant to those participants who had previously sent a message to the library. However, there were no significant differences between the familiarity levels for the categories using a Wilcoxon test for statistical significance.

Table 3: e-mail familiarity amongst participants

Type	Familiarity (1=completely unfamiliar / 5=completely familiar). N.B. percentages rounded down.				
	1	2	3	4	5
CFP	5%	13%	18%	30%	34%
Cin	25%	16%	29%	16%	12%
Spm	0%	0%	4%	29%	66%
NI	8%	16%	16%	37%	20%
Ord	4%	8%	12%	28%	45%
ITS	4%	4%	8%	32%	48%
Sem	12%	24%	12%	28%	20%
Lib	16%	12%	20%	32%	16%

Research questions

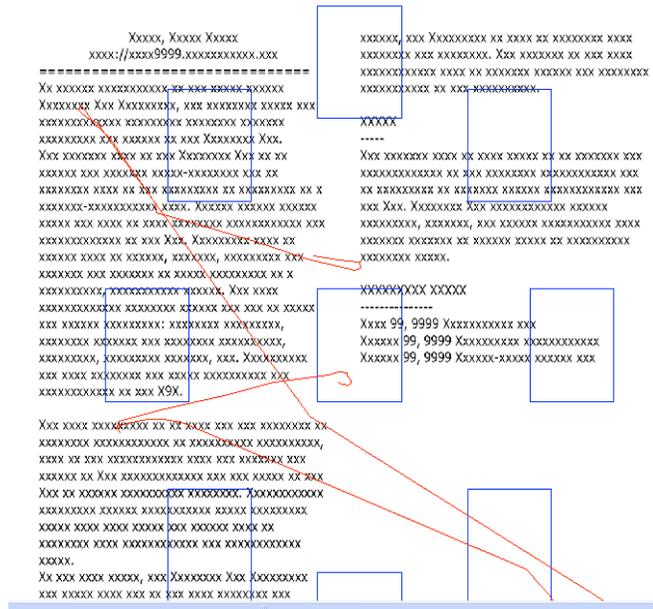
We were interested in the relative roles of purpose versus form in identifying e-mail genres. To test these we followed the same data formatting approach previously used in Toms, Campbell & Blades (1999) and Watt's (2009), later e-mail work. In this, form is the structural formatting of the e-mails whereas purpose is seen as the content. This is a simple treatment of communicative purpose, which is a more sophisticated concept. However, we use this approach for comparison against the previous work.

In total, 48 images of each genre of e-mail were created. For each image we created four representations:

1. the *original* e-mail with no formatting or content changes;
2. the e-mail with the original formatting but with semantic content replaced with X or 9s (Figure 1). This version retained possibly useful structural formatting clues but gave no content to identify the e-mail. Successful identification of genre based on this version would indicate the role of structural form in identifying genre;
3. the e-mail with the original textual content but all structural formatting removed. This version retains punctuation but presents the text as a stream of text;
4. the e-mail with all content replaced by Xs or 9s (as in condition 2) and all structure removed (as in condition 3).

This version gives no indication of content and acts as a baseline to measure participants' attempts to identify e-mail genre.

Figure 1: Screenshot shows calls for papers e-mail with content removed(x's and 9's), but structure. The red lines indicate the scanpath of the eye-tracker showing fixation and saccadic behaviour.



Methodology

Task

Each participant was shown a total of 64 e-mails, and asked to identify each genre by voice, while the eye-tracking system recorded the most discriminating features, which led them to identify each genre representation. The eye-tracking equipment was fixed to the desk; only a simple answer to identify the genre was possible, because detailed discussions (head/face movements) would have interfered with the eye tracking.

To reduce any possible order effects, the types of e-mail and their allocation were randomized twice: once manually and once also by the eye-tracker software facility to randomize the stimuli for each participant.

The order of activities was thus:

1. Calibrate participant on eye-tracker
2. Show each stimulus
3. Ask for an identification of type of genre whilst eye-tracker records viewing behaviour.

There was a rest break after 32 images for about two minutes and, after the break, the calibration was repeated.

Independent variables

The variables tested were as follows:

1. Purpose/type of genre: (Tables 1 and 2)
2. Form: represented in four ways (figure 2).

Measurements

The measurements used in the experimental design were:

1. Mean fixation duration is normally used as an indication of information complexity, that is, the higher the more complex the mental load/task.
2. Mean gazing time
3. Saccadic rate per second which differs depending on task difficulty/mental load changes, that is, if task is difficult and cognitive processing increases then saccadic rate per second decreases and vice versa.

The additional measurements used in the experiment were the number of e-mail genres identified correctly by each participant and the length of time it took for each participant to identify each e-mail genre. The number of genres correctly identified measured the effect of the genre type and the time taken measured the effect of the e-mail form on the participant.

Results and Analysis

Several types of numerical data were collected. We report our own observations, providing an analysis and our interpretation of the results for discussion along with statistical tests.

Rate of correct genre identification

On average, in the course of all the experiments, 53% of the e-mails were identified correctly. Breaking these figures down:

- identification of the original format averaged 78%,
- identification of the original format but with semantic content removed averaged 71%.
- identification of the semantic content with no structure averaged 44%

- identification of the version with no semantic content and no structure averaged 22%.

Table 4: Correct genre identification counts

Genre	Orig	Form no cont	No form with cont	No form no cont	Avg
CFP	79%	82%	47%	21%	57%
Cin	73%	56%	30%	18%	44%
ITS	79%	68%	44%	12%	51%
Lib	79%	71%	47%	18%	54%
NI	70%	59%	36%	15%	45%
Ord	79%	75%	47%	9%	53%
Sem	76%	65%	38%	38%	54%
Spm	85%	91%	62%	41%	70%
Avg	78%	71%	44%	22%	
SD	4.5	11.7	9.6	11.7	

Removing all identifying features, not surprisingly, had lowest performance. However, it is higher than the 12.5% (1 in 8) performance we might have expected from pure guesswork. We conducted a Wilcoxon significance test on the average response times for the four conditions. Retaining format over content did lead to more successful and accurate genre identification (original e-mails identified significantly more accurately than no form/no content, (Mdn=55.5) $z=-2.524$, $p=0.012$, $r=.51$; formatting with no content significantly more accurately than no form/no content, (Mdn=49.0) $z=-2.371$, $p=0.018$, $r=.48$; content with no formatting significantly more accurately than no form/no content, (Mdn=49.0) $z=-2.521$, $p=0.012$, $r=.51$). This indicates that retaining either content or form significantly improves genre identification over simply guessing a genre.

As might have been expected, a common genre which was encountered by the participants on a regular basis (Spam 70% Call for Papers 57% Seminars and Library both 54% were identified correctly most frequently while the Cinema genre 44% was the least identified. Newsletters were identified 45% of the time, Orders 53% whilst ITS Outages were 51%.

Identification of genre gaze time-form

We compared the time it took to identify a genre, Table 5. The time recorded was

measured from when the stimulus was shown to the time when the participant gave a genre identification of the image. The eye-tracker system recorded this to the thousandth of a second. Average measures were taken for each representation of genre of e-mail as well as standard deviation.

Some genres were quicker to identify than others: the Orders genre was recognized with the fastest (4.23 seconds on average), while Call for Papers took the longest to identify (6.30 seconds on average).

Table 5: Genre identification response in seconds

Genre	Orig	Form no cont	No form with cont	No form no cont	Avg
CFP	3.76	5.24	5.72	10.50	6.30
Cin	3.20	4.30	5.08	6.57	4.78
ITS	2.72	4.43	4.03	7.16	4.58
Lib	3.48	3.87	4.76	6.32	4.60
NI	4.31	4.54	6.70	6.13	5.42
Ord	3.03	3.43	3.89	6.57	4.23
Sem	3.13	4.63	4.20	5.66	4.40
Spm	4.14	5.02	6.27	6.42	5.46
Avg	3.47	4.43	5.08	6.91	
SD	2.19	3.24	4.23	7.79	

We conducted the same (as above) pairwise significance test on the average response times for the four conditions. Retaining either e-mail formatting or content led to significantly faster identification than removal of both (original e-mails identified significantly faster than no form/no content, (Mdn=4.99) $z=-2.524$, $p=0.012$, $r=.51$; formatting with no content significantly faster than no form/no content, (Mdn=5.45) $z=-2.371$, $p=0.036$, $r=.48$; content with no formatting significantly faster than no form/no content, (Mdn=5.93) $z=-2.521$, $p=0.012$, $r=.51$). This indicates that retaining either content or form significantly improves genre identification over simply guessing a genre.

The original e-mails were identified significantly faster than any other condition ((Mdn=3.82) $z=-2.533$, $p=0.012$, $r=.51$, against formatting with no content; (Mdn=3.96) $z=-1.970$, $p=0.049$, $r=.40$, against content with no formatting; (Mdn=4.99) $z=-2.524$, $p=0.012$, $r=.51$,

against no format with no content;) indicating that both form and content are important for fast identification of e-mail and the presence of both leads to optimal identification. The original e-mail was identified fastest for all genres.

The version of the e-mails that retained structure but no content was identified significantly faster than the version with content preserved but no formatting (Mdn=4.49) $z=-2.524$, $p=0.012$, $r=.51$.

The only genres for which the e-mails with structure preserved were slower to be identified were the ITS and Seminar that merits further study at a later date.

Eye-tracking metrics analysis

Now we turn to the eye-tracking data to investigate how the participants' eye movements shed light on how certain types of texts may cause certain types of processing.

According to Rayner(1998), amongst others, fixations have been linked to intense cognitive processing. In Table 6 we show the average fixation count for each genre. There are two statistical differences. The original version of the e-mail vs. no form with content (Mdn=6.48) $z=2.521$, $p=0.012$, $r=.51$ and the original e-mail vs. form no content had fewer fixations (Mdn=13.79) $z=-1.820$, $p=0.069$, $r=.42$.

Table 6: Genre identification mean fixation count

Genre	Orig	Form no cont	No form with cont	No form no cont	Avg
CFP	7.62	8.91	9.77	8.85	8.78
Cin	5.02	7.10	6.18	8.06	6.59
ITS	3.75	10.04	5.65	5.60	6.26
Lib	6.73	7.38	7.85	6.65	7.15
NI	7.85	6.20	8.91	5.45	7.10
Ord	4.94	5.73	5.22	7.02	5.72
Sem	4.49	9.46	6.23	6.09	6.56
Spm	8.75	9.97	9.83	5.45	8.50
Avg	6.14	8.09	7.45	6.64	
SD	5.875	8.145	7.04	6.37	

The mean fixation *duration* is measured in microseconds (ms) and Rayner(1998, p. 373) claims this provides a useful guide into the

typical speeds which the participants perform certain tasks.

In Table 7 we present the mean fixation durations for each genre. There were no significant differences between the fixation durations for the various formats.

Table 7: Genre identification mean fixation duration (microseconds)

Genre	Orig	Form no cont	No form with cont	No form no cont	Avg
CFP	270	282	305	948	381
Cin	269	370	253	286	295
ITS	299	313	302	334	312
Lib	617	294	277	309	374
NI	269	370	253	286	295
Ord	227	392	028	027	169
Sem	278	336	292	297	301
Spm	291	293	296	270	288
Avg	315	296	251	345	
SD	278	313	277	297	

The saccadic rates per second (Table 8) indicate less or no cognitive processing (Rayner 1998). The original e-mail had a significantly higher saccade rate than the versions with no formatting but not the version that retained the formatting (original vs. no form/content, (Mdn=21.43) $z=-2.521$, $p=0.012$, $r=.51$; original vs. form/no content, (Mdn=21.65) $z=-.980$, $p=0.12$, $r=.20$. The version that retained formatting but obscured the content had a significantly higher saccade rate than the version that had no formatting or content (Mdn=19.74) $z=-2.100$, $p=0.036$, $r=.42$.

Table 8: Genre saccadic rate per second(p/s)

Genre	Orig	Form no cont	No form with cont	No form no cont	Avg
CFP	41.13	49.01	29.39	22.85	35.59
Cin	27.51	21.08	19.38	21.61	22.40
ITS	21.25	21.29	21.23	19.84	20.90
Lib	20.34	18.98	18.67	15.24	18.31
NI	22.01	19.65	19.95	16.08	19.42
Ord	25.81	18.17	17.16	19.37	20.13
Sem	26.40	27.13	22.93	18.98	23.86
Spm	23.53	20.74	21.21	14.21	19.92
Avg	26.00	24.51	21.24	18.52	

SD	24.67	20.91	20.58	19.175	
----	-------	-------	-------	--------	--

Regressive saccades indicate signs of less meaningful visual cues or features causing the eye to search repeatedly over the same areas looking for clues of identification during the search (Table 9). The original e-mail cause fewer regressive saccades than either version with no formatting (original vs. no form/content, (Mdn=12.12) $z=-2.521$, $p=0.017$, $r=.51$; original vs. form/no content, (Mdn=9.73) $z=-1.400$, $p=0.012$, $r=.28$, indicating that the loss of formatting caused not only more searching but the participant returned to areas already scanned. The version that retained formatting but obscured the content had a significantly lower regressive saccade rate than the condition which removed both content and formatting (Mdn=12.88) $z=-1.960$, $p=0.012$, $r=.40$.

Table 9: Genre Regressive Saccades

Genre	Orig	Form no cont	No form with cont	No form no cont	Avg
CFP	12.07	10.20	15.55	14.48	13.08
Cin	6.47	9.90	13.58	14.60	11.14
ITS	9.05	8.98	12.67	20.33	12.75
Lib	9.55	11.35	12.21	12.16	11.32
NI	10.54	11.63	14.33	26.30	15.70
Ord	8.16	7.44	13.05	17.14	11.44
Sem	8.75	12.67	7.31	22.73	12.87
Spm	7.84	13.08	14.56	20.52	14.00
Avg	9.05	10.66	12.90	18.53	
SD	8.90	10.78	13.32	18.74	

Summarising the results according to the four versions of each genre:

- the original e-mail was identified correctly more often than any other version and was identified quicker.
- the versions with content removed, but formatting retained, were identified less quickly than the original e-mail. However there were no significant differences between this version and the original in terms of number of fixations, duration of fixations and number of saccades. This version also had a similar regressive saccade rate to the original e-mail,

indicating that the formatting gave useful structural clues.

- the versions with structure removed were identified correctly less often than the original or the versions with no content. They also resulted in more fixations than the original and fewer saccades indicating a higher level of cognitive processing.
- the version with no structure or content had a low level of correct identification, high fixation duration and low saccade rate. It is interesting in this case that participants were obviously, from the response time and eye-tracking data, attempting to make sense of these e-mails rather than simply guessing a response.

Discussion

Our three research questions were as follows:

1. **What are the relative contributions of form (structure) and purpose (content) in identifying genre?** From the results presented it was clear that form played a significant role in genre identification. From the eye-tracking data there were many features which were deemed important for each genre from the small sample collected (Table 10). This verifies Watt(2009) and Toms and Campbell's (1999) work where they contended that the 'attributes' of a document's genre enable it to be specifically identified and showed that genre attributes play a significant role in identifying documents.
2. **What oculographic methods do humans use when viewing and utilizing the invariant layout cues, such as white-space patterns or other formatting features that constitute genres?** We noticed different strategies for different representations of e-mail. When participants were shown a normal or normal with semantic content replaced representation, they used a circular scanning motion or indeed a 'cross' strategy which consisted of a left-right and then up/down behaviour. By using Microsoft Excel's XY scatter charts we were able to plot all the fixations and saccades which were recorded in the eye-tracking viewpoint files which gave us a good indicator of the useful features used by the participants (see figures

3 and 4 at the end of the paper). When introduced to the unformatted representations, the participants very often used the right side of the block of text in an up and down scanning pattern to identify the genre

3. **Do/can participants 'skim' the shape of e-mail texts by possibly using saccadic behaviour?** In some cases, the participants did indeed skim the shape of the texts in the formatted e-mail stimulus examples. When the e-mail text was heavily formatted (centred), e.g. in calls for papers, or aligned left (seminars), the shape of the text, according to the eye-tracker data, did seem to play an important role in the genre identification process. In contrast, when the text had all format removed, some participants also occasionally skimmed the shape of the large 'blob' of text but this could mean that they were looking for semantic content/keywords.

Table 10: Important features

Genre	Feature(s) deemed important
Call for papers	Dates, centred blocks of text in top title, lists, dates and capital letters, bold text and uppercase letters.
Cinema	The rectangular shapes of the movie details (made up from blocks of numerical content for showtimes).
Newsletter	Text blocks and URL's spaced apart.
Spam	Normal version keywords like 'LOTTO', Viagra and address format text. Drawn to emboldened text at top of text.
IT Outages	Mostly identified by the first six-ten lines including the title. The emboldened titles were used as cues.
Library	The normal representation was identified by the book information at the bottom (list format).
Orders	Decimalised numbers, lists, tables, £ symbology, address block.
Seminar	First ten lines were deemed important. As the text was scanned up and down continuously. Many participants were using the left alignment of

	the text.
--	-----------

Post Comments

All the participants agreed that they understood the task given to them, but they did experience some confusion with regard to the number of genres that they were 'being asked to remember'. It was pointed out that this was a task in which perceptual processes were being tested and which therefore required them to simply attempt to identify the genres.

There was an incidence of contradictory information during the analysis of the questionnaire and eye-track data. One participant reported that, on occasion, numerical appearance (heavily prevalent, for example, in the cinema genre) was an important clue as to the purpose of the genres. Examination of the eye-track trace behaviour, however, showed that the participant had not looked at the particular area containing the numerical data! This was an important finding since it reinforced the value of eye-tracking during the experiments: although even without the eye-tracking system, we are able to infer what is happening when people interact with texts, eye-tracking gives us more accurate data and thus more valuable insights as to what is really happening.

Conclusions

The work described in this paper is the beginning of a thorough investigation into how people use texts in terms of genre and perception. With the aid of eye-tracking data analysis we can gather clues as to the processes that are used through the recorded oculographic behaviour hinting as to the amount of cognitive processing (fixations) or suppressions (saccades). It is perfectly feasible to say that purpose and form are very important for explaining the interaction with textual documents so that a 'community of practice' (Wenger 2000) can operate efficiently and also re-enforces the previous experimental findings by Toms & Campbell (1999) and Watt (2009).

We found, in our sample of a university 'community of practice', that readers can determine the purpose and form of genres and that during this process some readers do skim the shape or layout of the e-mails (form) whilst using the scanning method to find specific

information. Our future work will involve taking what we have learnt of this human oculographic behaviour and demonstrating how human categorization behaviour can be emulated by training a machine for automatic retrieval for example, text categorization

Acknowledgements

The authors would like to thank Professor Richardson and Dr Laura Muir from the School of Engineering at the Robert Gordon University. Without the kind permission to use their eye-tracking equipment we could not have collected such useful data as we have for this study.

Malcolm Clark would like to thank his mother Anne Mackay for proof reading the many drafts of the paper and casting her very neutral eyes over the flow of the content.

References

- Braisby, N & Gellatly, A 2005, *Cognitive Psychology*, In association with The Open University, Milton Keynes.
- Buscher, G, Dengel, A & Elst, Lv 2008, 'Query expansion using gaze-based feedback on the subdocument level', paper presented to Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval, Singapore, July 20-24, 2008.
- Clark, MJ, Ruthven, I & Holt, POB 2008, 'Genre Analysis of Structured E-mails for Corpus Profiling', paper presented to Workshop on Corpus Profiling for Information Retrieval and Natural Language Processing, BCS HQ, London, October 19.
- Gibson, J 1986, 'The Theory of Affordances', in *The ecological approach to visual perception*, 2nd edn, LEA, Hillsdale, New Jersey, p. 127.
- Gibson, JJ 1986, *The ecological approach to visual perception*, 2nd edn, LEA, New Jersey.
- Goldberg, JH & Kotval, XP 1999, 'Computer interface evaluation using eye movements: methods and constructs', *International Journal of Industrial Ergonomics*, vol. 24, no. 6, pp. 631-45.
- Gregory, RL 1966, *Eye and Brain the psychology of seeing*, 1st edn, World University Library, London.
- Joachims, T, Granka, L, Pan, B, Hembrooke, H & Gay, G 2005, 'Accurately interpreting clickthrough data as implicit feedback', in *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, Salvador, Brazil, pp. 154-61.
- Lakoff, G 1987, *Women, fire, and dangerous things: what categories reveal about the mind*, University of Chicago Press, Chicago.
- Michaels, CF 2007, *Gibsonian Perspective of Genre*, E-mail from Prof Claire Michaels edn, Storrs, Connecticut, 7 December 2007.
- Muir, LJ & Richardson, IEG 2005, 'Perception of sign language and its application to visual communications for deaf people', *Journal of Deaf Studies and Deaf Education*, vol. 10, no. 4, pp. 390-401.
- Nielsen, J 2006, *F-Shaped Pattern For Reading Web Content*, Nielsen Norman Group, viewed 22 September 2007 2007, <http://www.useit.com/alertbox/reading_pattern.html>.
- Poole, A & Ball, LJ 2006, 'Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects', in C Ghaoui (ed.), *Encyclopedia of Human-Computer Interaction*, Idea Group Inc, Pennsylvania.
- Rayner, K 1998, 'Eye Movements in Reading and Information Processing', *Psychological bulletin*, vol. 124, no. 3, pp. 372-422.
- Toms, EG & Campbell, DG 1999, 'Genre as interface metaphor: exploiting form and function in digital environments', in *Proceedings of the 32nd Hawaii International Conference on System Sciences*, Hawaii, p. 2008.
- Toms, EG, Campbell, DG & Blades, R 1999, 'Does Genre Define the Shape of Information? The Role of Form and Function in User Interaction with Digital Documents', *Proceedings of the ASIS Annual Meeting*, vol. 36, pp. p693-704.
- Underwood, G, Humphrey, K & Foulsham, T 2008, 'Knowledge-Based Patterns of Remembering: Eye Movement Scanpaths Reflect Domain Experience', in *Proceedings of the 4th Symposium of the*

Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for Education and Work, Graz, Austria, vol. 5298, pp. 125-44.

Watt, SNK 2009, 'Text categorisation and genre in information retrieval', in A Goker & J Davies (eds), *Information retrieval: Searching in the 21st Century*, John Wiley & Sons, Chichester, pp. 159-79.

Wenger, E 1999, *Communities of Practice: Learning, Meaning, and Identity*, Cambridge University Press.

— 2000, 'Communities of Practice and Social Learning Systems', *Organization*, vol. 7, no. 2, pp. 225-46.

Yates, JA & Orlikowski, WJ 1992, 'Genres of organizational communication: a structural approach to studying communication and media', *Academy of Management Review*, vol. 17, no. 2, pp. 299-326.

— 2002, 'Genre Systems: Structuring Interaction through Communicative Norms', *Journal of Business Communication*, vol. 39, no. 1, pp. 13-35.

Figure 3: Screenshot shows order e-mail with normal structure but no semantic content with fixations and saccades for participants 1-2 (data size reduced for reasons of clarity for the reader).

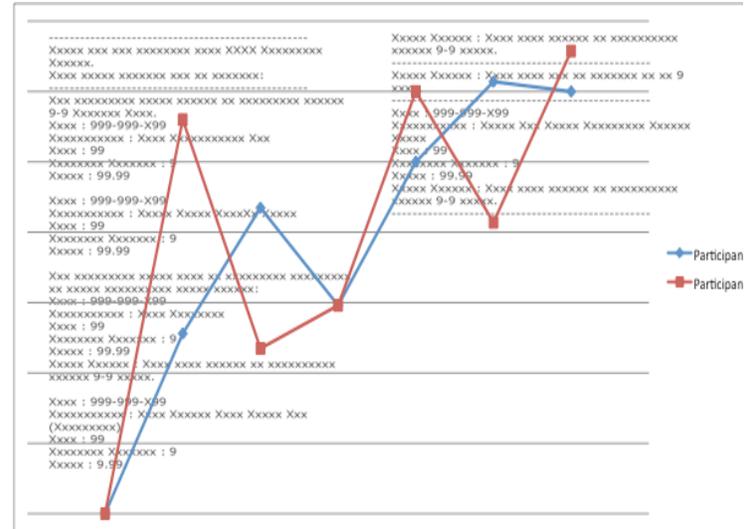


Figure 2: Screenshot shows original call for papers e-mail with fixations for four participants, each represented by different symbols.

