Task Impact on Cognitive Processing of Narrative Fiction Film

Lund University, Department of Philosophy

Master's Thesis in Cognitive Science, May 2006

Author: Janna G. Spanne
janna.geggus-spanne.675@student.lu.se

Advisor: Kenneth Holmqvist, LUCS / LU Humanities Laboratory

Abstract

The purpose of this study is to investigate a segment of the process by which a film spectator extracts information from moving images and treats it cognitively to generate an integral experience of a film. A comparison is attempted between objectively measurable eye movement data and the participants' own reports of their impressions of a film sequence under three different task conditions.

The results show that the cognitive processing of film information is influenced not simply by the presence or absence of a task, but by a combination of specific task content and the ways in which information is presented in the film, that is, its style.

These findings are tentatively connected to theories of frugal heuristics in everyday decision-making and their relevance to the cognitive processing of film information.

Acknowledgements

I am greatly indebted to:

- Kenneth Holmqvist, my advisor, for introducing me to the possibilities of eye tracking methods in film studies and for being a friendly and efficient guide in the empirical maze;
- Lenisa Brandao, Kerstin Vikström, and Paulina Lindström, the participants of the Spring 2006 eye tracking seminar in the LU Humanities Lab, for valuable comments on several drafts of this paper and practical help in carrying out this study;
- Petra Björne, LUCS PhD candidate, for important pointers on experiment design, for her contribution towards improving the initial project description, and for moral support in times of need;
- Marcus Nyström, doctoral student in Information Technology, LTH, for several novel points of view on the interpretation of the data, for clarifying discussions of data analysis and for performing the MATLAB calculations;
- Sven Spanne, for asking the right uncomfortable questions at the right moment.

1. Background and goals

Traditional film theory consists in philosophical debates about the nature of film art, analyses of its means of expression, biographical studies of individual film artists, abstract application of various psychological and social models on film viewing, or study of the economic and marketing aspects of the film industry. – A cognitive psychologist, on the other hand, may oversimplify by considering film a "dynamic real-world display" (Germeys 200?), disregarding the various ways in which the spectator is conscious of narrative fiction film as a fiction and an illusion (Anderson 1996).

Both these approaches suffer from the lack of empirical data relating to the cognitive processes involved in film comprehension, and there is in fact very little such data available. Cognition and film scholars agree that the cognitive processing of film information is an interplay of top-down and bottom-up mechanisms (Bordwell 1987, Anderson 1996, Grodal 1997, May, Dean & Barnard 2003, Nyström, Holmqvist & Novak 2005). Just how these mechanisms interact and how much each of them contributes in a given viewing situation is a broad empirical issue so-far rarely addressed.

Applying empirical methods from cognitive science in the field of film studies is a novel approach. It should facilitate interdisciplinary exchange between film studies and cognitive science by exploring the possibilities and limitations of using film in the study of human attention and decision-making, and by supporting cognitive film theories with empirical data. Such data might prove constructive to applied issues such as the debate about the impact of film on moviegoers' attitudes and behavior.

In the context of this study, the term "film" refers to narrative fiction film. Issues such as film as an abstract art or the delimitations of the documentary are not taken into account.

1.1 Motives and theory

The present work is intended as a starting point for the empirical study of some cognitive processes involved in film comprehension. It should indicate whether tasks given to the spectators influence how they will assign visual attention in following a film sequence, how such tasks influence spectators' conscious evaluation of the sequence and, hence, how visual attention relates to the spectators' impression of a film.

1.1.1 Eye tracking and moving pictures

In the past it has been extremely difficult to obtain information about those aspects of the film-viewing process that aren't accessible to introspection and self-assessment. Recent advances in technology have greatly improved the possibilities to collect and process the data necessary for empirical studies of the cognitive processing of film, allowing for a view of the spectator as a complex organism interacting with a complex environment. In particular eye tracking methods have evolved to be sufficiently non-restrictive to approximate a natural film-viewing situation, and we have processing methods capable of dealing with the quantities of data yielded by eye tracking studies of dynamic stimuli.

Eye tracking methods are based on the fact that humans make detailed visual observations in an area extending over a visual angle of only about 2°. (Glenstrup & Engnell-Nielsen 1995) Thus humans have to move their eyes to scan different parts of a scene, and these eye movements are closely connected to visual attention. (Deubel & Schneider 1995) Eye movements are said to be "an overt behavioral manifestation of the allocation of attention". (Henderson 2003) As it does not rely on introspection, eye tracking may be the best single method currently available for the empirical study of the film-viewing process.

There has been very little empirical work to-date in the field of film theory, eye tracking related or otherwise. As the overview in Henderson (2003) shows, a considerable amount of

research has been done on human eye movements and gaze control in the perception of real-world scenes and still photographs, including scene statistics, saliency maps, the role of memory, background knowledge and task relevance. None of it, however, seems to have been applied to film stimuli. Duchowski (2002) cites several eye tracking studies on the perception of art, and one study by d'Ydevalle et al. on editing "errors" in film.

A thorough metatheoretical overview of film studies, Björklund & Jansson (2006), shows that established film theories generally share the view of the spectator as an abstract entity to be ascribed attitudes and reactions suitable to each theory. The communicative gap that undoubtedly exists between film theorists and researchers in cognitive and behavioral science is largely due to differences in attitudes towards empirical studies of spectator behavior.

For reasons of the history of film studies, rooted in subjects such as aesthetics, literary and theatre studies, there is no tradition of empirical work in the field. A relatively new branch, cognitive film theory, is represented by Bordwell (1987), Branigan (1992), Anderson (1996) and Grodal (1997). It focuses on high-level cognitive processing of film information, asking question such as "How do film viewers turn the information provided in a film into a coherent narrative?" or "How does a film generate emotional responses in a spectator?" Unfortunately, even though cognitive film researchers often refer to findings from cognitive science, neuroscience and evolutionary psychology, they rarely give more than second- or third-hand references to actual empirical work done in those areas. But cognitive film theory does differ from traditional film theories in that it enables researchers to formulate falsifiable hypotheses and so, in principle, opens for the possibilities of empirical study.

A fundamental issue is what properties of the spectator influence his/her visual attention in viewing a film. Goldstein et al. (2004) and Nyström, Holmqvist & Novak (2005) indicate that, in free viewing, spectators attend to largely the same areas, regardless of gender and age. These findings suggest that differences in spectators' impressions of a film arise on a higher cognitive level than that of visual attention. It is tempting to view these results as contradicting traditional feminist film analysis (Mulvey 1976, Stam 2000) with its notion of an inherently masculine "gaze", but, in fact, gender theorists do not provide a falsifiable definition of "gaze". Further eye tracking studies, combined with spectator surveys, might help set up an empirical foundation for a gender theory of film perception.

Other empirical studies carried out in connection with technical applications also suggest approaches potentially useful to film theorists:

May & Barnard (2003) propose to use film editing techniques to improve computer interfaces. Their study is a so-far unique empirical work relating visual processing to a filmmaking technique. They have set up a taxonomy of film editing based, not on a specific film theory, but on the position of important items on the screen before and after a cut.

d'Ydevalle et al. (1991) study the impact of subtitles on eye movements, showing that attention to subtitles is automatic and compulsory. Their study considers film as multi-modal through its attention to both spoken and written (subtitled) dialogue.

Nyström, Holmqvist & Novak (2005) propose a mathematical method for computing gaze density functions (gdf) from a collection of data by numerous viewers, and for analysis by tracking and visualizing changes in gaze density function values over time, in order to predict which parts of a picture should provide high-quality visual input. Their method is designed to be highly reliable and seems eminently suitable for use not only in video compression applications, but also in general studies of visual processing of film information. Although the present study does not apply their model to its full extent, their article was instrumental for designing the experiment, and calculations originally developed for their compression method were used to analyze the data discussed here.

1.1.2 Still vs. moving pictures

Through his experimental work in the 1950s and 60s A.L. Yarbus conclusively showed that, given different tasks, a viewer's gaze will take different paths in scanning a still picture (Yarbus 1967). Eye movement patterns while viewing a given still picture vary considerably in the same individual depending on the task assigned, for example 'Just look at the picture' / 'Give the ages of the people' /. 'Estimate the material circumstances of the family'. (Figure 1:1)

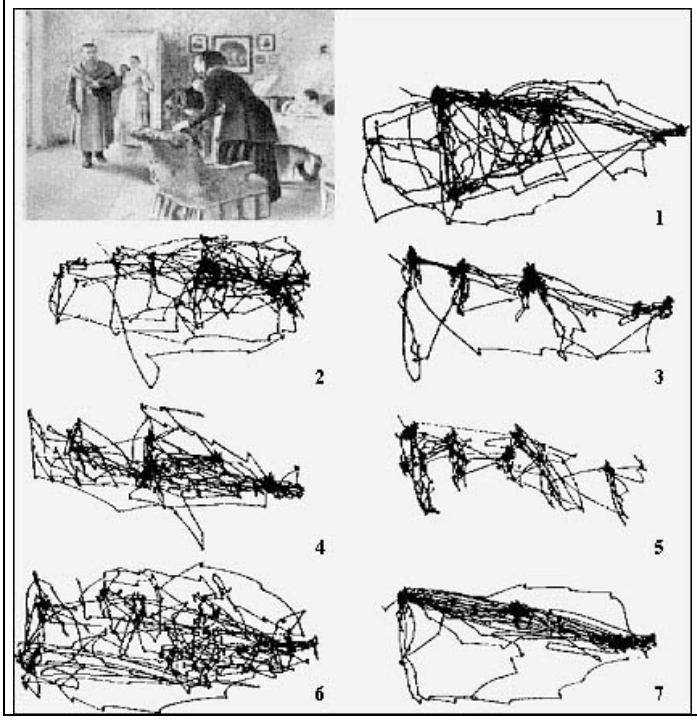


Figure 1:1: Impact of task on the viewing of still pictures (Yarbus 1967)

1. Free examination
2. Estimate the material circumstances of the family
3. Give the ages of the people
4. Surmise what the family had been doing before the arrival of the "unexpected visitor"
5. Remember the clothes worn by the people
6. Remember the position of the people and objects in the room
7. Estimate how long the "unexpected visitor" had been away from the family

It is easy to assume that something similar should apply for moving pictures as well. There are, however, good reasons to believe that the mental processing of moving images differs from that of stills. Movement as such tends to capture attention. Viewed in a perspective of evolution and adaptation to an environment, survival situations (feeding, predator avoidance, mating, social interaction) are typically dynamic, while static pictures are a rather late human construction related to symbolic communication (Persson 2005).

Although a film does not equal the reality that it purports to depict, film resembles the physical world more than still images do, in ways that are crucial to perception and cognitive processing:

1) **Movement** serves as an attentional cue in perception of dynamic situations (Johnson & Proctor 2004). This is true for natural situations, and from the point of view of sensory perception, film is a part of our physical world (Anderson 1996).

2) Like real-life dynamic situations, film has its own **temporal** dimension: information is presented in a temporal sequence at a determined speed, over which the spectator has no influence (Bordwell 1987). In the perception of still images, the temporal dimension is provided by the viewer scanning the picture. (Figure 1:2)

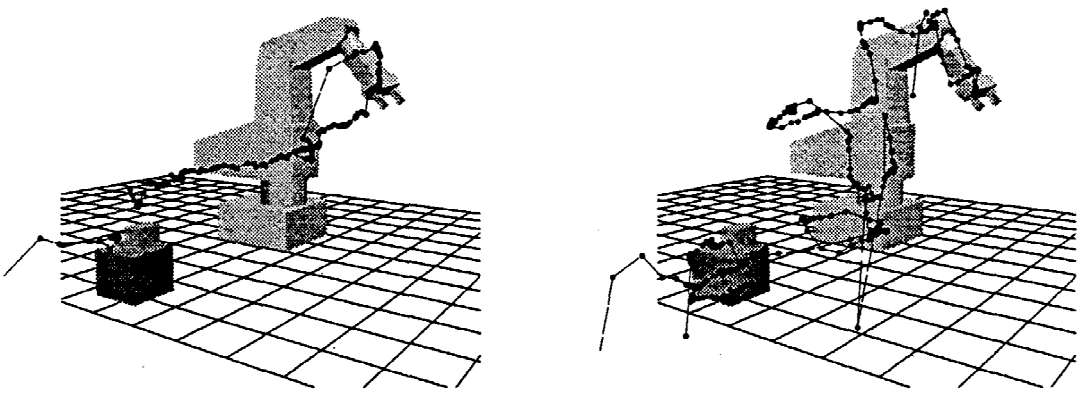


Figure 1:2: The viewer's scan path provides a temporal dimension a still picture.
"Eye movement while viewing a dynamic image sequence (left) and a static snapshot of the same scene (right). Recorded eye position data is indicated by single dots connected with lines between measurements." (Blackmon et al. 1997)

3) Like the physical world, film presents information **multimodally**, while a still picture is typically a strictly visual experience. The modalities are fewer than in real-life situations; mainstream fiction film does not purport to impact the audience's sense of smell, taste or touch, but it does impact their hearing. Sound is of crucial importance for the ways a film affects a spectator (Bordwell & Thomson 2003), and not even so-called silent film was ever intended for display without the accompaniment of sound. Even if the present study is limited to visual perception, the impact of sound must be taken into account in any complete model of film comprehension (Johnson & Proctor 2004).

1.1.3 The whole story

As every moviegoer knows, and as film theorists agree, there is much more to the film experience than the cognitive processing of visual percepts. If the goal of a cognitive film theory is to construct a model of how spectators arrive from the perception of colored patches on a screen to the understanding of a narrative and to an emotional experience related to viewing a film, efforts must be made to examine how the various steps in the cognitive processing of film information relate to the spectator's film experience as a whole.

Eye tracking experiments may say a lot about how task priming affects visual attention, but, in order to relate an eye tracking study to the actual film-viewing process, it should be complemented by an examination of the participants' subjective evaluation of the stimuli. That is why the present study contains not only a task-related eye tracking experiment, but also an attempt to examine the impact of the same tasks on the participants' evaluation of the film sequences presented to them.

1.2 Specific goals

For reasons presented in Section 1.1.2, we cannot assume that findings from still picture viewing will be directly applicable on moving pictures. Also, similarities between film and natural scenes suggest that theories relating to cognition in everyday situations will have a bearing on how we view the cognitive processing of film. The present study constitutes a step towards an adequate hypothesis about task influence on the visual processing of moving pictures. It also purports to show that theories of simple/frugal heuristics may provide a tool for the analysis of cognitive processing of film.

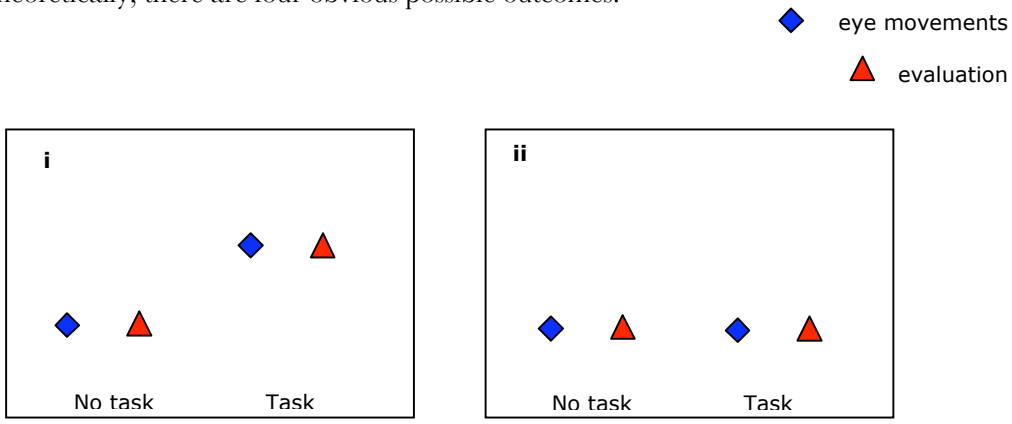
It does so by addressing the following questions:

a) Does tasks priming yield significantly different eye-movement patterns?

b) Does tasks priming yield significantly different evaluations of the stimulus sequences?

c) What are the connections, if any, between eye movements (a) and the spectators' conscious evaluation of a film sequence (b)?

Theoretically, there are four obvious possible outcomes:



i) eye movements and evaluation both vary with task

ii) neither eye movements nor evaluation are influenced by task

iii) eye movements but not evaluation vary with task

iv) eye movements do not, but evaluation does vary with task

While (i) and (ii) are fairly straightforward, the task makes no difference / makes a difference all the way, (iii) and (iv) would yield questions relevant to cognitive film theorists:

(iii) How do spectators arrive at the same conclusions in spite of looking at different parts of the image?

(iv) How do spectators arrive at different conclusions from looking at the same parts of the image?

Other relevant questions are:

d) How does the form of the stimulus affect the impact of task? (see Section 2.1)

e) How does the content of the task affect its impact? (see Section 2.3)

f) How do the subjects' backgrounds affect the impact of task?
This last question was not addressed in this study, but efforts were made to control background factors that might be important. (see Section 2.2)

2. Experiment description

2.1 Stimuli

Since properties and content of an object or a situation, such as color, onset of movement, presence of human faces and hands, influence visual attention (Yarbus 1967, Johnson & Proctor 2004, Wolfe & Horowitz 2004), two film sequences with different visual properties were used in the actual experiment.

The participants in the experiment watched the same two film sequences of 10 minutes each, both including a man-woman dialogue scene of 2-3 min. The order of presentation was reversed for half the participants. The entire viewing situation in the experiment was not "natural" in the sense that the subjects did not have access to the entire narrative context preceding the stimulus sequences. The study's main goal, however, was to examine how spectators extract information from a scene, not how they incorporate it into a coherent narrative.

The sequences came from two mainstream American films, *Die Hard 1* (1988) and *Armageddon* (1998), both box-office successes in the action genre, with a wide international distribution. The intention was to choose films aimed at and successful before a large general audience.

2.1.1 A matter of style

The style of a film comprises the film's various ways of presenting content, such as: editing speed and rhythm, camera movement, angles and framing, actors' speech and body language, music and other sound, sets, costumes. Style is of crucial importance for a film's impact on a spectator's visual attention (Bordwell & Thompson 2003). In order for the experiment to measure the impact of task and not just the impact of a particular style it was necessary to choose two stimuli obviously unlike each other, while remaining within the framework of mainstream entertainment. A previous study (Spanne 2004) shows that the two films from which the stimuli were taken differ both in style and in ideology, *Armageddon* adhering closer to the High Concept mode of film making and marketing and being politically more conservative than *Die Hard 1.*

The ideas of High Concept (Wyatt 1994), where film is primarily a marketable commodity rather than a narrative medium, have to some extent impacted all mainstream Western filmmaking in the past 20 years. If we consider classical Hollywood narrative and High Concept as two end points of a scale, the corresponding values of their defining traits are summarized in the following table.
(After Bordwell & Staiger1988, Wyatt 1994, Hedling 1999, Spanne 2004):

	Classical Hollywood	High Concept
Narrative	**Causality** Story and character development	**Attraction** Exposure of attractive objects, people and places
Characters	**Individuals** Personal motives and development as narrative causal agents	**Stereotypes** Stereotyped motives, little discernible development
Actors	**Character personifiers** Craftspeople expressing the characters' individual motives and development	**Iconic stars** Schematic figures interconnected with their previous films, other pop culture references and exposed private lives
Music	**Narrative support** Integrated into the narrative, illustrates environment, mood, relationships between characters	**Marketing synergy** Soundtrack CD planned as a marketable product in its own right as well as an aid in marketing the film
Style	**Credibility** Convincing images of objects and environments relevant to the narrative and supporting causal connections	**Excess** Surface, design, technical sophistication, product placement

The classical Hollywood end of the spectrum, represented here by *Die Hard*, is a style where most formal aspects are at the service of the narrative (Elsaesser & Buckland 2002, chapter 2). Editing rhythm varies with the needs of the narrative, and extremely fast cutting is rare. A typical dialogue sequence has an average shot length (ASL) of 2 seconds, a very frequent value for narrative film (Elsaesser & Buckland 2002, chapter 3). Camera angles depict the characters' status and relationships; music creates a seasonal mood or tells of the villains' European origins; if an expensive wristwatch is mentioned early on, it will have a role in saving its wearer later; apart from one semi-famous star the actors simply represent their characters. (Prior to *Die Hard*, Bruce Willis had a lead role in one romantic comedy and a television series.)

In extreme High Concept, represented by *Armageddon*, style is a marketing tool; the film has been called "the first 150-minute trailer" (Ebert 1998?). Editing is adapted to the habits of the MTV generation, the film's primary intended viewer group, with many extremely fast-cut sequences (ASL 0,5 sec) consisting largely of light flashes (explosions, meteor impacts). Camera work serves to present celebrities on screen to their greatest advantage; music composed by other celebrities is adapted for record sale, advanced technology is displayed as "eye candy" without contributing to the narrative; the spectators are expected to be knowledgeable about the stars' private lives so as to add an extra emotional dimension to the story (for example, Bruce Willis, father of three daughters, plays the loving father of the character played by Liv Tyler, real-life daughter of Steven Tyler of the group "Aerosmith", who provide the music for the film).

The two dialogue sequences differ greatly in their depiction of the women:

In *Die Hard*, the hero's wife, who is a high-ranking executive, confronts the villain in a territorial contest, making demands, taking initiatives in the dialogue; she is seen closing in on him and getting him to look up on her.
In the *Armageddon* sequence, the hero's daughter, who has never been seen working or studying, says goodbye to her father at a distance, over a video link from space; she is fragile, surrounded by cold machinery, looking up at a huge screen that displays his image.

A pilot study with four participants, prior to the current experiment, showed that these spectators (although film students conscious of the impact of style) really did respond to the style clues and rated the two women according to expectations.

2.2 Subjects

The experiment comprised 33 subjects, with normal or corrected-to-normal vision, able to follow a mainstream American film without subtitles and to answer a written questionnaire in English. Eye tracking data from two participants was lost due to technical problems, leaving data from 31 subjects to analyze.

Rather than using a group of first-year film students, the intention was to approximate a "normal" cinema audience, even if people with an explicit interest in film would necessarily be somewhat over-represented in a group of volunteers for an experiment of this kind. In a background questionnaire given in advance of the data collection session, the subjects were asked to provide some preliminary information: age, gender, primary language, educational background, film viewing habits and preferences. (Se Appendix I). Based on their replies to the questionnaire, the subjects were divided into three groups, as far as possible matched for age, gender (balanced in each group) and film viewing preferences.

A question asking the participants to rate a number of mass-market films was included, not only in order to assess the participants' attitude towards mainstream entertainment film, but also to find out whether they remembered having seen (and liked /disliked) the films that the stimulus sequences were taken from. The groups could not be exactly matched for pre-knowledge, but each group had several subjects who had not seen either or one of the films, and there was a comparable within-group variation as to the subjects' remembered opinion of the ones that they had seen.

The participants were informed that their eyes would be filmed. No other advance information about the experiment was volunteered, but the eye tracking work at the Humanities laboratory is widely known among the population from which the sample was taken. It is, however, considered highly unlikely that subjects would be able to consciously manipulate their eye movements over 20 minutes so systematically as to influence the outcome. The participants had no advance knowledge of the specific hypotheses that the experiment was intended to test.

After the test, subjects were directed to a web site where a description of the project was available, and offered a copy of the resulting thesis. They were also given contact information to the researcher for further information.

The subjects were made aware that the collected data might be explored and analyzed further in the future, and asked to give their written consent. They were again reminded that they could, without motivating, at any point request their own data to be deleted from the material.

2.3 Tasks

Yarbus' study (1967) shows that eye movements vary depending on the contents of the task given to the person examining a still picture, not simply due to the presence/absence of a task. If the main point of the study had been to apply Yarbus' questions on dynamic material, several specific tasks not involving the spectator's attitudes would have been set.

In a design involving simple presence/absence of a task, it would be impossible to ascertain if an effect was due to the presence of a task as such, or the nature and content of the specific task given. Thus, two different tasks were designed in addition to the null task.

As the study was geared towards examining the cognitive processes in a typical film-viewing situation, that is one connected to emotions and values rather than the transmission of factual information, it was important for the tasks to be evaluative.

Group 0 was asked to watch the sequences without further specification (*Condition 0*)

Group 1 was asked to determine whether the sequences made them want to see the entire films (*Condition 1*)

Group 2 was asked to form an opinion about the female characters (*Condition 2*)

Condition 0	Watch the following two film sequences.
Condition 1	Watch the following two film sequences. Decide if you would be interested in seeing the films that they come from.
Condition 2	Watch the following two film sequences. For each one, form an opinion about the woman who takes part in the dialogue.

"Wanting to see the entire film" was used as an indicator of a spectator's general impression, his/her overall attitude to the film as elicited by the sequence. The intention was to set a general evaluative task.
"Opinion about the female character" was intended as a specific, attitude-related task, as opposed to a possible specific, attitude neutral task (for example, "How many long-haired people are there in each sequence?").

2.4 Procedure

To approximate a natural viewing situation as closely as possible, eye movements were monitored with the aid of a head-mounted eye tracker, rather than table-mounted apparatus that restricts the subject's movements. Apart from instruction and calibration periods, the subjects were left to view the stimulus sequences on their own.

The stimuli were presented by a video projector on a screen, giving an image size of 105x83 cm. Participants were placed in an easy chair at a distance of approximately 2 m from the screen.

Each data recording started with a spatial calibration scheme where subjects were in instructed to move their gaze to each of 13 static points presented simultaneously on the display and pointed out to them sequentially.

The task assigned to his/her group was presented to each participant on paper, after the spatial calibration. The participants were requested to read the task at their own pace and to ask for clarifications if needed. Then the task sheet was removed and temporal calibration data collected: subjects were asked to follow with their gaze a red dot making an oval motion across a black background. The temporal calibration could then be made off-line by matching each viewer's gaze spatially with the red dot at each frame. The temporal calibration stimulus was presented as a part of the video sequence and the participant was left alone to view the entire video material.

The eye-tracker was an SMI iView using a pupil/cornea reflex system, worn as a bicycle helmet, compensating for head movements by means of a Polhemus magnetic motion tracker. Eye-movements were sampled at a rate of 50 Hz; hence two points of gaze were obtained from each subject and video frame. The output was recorded as a coordinate file in the coordinate system of the film projection.

After watching the two sequences once, the participants were asked to fill in two questionnaires for each sequence: one to give their general opinion of the film, the other calling for their evaluation of the female character. In order to elicit spontaneous replies, the response time was limited to one minute for each questionnaire. (See Appendix II)

Each questionnaire consisted of a set of twelve pairs of words or phrases representing the end points of a value scale, such as "beautiful – ugly", "shallow – deep", "worth seeing – not worth seeing". The general evaluation questionnaire not only called for the subject's personal impression, but its wording also evoked some positive/negative stereotypes frequently found in film reviews geared towards the general public. The female character questionnaire was intentionally designed to evoke the sexist stereotypes that mainstream American film is generally said to propagate (Stam 2000, Elsaesser & Buckland 2002).

2.5 Data

In the experiment three categories of data were collected and analyzed: replies to the background questionnaire, eye tracking data, and replies to the task-related questionnaires rating the films and the female characters. The data was treated anonymously, under a code system enabling the researcher to identify and delete a subject's data should the subject request it.

2.5.1 Background and film habits data

The background questionnaire data was intended to provide a basis for dividing the subjects into three balanced groups relative to age, gender, film-viewing habits, and previous knowledge of the films that the stimulus sequences were taken from. The background data was used solely to obtain three participant groups that were as balanced as possible, and was not analyzed further for correlations with the experiment results in the context of this study

The participants' age varied between 18 and 51 with an average of 33. There was a certain over-representation of women, with a female/male ratio of 1,4.

Each possible reply related to education and film-viewing habits was assigned a value in points, where high values indicated a deviation from the pattern of a regular consumer of mainstream entertainment film. Points were assigned for education in aesthetic subjects, regularly seeking out non-mainstream film and giving low ratings to the listed mainstream film titles.

The highest possible rating was 70 points; the participants' ratings varied between 17 and 45 with an average of 30,8.

2.5.2 Eye tracking data

Recording eye movements by means of head-mounted equipment while showing the stimulus sequences on a large screen goes a long way towards approximating a natural film-viewing situation, as the subjects can be seated in a relaxed position and their head movements are minimally restricted.

The recordings yielded a separate data file for each participant, containing a set of coordinates representing the positions of two fixations for each frame. This raw data was used as input to a Matlab application for analysis by condition, referred to as condition/group 0, 1 or 2 based on the task assigned to each participant.

2.5.3 Film and character rating data

The task questionnaires asked for the subjects' personal impression of the two films and of the female characters, that is questions related to the tasks used in the experiment. Each participant answered both the questionnaire relevant to his/her task, if any, and the "irrelevant" one.

The questionnaires were evaluated so as to yield interval data, with an arbitrary zero at the conventionally negative end of each scale, and a maximum rating of 9 at the conventionally positive end. A mean value was calculated for each questionnaire, yielding a highest possible rating of 9. For example, a person who moderately disliked the *Die Hard* sequence would give replies that returned a mean rating of 3,25, one who very much liked the woman in *Armageddon* would reply in ways that yielded a mean rating of 7,33.

The raw data was coded into an Excel spreadsheet for further analysis by condition, as in the eye tracking data.

2.6 Analysis

a) One main issue of the study was whether the task given to the subjects would influence their eye movement patterns while viewing the film sequences.

The null hypothesis was that all three groups would foveate largely the same areas in each scene. If the presence/absence of a task and/or its content does matter, there would be systematic, statistically significant differences in fixation patterns among the three conditions. As gaze density functions are a systematic way of determining what eye movements are "normal" in a given group, such differences should be discernible as variations in correlation between gaze density functions of the different groups.

b) In parallel, it was important to determine whether the tasks would influence the subjects' evaluation of the film sequences and the female characters.

The null hypothesis was that there would be no significant difference in the evaluations under the different task conditions. Otherwise, a condition might yield higher response values for its relevant questionnaire than for the irrelevant one, or vice versa.

c) Finally, to determine how visual attention relates to subjective evaluation, results of the eye movement data analysis would be compared to those of the evaluation data analysis. Theoretically, four outcomes were possible, as presented in Section 1.2.

As the stimulus consisted of two film sequences with considerably different formal properties, the data was analyzed in its entirety, but also separately for each sequence, in order to examine how the manner of presenting content impacts the cognitive process.

2.6.1 Eye tracking data analysis

The eye movement data from each group of subjects was processed in a Matlab application by the method proposed in Nyström, Holmqvist & Novak 2005, to yield a gaze density function for the entire group.

Prior to comparing gaze density functions for the various groups, in order to determine when two such functions should be considered different, self-similarities were generated. These are a way of determining for each gaze point how high the probability of its position is in a given gaze density function, that is how "typical" it is of a given set of gaze points, or how great the variation in gaze point position is within a condition.

As the first step in the analysis, an ANOVA for the entire material by condition was calculated, to examine differences in self-similarities between the three groups

In the next step, gaze density functions were compared: i) between the control group and each of the task groups, and ii) between the two task groups. The initial comparisons were made by three different methods: point dropping, Mean Square Error, and correlation (Nyström, Holmqvist & Novak 2005). These are three alternative ways of expressing the same relationships. As the correlation graph was easily readable and the term "correlation" has an appropriate intuitive meaning, only correlation was used for the continued analysis. The variations in correlation were mapped over time and examined for significantly high/low correlation values and for differences in correlation between the various pairs of conditions.

Since the temporal dimension is one of the most important features distinguishing moving pictures (film) from stills, the next step was to consider the variations in self-similarities for the various conditions over time. The differences summarized in the ANOVA box plot were plotted over time, using a sample of one frame in ten from the stimulus sequences, in the order 1) *Armageddon*, 2) *Die Hard*.

Pairs of conditions were then examined for correlations and graphed over time. This may be considered a type of post-hoc analysis when the analysis of variance of self-similarities yields significant results.

For continued analysis the data was broken down to examine each film segment separately in the manner described above: box plot of self-similarities, variation in self-similarities over time, and pair-wise correlation between conditions over time.

Salient features such as conspicuous differences or similarities in the graphs were identified in order to link them to the content of each relevant sequence. To facilitate this, a video file was generated of the sampled frames (one in ten from the original stimuli) with overlaid markers for individual fixation points color-coded by condition.

2.6.2 Film and character rating data analysis

The evaluation questionnaires were rated in points as described under "Data" in Section 2.5.3.

The raw data consisting of a mean value for each questionnaire and participant was coded into an Excel spreadsheet for further analysis by condition. Another set of data was obtained by subtracting the *Armageddon* rating from the *Die Hard* rating for each task and participant.

participant	condi-tion	*Armageddon* film rating	*Die Hard* film rating	film rating *Die Hard* - *Armageddon*	*Armageddon* woman rating	*Die Hard* woman rating	woman rating *Die Hard* - *Armageddon*
adh001	0	3,67	3,00	-0,67	4,25	6,58	2,33
adh003	0	2,42	3,50	1,08	2,33	7,58	5,25
dha104	1	2,83	3,08	0,25	1,67	6,67	5,00
dha106	1	6,42	5,42	-1,00	3,83	4,33	0,50
dha107	1	2,67	5,58	2,92	3,75	6,17	2,42
adh201	2	2,40	5,83	3,43	6,25	8,25	2,00
adh202	2	3,08	3,92	0,83	4,50	6,58	2,08

Figure 2:1: Example of film and woman evaluation data.
Ratings of the films, the female characters, and rating differences between the two segments (*Die Hard*-based rating – *Armageddon*-based rating)

To get a rough general idea of the tendencies in the data, overall means were calculated for each question and film, and for the rating differences between the films. Then a statistical analysis was performed in SPSS, yielding scatterplots, box plots and ANOVA tables. These were examined for significant differences.

2.7 Results and summary

2.7.1 Eye tracking data

A first analysis of the eye tracking data showed that it was approximately normally distributed. An ANOVA test was performed to determine if there were any significant differences, that is, whether or not the three task groups could be said to represent different population. ANOVA for the entire material by condition showed significant differences in self-similarities for the three groups, that is, the three gaze point samples did not come from the same population. The ANOVA yielded a value of F= 289,21.
Thus the overall null hypothesis could be rejected.

As Figure 2:2 shows, the fixations in condition 1 were more dispersed than in condition 0, and those in condition 2 significantly more dispersed than in both 0 and 1.

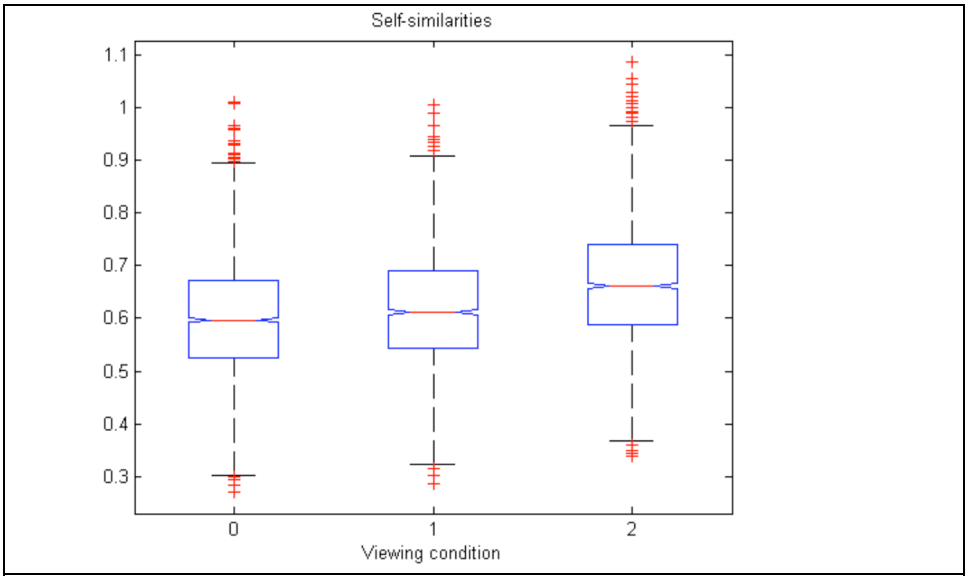


Figure 2:2: Gaze dispersion as a function of viewing condition for the entire set of data.
Box plot with wedges at different levels shows significant differences in self-similarities for all three conditions, with Condition 2 differing markedly from Condition 0 and Condition 1.

The temporal dimension being extremely important when analyzing dynamic material, the variations in self-similarities summarized in the box plot were then plotted over time. one frame in ten from the stimulus sequences was used for analysis, in the order 1) *Armageddon*, 2) *Die Hard*. (Figure 2:3)

Figure 2:3: Gaze dispersion for each condition over time.
Transition between the sequences is marked by a pale blue line.
1) *Armageddon*, 2) *Die Hard.*
Note the difference in the self-similarities between the two sequences.
Conditions: 0 = no task, 1 = film evaluation, 2 = woman evaluation

The graph revealed considerable variation both within each condition and between conditions, relating to which segment was being analyzed. In particular there was virtually no overlap between condition 2, the female evaluation group, and the remaining two, in the *Die Hard* segment: condition 2 showed consistently higher gaze dispersion throughout the segment, with particularly high dispersion values in a sequence just past the middle of the segment.

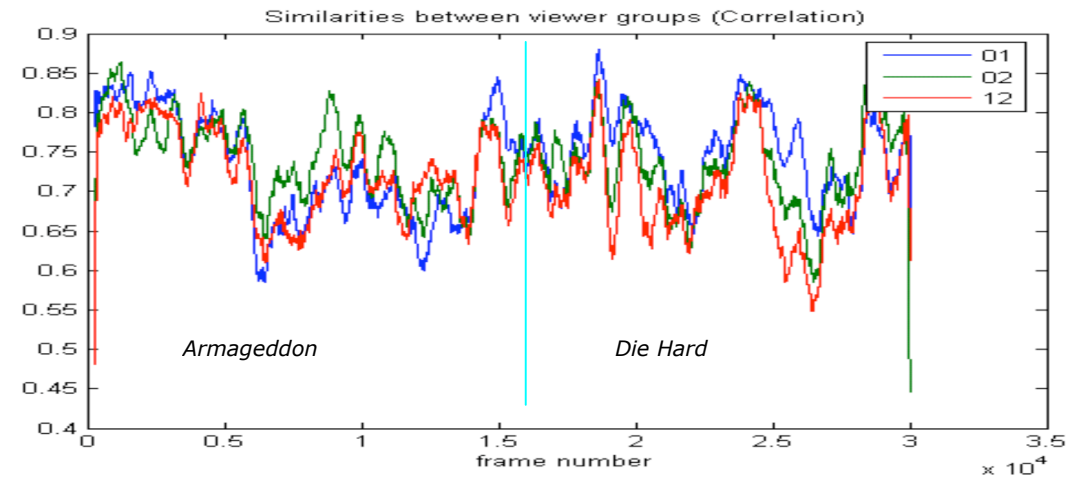


Figure 2:4: Correlation for pairs of conditions over time. The overall degree of correlation and individual pair-wise correlations both display a greater variation in the *Die Hard* segment

The "post-hoc" analysis, examining pair-wise correlations between conditions over time (Figure 2:4), shows a considerable overlap in the *Armageddon* segment; in the *Die Hard* segment, even though the pair-wise correlations vary in parallel, the overall variation is greater than in the *Armageddon* segment, and the correlation between condition 0 and condition 1 (blue) is consistently higher than that of the other pairs.

In either the self-similarities diagram or the pair-wise correlations diagram the transition between the *Armageddon* and the *Die Hard* segment is easily discernible. As the segments have been chosen for differences in form, eye movement patterns are obviously connected to the form of the film stimuli.

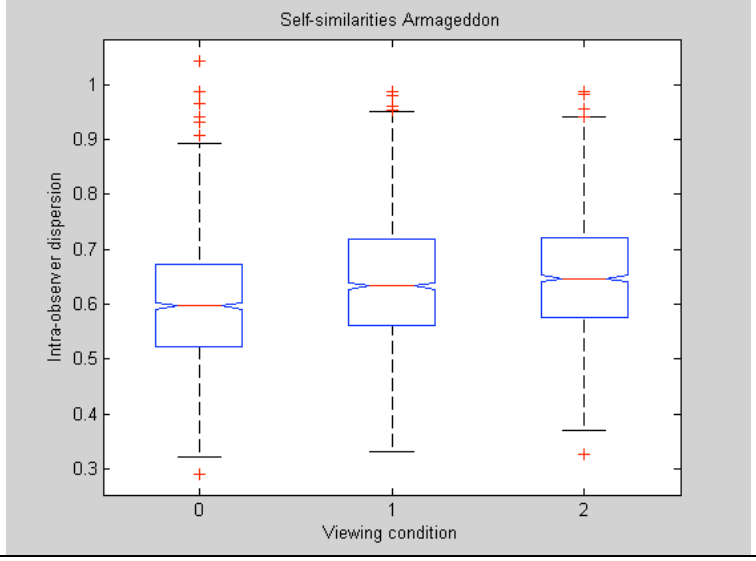


Figure 2:5: Gaze dispersion as a function of viewing condition (*Armageddon*). For condition 0, the no-task condition, the dispersion is significantly lower than for condition 1 and condition 2, indicating that it is the presence/absence of a task that makes a difference.

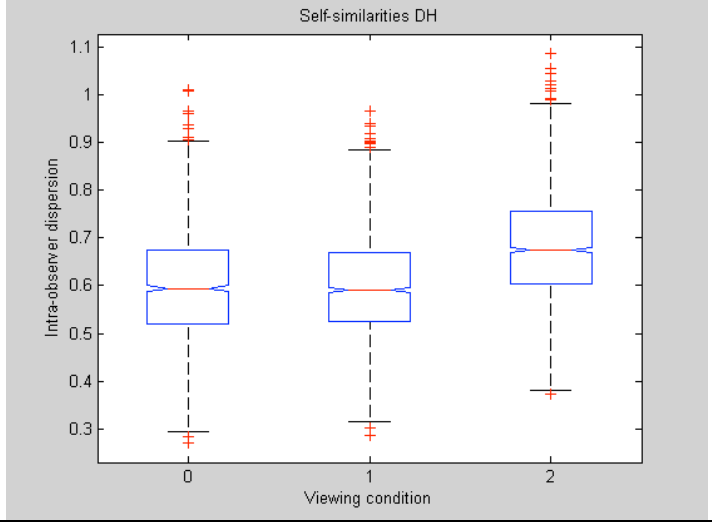


Figure 2:6: Gaze dispersion as a function of viewing condition (*Die Hard*). For condition 2, the woman evaluation condition, the dispersion is significantly higher than for condition 0 and condition 1, indicating that it is the content of the task that makes a difference.

A comparison the box plots of self-similarities for the two segments shows an interesting difference:
For *Armageddon*, the gaze dispersion in condition 0 is significantly lower than that in either condition 1 or condition 2, while condition 2 still yields a higher dispersion than condition 1. (Figure 2:5) (F=72,52)

For the *Die Hard* segment, on the other hand, dispersion under condition 2 is significantly greater than under 0 or 1, while that under condition 1 is slightly *lower* than under 0. (Figure 2:6) (F=303,98)
In other words, while the presence/absence of a task makes a major difference for variation in the subjects' fixation patterns in *Armageddon*, it is the nature of the task that matters most in *Die Hard*, where the specific task yields a considerably higher variation.

A graph of variations in self-similarities over time for each segment (Figures 2:13 and 2:14) agrees with the box plots: while gaze dispersion is consistently lower for condition 0 than the other two throughout the *Armageddon* segment, during the *Die Hard* segment it is consistently higher for condition 2.

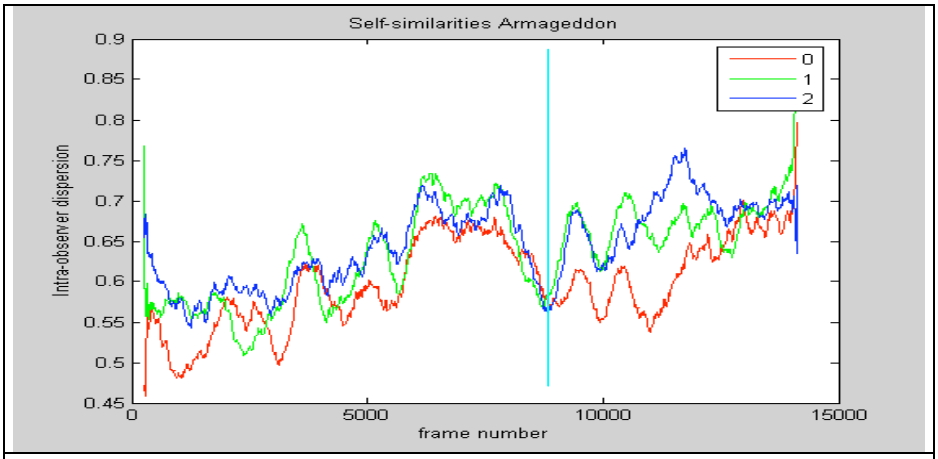


Figure 2.7: Gaze dispersion for each condition over time (*Armageddon*).
Cyan line marks a low point in variation within all groups, corresponding to a close-up of the protagonist having sacrificed himself for the good of others.

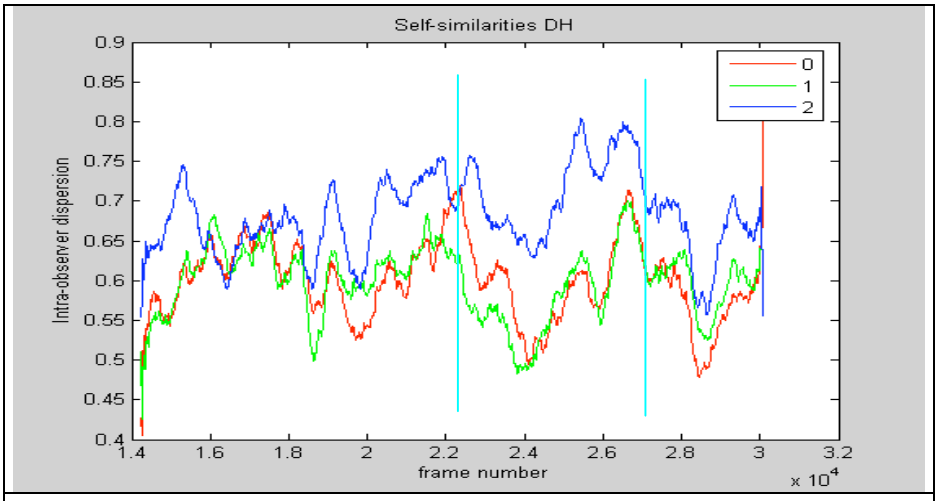


Figure 2:8: Gaze dispersion for each condition over time (*Die Hard*).
Cyan lines indicate some of the notable irregularities in the variation.

In the *Armageddon* segment, the pattern of self-similarities changes considerably at around frame 9000 (Figure 2:7, cyan line). In the *Die Hard* segment there is no similarly dramatic transition, but in the sequence delimited in the diagram (Figure 2:8, cyan lines) condition 2 exhibits an exceptionally high degree of dispersion. How do these patterns relate to patterns of correlation between conditions ("post-hoc" analysis)?

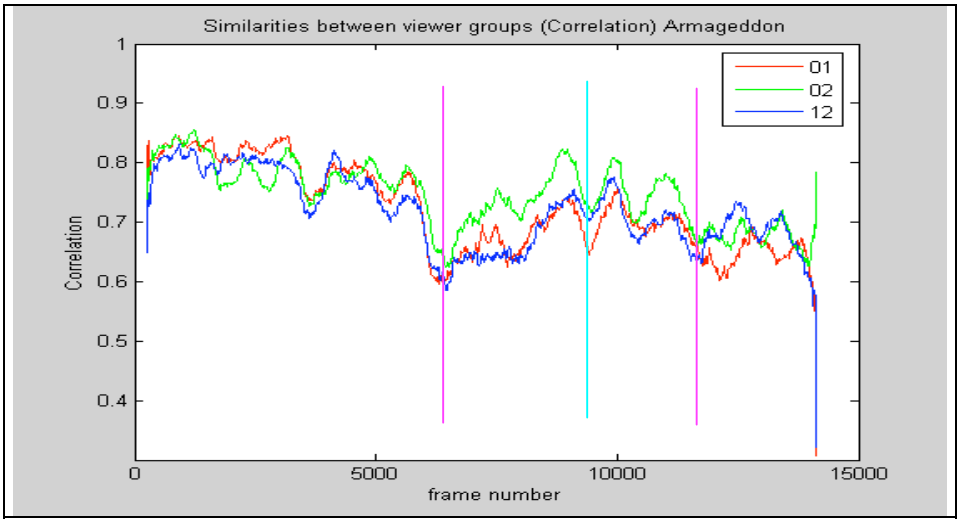


Figure 2:9: Correlation for pairs of conditions over time (*Armageddon*).
Cyan and magenta lines indicate notable irregularities in the correlation pattern.

The dip in gaze dispersion in the *Armageddon* segment does correspond to a dip in all three correlations (all the groups "agree to disagree"), but an even stronger dip occurs at about frame 6500 (Figure 2:9, leftmost magenta line). From there the correlation between conditions 0 and 2 goes up and remains relatively high until frame 1200 (Figure 2:9, rightmost magenta line).

Figure 2:10:

Armageddon (ca frame 6500):
low correlation between viewer groups

Red = condition 0
Green = condition 1
Blue = condition 2

On screen, the frames around 6500 correspond to the moment where the hero's daughter's fiancé has drawn the short straw to die saving the world and the hero first contemplates saving him. (Figure 2:10) The scene is preceded by a transition from on-screen pyrotechnics to dialogue, accompanied by a drop in all pair-wise correlations. The dialogue between the hero and his daughter, relevant to the task in condition 2, does not begin until about frame 1200. As Figure 2:9 shows, the pair **not** involving the task relevant groups displays a relatively lower correlation in this final sequence (red graph line). Figure 2:7 indicates that condition 2 also exhibits a higher dispersion from this point on.

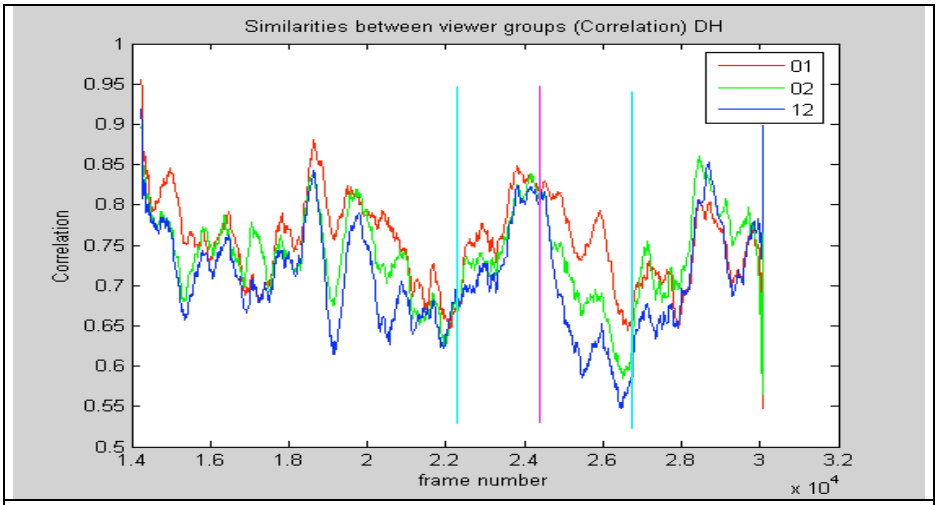


Figure 2:11: Correlation for pairs of conditions over time (*Die Hard*). Cyan and magenta lines indicate some notable "meeting points" in the correlation pattern, to be related to parts of the narrative.

For the *Die Hard* segment, the points that limit the period of particularly high dispersion in condition 2 correspond to "meeting points" in correlations: points where all pairs of conditions exhibit approximately the same degree of correlation (Figure 2:11, cyan lines). An even more interesting transition occurs at about frame 24500 (Figure 2:11, magenta line), after which the pairs containing condition 2 exhibit a considerably lower correlation than the remaining pair.

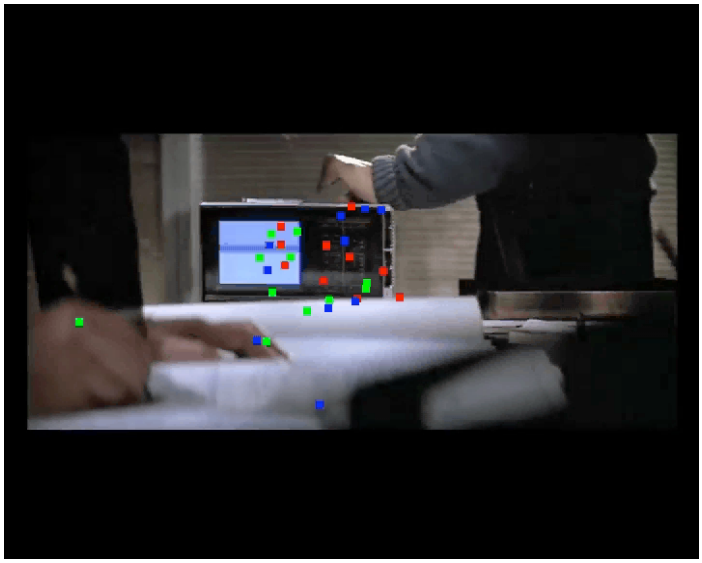


Figure 2:12:

The appearance of a TV set attracts unanimous visual attention. Low within-group dispersion, high between-groups correlation. Corresponds to magenta line in Figure 2:17 (*Die Hard*)

Red = condition 0
Green = condition 1
Blue = condition 2

The high point in correlations marked by the magenta line in Figure 2:11 corresponds to the appearance of a television monitor (Figure 2:12) just prior to the dialogue between the villain and the woman in *Die Hard*, that is the sequence relevant to the specific task (condition2). An ANOVA test of the self-similarities in the following sequence yields a value of F=138,05, indicating a high task impact. Figures 2:13, 2:14 and 2:15 show the corresponding box plot, graph of self-similarities over time, and pair-wise correlations between task groups.

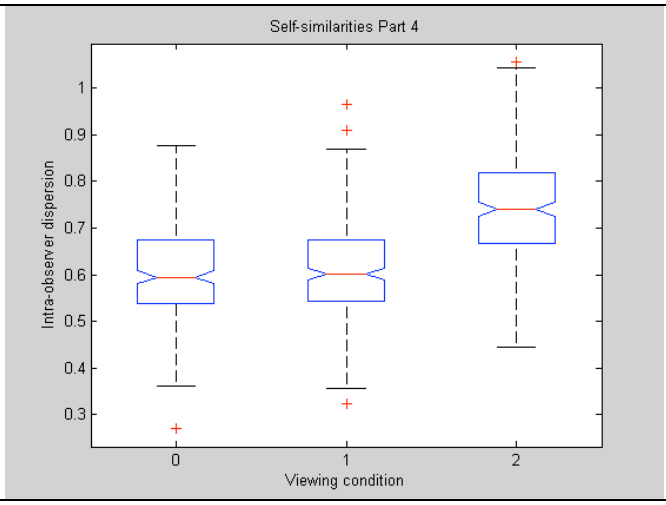


Figure 2:13: Box plot of self-similarities (variation within groups) for the woman dialogue sequence in *Die Hard*.

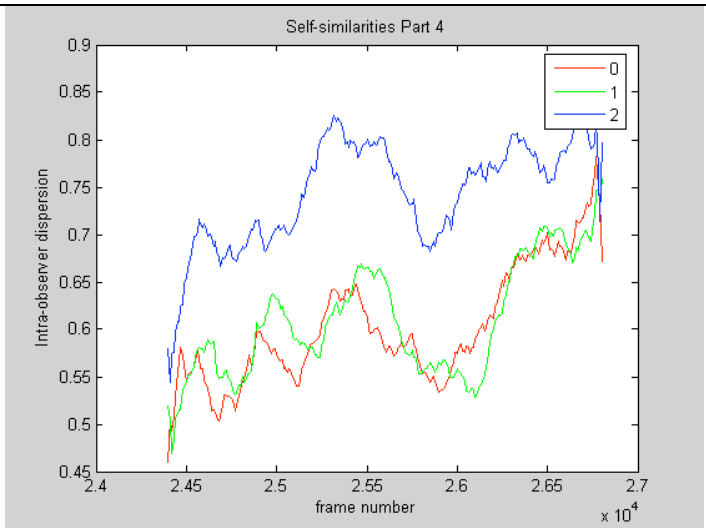


Figure 2:14: Variation in self-similarities over time for the woman dialogue sequence in *Die Hard*. Condition 2 has a significantly higher dispersion than conditions 0 and 1.

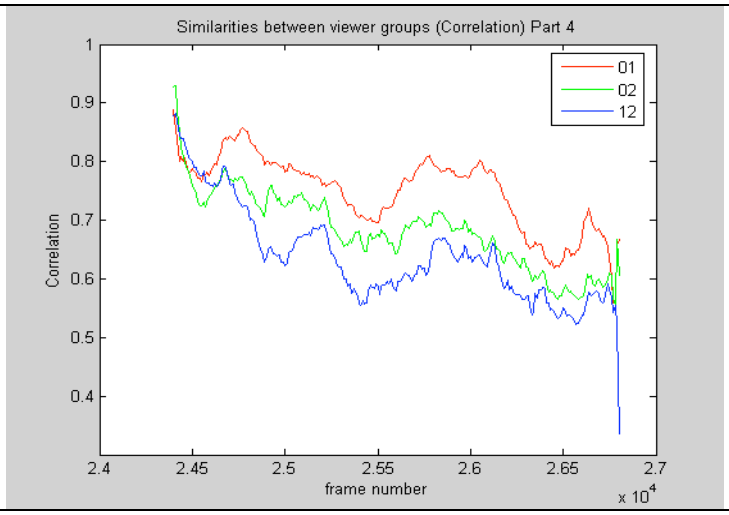


Figure 2:15: Correlation for pairs of conditions over time for the woman dialogue sequence in *Die Hard*. The pairs containing condition 2 show a consistently lower correlation than the 0-1 pair.

2.7.2 Film and character rating data

The following tables contain mean values of the film and character ratings, and rating differences between the two film sequences, under each of the three conditions:

Film ratings

Armageddon		*Die Hard*		Difference *Die Hard – Armageddon*
Overall mean	3,35	Overall mean	3,41	0,05
Condition 0	2,90	Condition 0	2,44	-0,46
Condition 1	3,71	Condition 1	3,91	0,20
Condition 2	3,37	Condition 2	3,82	0,45

Female character ratings

Armageddon		*Die Hard*		Difference *Die Hard – Armageddon*
Overall mean	3,85	Overall mean	5,93	2,08
Condition 0	3,87	Condition 0	6,15	2,28
Condition 1	3,58	Condition 1	5,81	2,23
Condition 2	4,15	Condition 2	5,85	1,70

The data seems to indicate that the general film ratings vary very little by condition, but that there is a noticeable variation in the female character ratings.

To examine if these differences are statistically significant, the data was depicted graphically in scatterplots and box plots, and one-way between-subjects ANOVA was performed. Although the scatterplots show that much of the data approximates normal distribution only poorly, ANOVA is considered robust and it is the usual inferential statistical procedure for an experiment involving three levels of one factor (Heiman 2006).

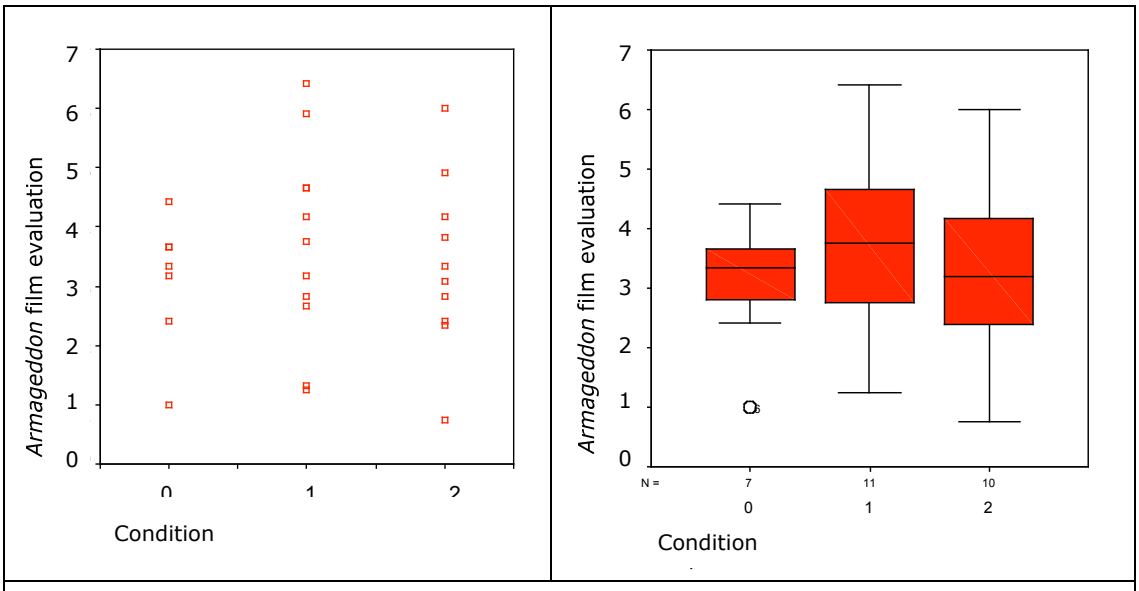


Figure 2:16: Scatterplot and box plot of the *Armageddon* film evaluation.
No statistical evidence that the three condition groups represent different populations.

For the *Armageddon* film evaluation, the scatterplot and box plot give no indication that there should be significant differences by condition. (Figure 2:16)

The ANOVA test showed no significant result, yielding F=0,389.

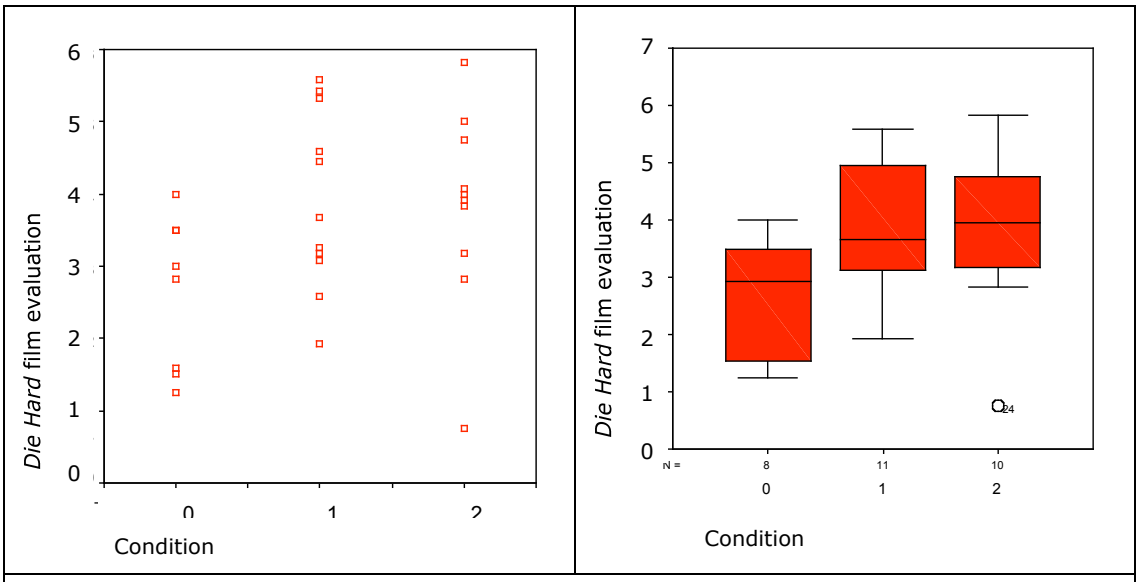


Figure 2:17: Scatterplot and box plot of the *Die Hard* film evaluation.
Statistically significant evidence that the three condition groups represent different populations.

The low rating of *Die Hard* by group 0 (no task) was significant, as the one extremely low rating in group 2 was treated as an outlier (Figure 2:17).

The ANOVA test returned the value F=2.798.

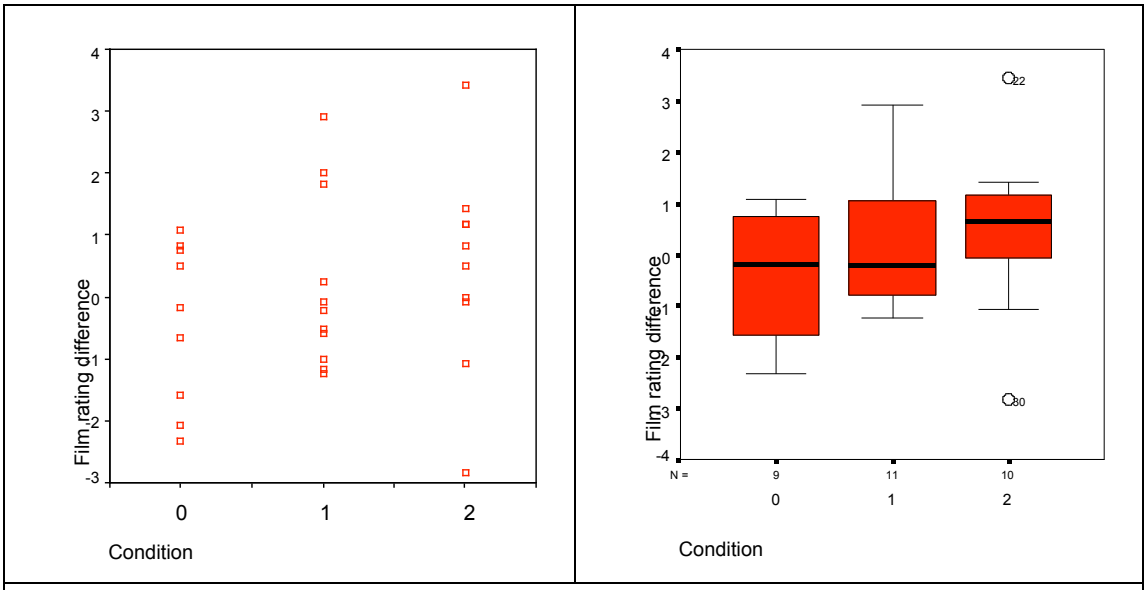


Figure 2:18: Scatterplot and box plot of the differences between the *Die Hard* and *Armageddon* film evaluation.
No statistical evidence that the three condition groups represent different populations.

Another non-significant result was that of differences in ratings between the two films, Rating (*Die Hard*) – Rating (*Armageddon*). The scores under condition 0 (no task) are particularly low, but not significantly so (Figure 2:18).

The ANOVA test yielded F=0,847.

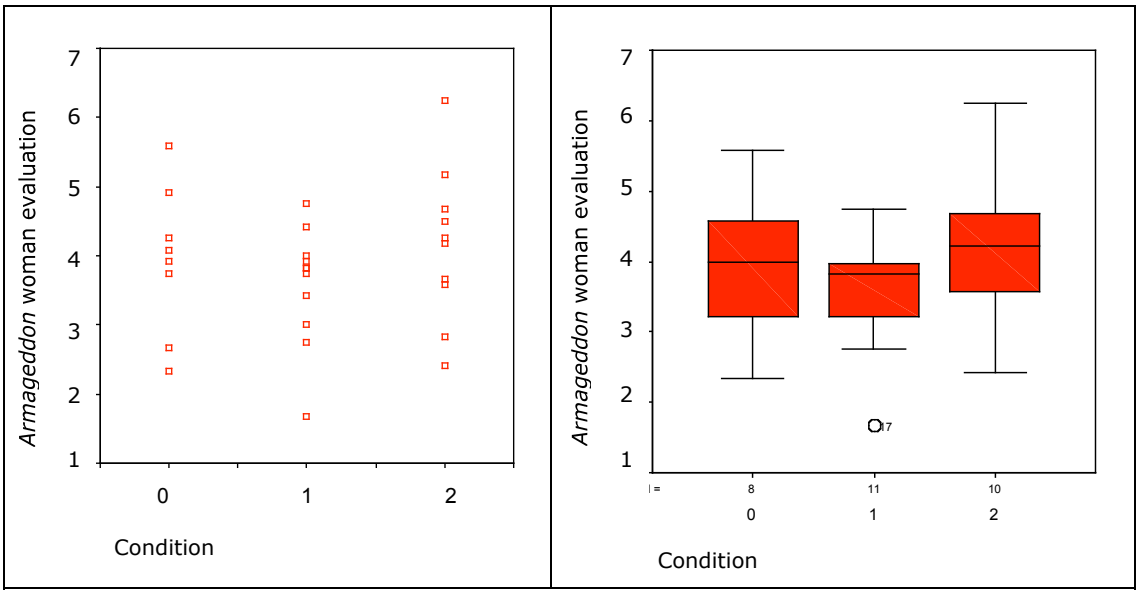


Figure 2:19: Scatterplot and box plot of the *Armageddon* woman character evaluation.
No statistical evidence that the three condition groups represent different populations.

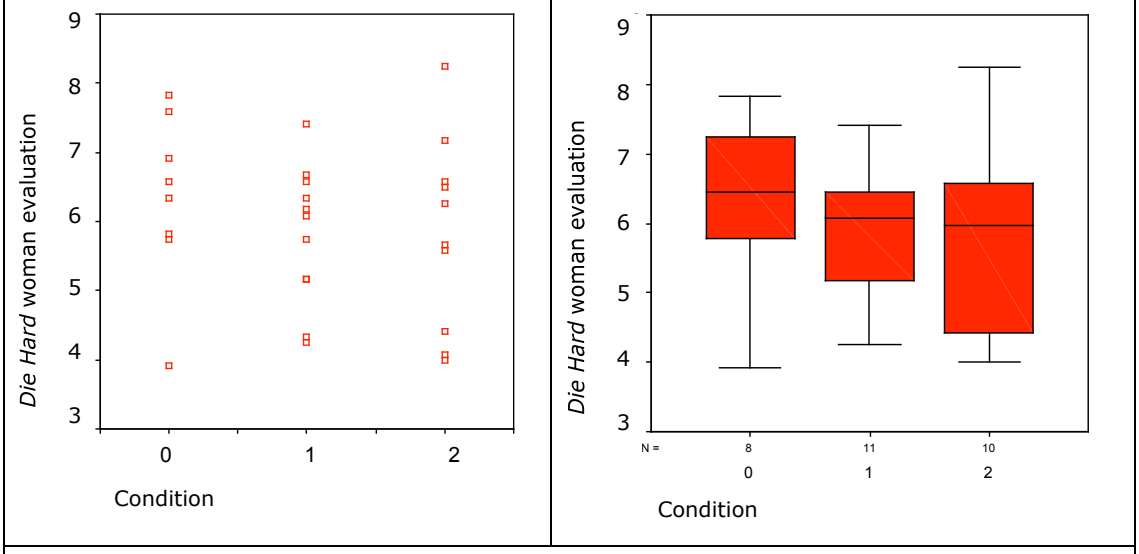


Figure 2:20: Scatterplot and box plot of the *Die Hard* woman character evaluation.
No statistical evidence that the three condition groups represent different populations.

Figures 2:19 and 2:20 illustrate two other non-significant results, showing no significant task impact on the evaluations of the woman characters in the two film sequences.
The ANOVA F values are F=0,875 (*Armageddon*) and F=0,527 (*Die Hard*).

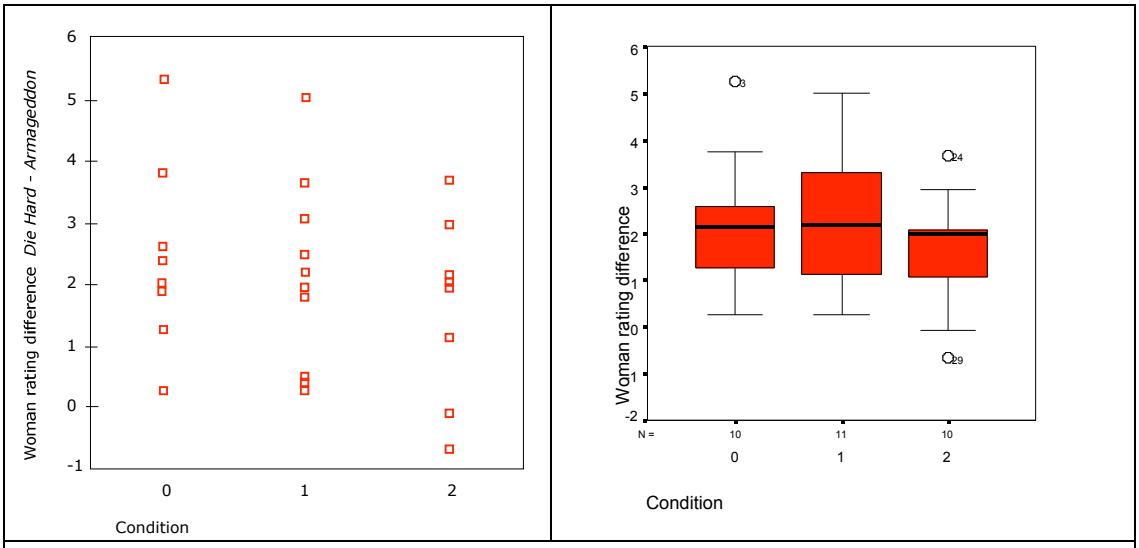


Figure 2:21: Scatterplot and box plot of the differences between the *Die Hard* and *Armageddon* female character evaluations.
No statistical evidence that the three condition groups represent different populations.

The comparisons between the female character evaluations seemed to yield an important result, where the relevant task group (condition 2) showed a smaller difference between the *Armageddon* and *Die Hard* ratings than the remaining groups.

With a value of F=0,538, the material shows no significant result under the ANOVA test, just as the box plot illustrates. (Figure 2:21)

We cannot even simply conclude that there is no task effect – in that case the ANOVA test should yield F=1. All the above analyses except the one for ratings of *Die Hard* give F values between 0 and 1, the highest being F=0,875. The error variance within each group is high compared to that between groups, indicating that either or both error estimates are inaccurate. (Heiman 2006)

2.7.3 Summary of results

The eye tracking data shows, most importantly, that the **form of the film stimulus** matters for the cognitive processing of its content, including the impact of task:

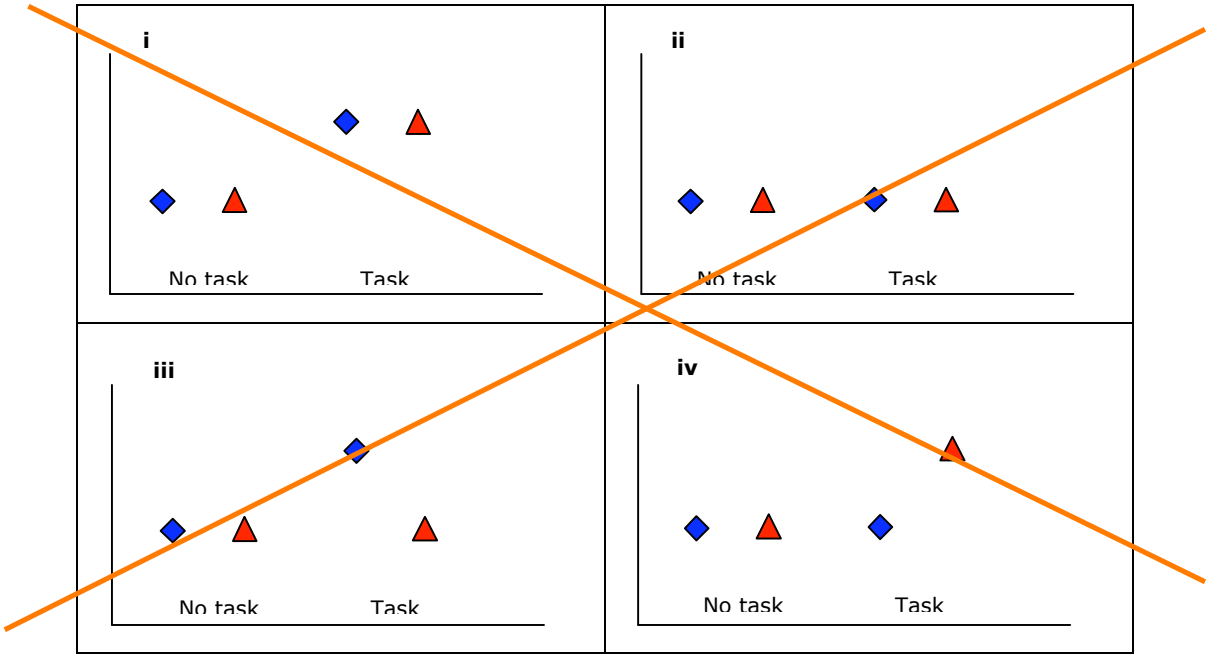
a) In the **attraction**-driven segment (*Armageddon*), there is little difference in gaze density functions among the conditions in the first part of the material. In the latter part, the fixation points are less dispersed for the no-task condition than for the other two, that is, spectators without a task "agree" significantly more on where to look.
This indicates that it is the **presence/absence of a task** that makes a difference.
The correlation data is more ambiguous, conditions 0 and 2 being highly correlated in about 1/3 of the segment.

b) In the **narrative**-driven segment (*Die Hard*), the gaze dispersion is higher under the specific task condition, that is, spectators with a specific task in mind "disagree" significantly more on where to look.
This indicates that it is the **specificity of the task**, not the simple presence or absence of one, that makes a difference.
The correlation data consistently displays conditions 0 and 1 as more highly correlated than the other pairs, even if the degrees of difference vary.

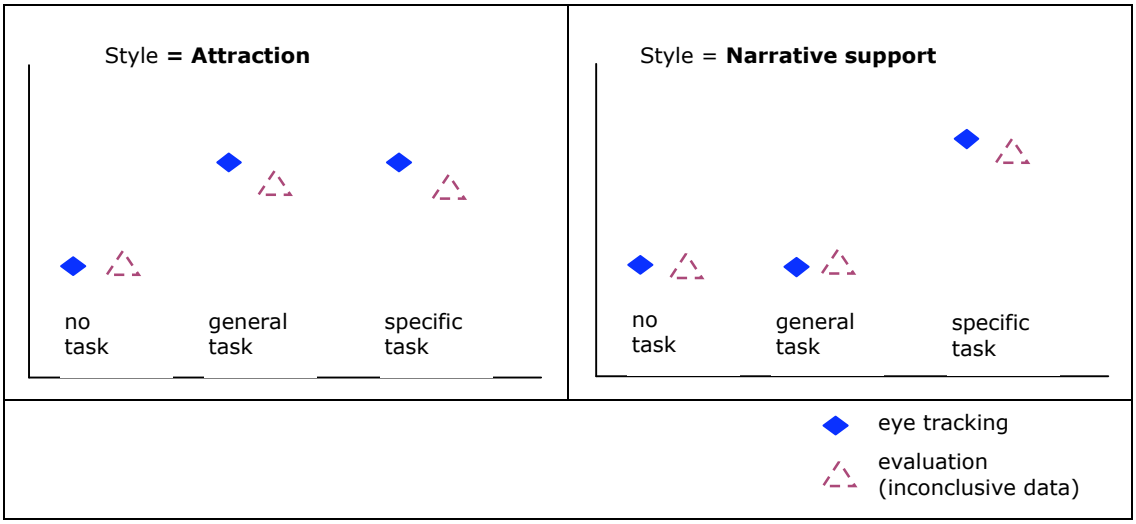
The character and film evaluation data yields no conclusive results, although the initial examination indicated otherwise.

3. General discussion

3.1 Form, content and visual attention



The assumptions about possible results in Section 1.2 were obviously oversimplified and none of the options i.-iv. apply. The data points towards a more complex picture.



Degrees of variation in fixation patterns vary between sequences with different visual properties. Salient highs/lows and differences in eye tracking data relate to both form and content. The transition between the two film sequences in the stimulus is easily traced in correlation diagrams. So are transitions within the sequences between "eye candy", fast-cut general action scenes, and dialogues significant for plot development. (see Figure 2:7, 2:8, 2:9 and 2:11)

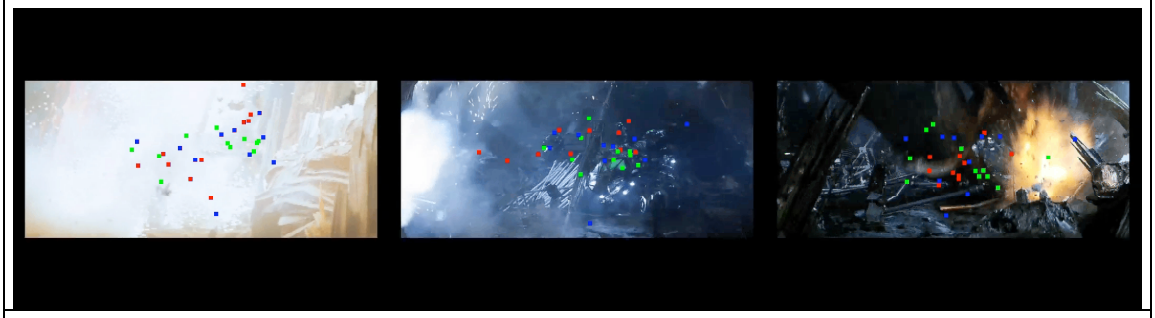


Figure 3.1: Visual excess in High Concept: movable camera, pyrotechnics and fast cutting collect fixations in the center of the image.
Armageddon, single shot examples with fixation points from the three task groups.

The sample images in Figure 3:1 are all taken from the sequence up to frame 5000 in the *Armageddon* segment. In such a highly dynamic display, with fast cuts, mobile camera and action, the test results show that there will be less variation (dispersion) within viewer groups and a higher correlation between them, allowing relatively little task-related overall variation in visual attention. (See Figures 2:9 and 2:11)

As the function of style in a High Concept film is to display attractive object and people, framing has comparatively little narrative function. We can expect important items to appear in the center of the picture, as Figure 3:2 illustrates.

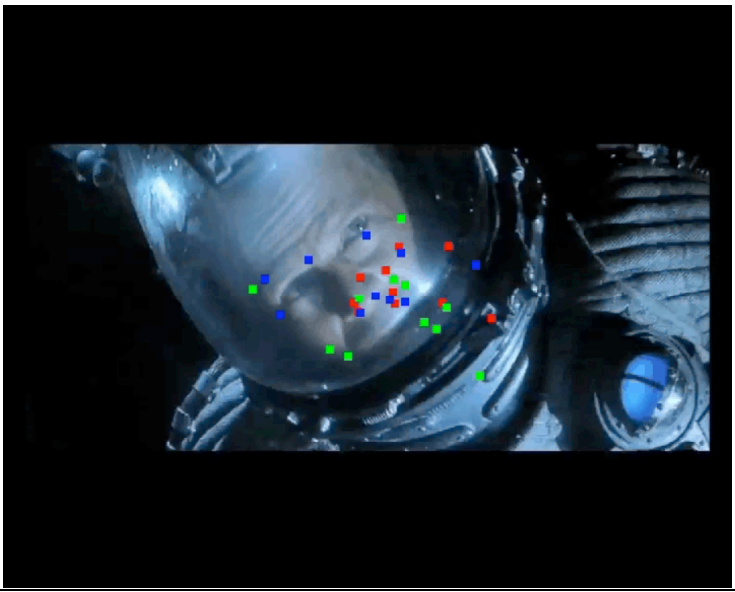


Figure 3:2:

Important object framed in predictable position in High Concept: the hero's face in the middle of the frame during the pyrotechnics sequence in the *Armageddon* segment.

In a narrative/dialogue sequence the speed of presentation is regulated by the characters' speech, their actions, or other narrative content, by means of methods such as camera angles adapted to turn-taking in dialogues, or match-on-action cutting. As framing has a narrative function, important items are not necessarily placed in the center of the picture. In terms of eye tracking, visual attention will be allocated to items with a potential dramatic impact, such as human faces and hands. (Figure 3:3) (Cf. Glenstrup & Engnell-Nielsen 1995, Yarbus 1967)

Figure 3:3: Faces and hands attract visual attention in a classical narrative. *Die Hard* shot sequence with fixation points from the three task groups.

In the present study such narrative sequences display a higher degree of variation within viewer groups and lower between-group correlation. This variation is partly task-related and can to some degree be accounted for by the impact of task priming.

The results indicate that the nature of the task sometimes matters more than simple absence/presence of a task. In a "classical" Hollywood film, where the form and style are wholly at the service of the story, priming by a specific task can be said to initiate a search that spectators perform in individual ways. This may explain that the group primed by the specific task in the experiment displays a greater internal variation and a smaller external correlation than the other task groups.

Thus the key issue isn't the presence or absence of a task. A search will be made if 1) the **nature of the task** demands it and 2) the **form of the stimulus** permits it.

Figure 3:4:

Searching for the woman? Dispersed fixation pattern in condition 2, under a specific task. Conditions 0 and 1 follow the regular pattern of faces attracting visual attention.

To sum up, there seems to be a strong connection between the ways information is presented in a film and the susceptibility of the cognitive process to task influence.

In a High Concept pyrotechnics sequence, the speed imposed on the spectator's perception causes fixations to remain in the center of screen. Important items are displayed centrally, so that framing is predictable and the presentation does not invite individual search patterns.

A classical Hollywood narrative has a higher density of information: cutting, framing and camera movement contribute often non-redundant information to the narrative. The fact that various options are open is conductive to visual search, and the relatively low-speed

presentation allows for alternative searches depending on task. (Figure 3:4) So the impact of task will be greater in a classical narrative sequence than in a High Concept attraction-driven one.

Lastly, researching the impact of task on film viewing may not seem particularly relevant to the endeavor of examining the cognitive processes involved in normal viewing of entertainment movies. But, as Section 2.3 describes, the tasks were formulated so as to approximate evaluations that spectators might be expected to engage in when discussing a film. Documenting and analyzing casual conversations about films would be a research project in itself. In an explorative context such as that of the present study, documented viewer comments may serve as a fair approximation. Browsing through a number of film message boards and viewer comments on the Internet Movie Database (www.imdb.com) shows that wanting/not wanting to see more of a film, and evaluating characters, are in fact very frequent remarks. People may be impatient for a sequel, or dreading it, or they may express opinions about specific characters' believability, intelligence, sexual attraction and the like. So it is quite plausible to assume that a spectator may take part in a conversation along these lines before watching a movie, and that such a conversation may function as a priming "task", influencing the spectator's ways of processing the content of the film.

3.2 Evaluation study: the need for a broader picture

From a first look at the simple arithmetic means of the film and character ratings, as given in Section 2.7.2, it seems easy to conclude that:

1) there are only minor differences in the ratings of the films, that is evaluations corresponding to condition 1 ("Decide if you would be interested in seeing the films that the sequences came from."). The ratings do not vary spectacularly by condition, and the rating differences between the two films are negligible. The rating made by the relevant condition group does not deviate noticeably.

2) the female character evaluations, corresponding to condition 2 ("For each sequence, form an opinion about the woman who takes part in the dialogue.") do not vary remarkably by condition, and the rating by the relevant group doesn't deviate. The rating differences between the two films, however, are noticeable: the participants rate the *Die Hard* female character by about two points higher than the *Armageddon* character, that is by 1/5 of the maximum possible rating. The smallest rating difference is displayed by the task relevant group, Condition 2.

The obvious question was whether these initial impressions would stand up to proper statistical scrutiny. As the results in 2.7.2 show, they did not. The statistics indicated inaccurate error estimates. An examination of the scatterplots disclosed that the data displayed only a crude approximation of normal distribution, not the kind that would be expected under a good experiment design.

The construction of valid and reliable tests in social sciences is an entire discipline in itself (Passer & Smith 2003, Wedin & Sandell 1999, Kline 1993). Thus it would have been extremely fortunate if a hastily designed questionnaire submitted to a small number of subjects, as was the case in the present study, had yielded significant results.

Like many psychological tests, a film evaluation test uses the participants' subjective ratings to assess the movie's effect on them. Concerns about reliability and validity are relevant in all testing, and are routinely discussed in contexts of psychological research. Based on Kline's (1993) guidelines for psychological testing, and leaving aside the technical aspects of his book, the film and character rating questionnaires in this study may be criticized on both counts, poor validity and poor reliability.

– If *validity* is said to mean that "the test measures what it is supposed to measure", the present test may have measured little more than the participants' interpretation of the vocabulary of

the test items, instead of measuring their evaluation of the films and characters. Even with a native-like command of the language, some of the terminology may vary in interpretation depending on the participant's background. A case in point is the value pair "worthless – valuable". In the context of film studies and criticism, these are often employed in an aesthetic and/or ethical sense. A person not used to that context may well interpret them as referring to a film's production value and an "expensive" look. A closer examination of the questionnaires indicates that this may indeed have been the case for some of the subjects.

– *Reliability* refers to "how well a test measures what it measures", and, since the ANOVA test showed poor error estimates, the present test did not do so particularly well. Evaluations of spectators' general impressions of a film or of its characters are by necessity difficult to define operationally; the more important is it that care be taken in selecting test items and designing the tests.

Each of the questionnaires had only 12 questions, all addressing different aspects of the variable being tested. There was no redundancy, no alternative formulations, hence no way of checking for internal consistency. Kline considers10 items as the minimum in a test, some 20 items being preferable. In view of the many-faceted variables tested in this study, using more than12 items would have been strongly advisable.

The 12 test items were postulated by the researcher, not selected from a larger set with the aid of a pilot study using a large and representative sample of the population. The only pilot study in the present experiment was conducted on a small sample entirely composed of first-year film students, thus pointless for preventing the terminology problem.

In the attempts to construct a cognitive model of the film-viewing process it is necessary to relate experiments concerning local cognitive mechanisms to the spectators' experience of the film, since this is the final "product" of the viewing process. But in order to examine this experience in a meaningful way it is crucial to develop valid tests that yield reliable data, and, whenever possible, to complement these tests with methods not relying on introspection, such as physiological measurements (heart rate, skin conductivity response, ERP).

3.3 Decisions under constraints: a case for simple heuristics

A suitable form of data presentation, such as that of a classical narrative, may be a necessary condition for a task-related allocation of visual attention. But it does not explain **why** such allocation actually takes place. The question remains: why would visual attention, as indicated by gaze density functions, be more dispersed under condition 2, with the more specific task, than under conditions calling for either neutral viewing or a general evaluation?

Figure 3:5:

The frugal heuristic "Look at the face!" is modified by the demands of specific task.
Condition 2 calls for an evaluation of the woman, and she is allocated some visual attention even when her back is turned.

Die Hard, the woman dialogue sequence.

Joseph Anderson repeatedly points out that humans process film information by the same perceptive and cognitive tools as other information that reaches them – or that they select – from their physical environment. (Anderson 1996) What general theory of human information processing may help account for the specific condition effect? And how can a film theory tie in with such an account?

There is no single, unified cognitive film theory. Early cognitive film theorists such as Bordwell (1987) and Branigan (1992) focus on the process by which spectators construct a coherent narrative. Anderson (1996) examines the cognitive processing of film in terms of evolutionary psychology, Torben Grodal (1997) refines the model by adding the element of emotion, in agreement with the findings of neuroscience that emotions are of great importance for decision-making. (Damasio1994) But all varieties of cognitive film theory share certain common features:

The basic assumption is that the spectator generates a mental model of the film's content based on the information provided by the film. One important distinction, that film theory shares with theories of other narrative media, such as literature, drama, comics or ballet, is that between *story* and *plot*. (Alternative terminologies are used depending on the tradition that each specific film theorist works in.) The term *story* refers to the film's content ordered in a natural causal and temporal sequence, while the term *plot* refers to the content as it is presented in the film. Consider the following example:

> *Story*: Rodney comes from Boston to Longhorn City wanting to be a cowboy. At a church picnic he meets Jennifer and falls in love with her. Her father, a tough rancher, despises greenhorns. A band of cattle thieves is active around Longhorn City. Rodney uncovers them and is at first humiliated by them. Then he outsmarts and liquidates them, saving the rancher's prize herd. The rancher gives him Jennifer in marriage.

The opening sequences of two different plots presenting the story:

> *Jennifer's tale*: Jennifer, middle-aged, in the parlour by the piano, tells her daughters that Longhorn City was not always the civilized place that it is now. When Rod once came from Boston, his life in the West was not easy. She first met him at a church picnic…

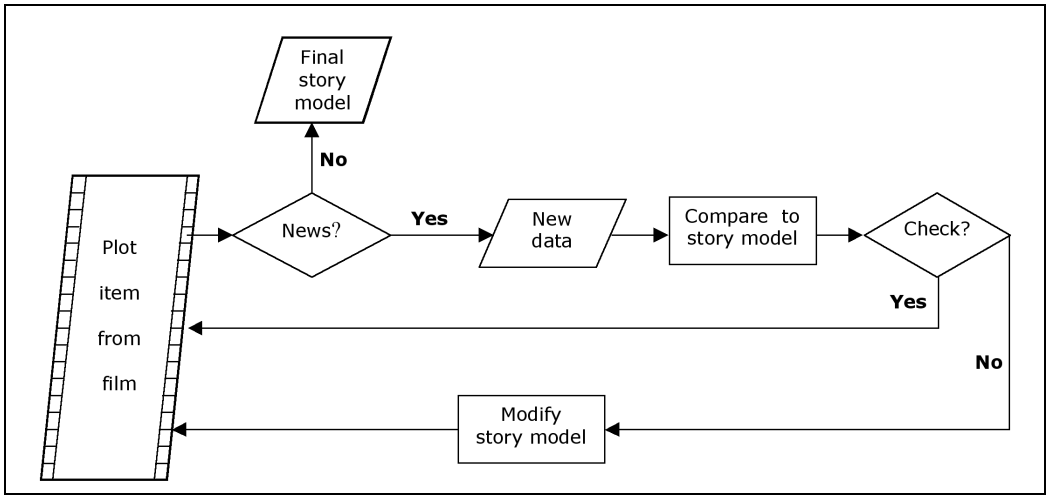
> *The cattle thieves' tale*: Tom, Harry and Sancho come riding into town to celebrate their latest cattle raid. In the saloon they run into a little Eastern runt who starts preaching morals at them, so they shoot his pants down. It turns out that he's been seen making eyes at Jennifer, who Tom sees as his own property…

Who is shown on screen, and what, if anything, the characters say, are not the only ways of presenting story information. Conventionally romantic music communicates to the spectator that Rod and Jennifer have fallen in love; framing Jennifer in the middle of an empty landscape can illustrate her loneliness, a low camera angle displays the cattle thieves as big and threatening.

The basic concept of cognitive film theory is that the spectator uses such plot information to generate a mental model of the story. This is a dynamic process. The spectator enters it equipped with a set of expectations, values and experiences that will influence her processing of the film information. Early on in the film she generates a rough mental model of the

content, and then compares each new plot element to it. If it agrees, the model will be maintained, if not, the model will be modified, and so on, for all information presented gradually during the film.

The following is a very rough schematic representation of the process from plot elements to a mental model of the story:



Individual spectators have varying backgrounds, attitudes and habits, so they collect and process film information in idiosyncratic ways, but there are significant tendencies over groups. A mass-market genre film, largely relying on widespread conventions, will generate very similar cognitive responses in major segments of its target population. (An informal experiment that anyone can try at home is watching the opening sequences of a few genre films, such as westerns, romantic comedies or action films. Very often it is possible to get a roughly correct idea of the entire story from about the first seven minutes.) Delaying plot elements is a common way of creating suspense in thrillers and horrors. – Less formulaic films exploit the story-plot relationship by providing contradictory plot elements and ambiguous story clues, and forcing more frequent modifications of the spectator's mental model. Such films are also likely to facilitate generation of more varied mental models in individuals in the same target population, that is, viewers may disagree as to what the movie "is about".

What makes cognitive film theory particularly interesting in the context of the present study is the fact that it models film viewing in terms of **choices** and **decision making**: should a new plot element modify the mental model of the story, and if so, in what direction? Such decisions are not conscious, that is, people do not "go to films to make inferences" (Stam 2000), but can be seen as analogous with the extremely rapid and non-conscious choices that speakers make when selecting words and articulation settings in casual speech in their native language.

The film spectator's access to information is temporally **constrained** in ways that a book reader's or a still picture viewer's is not, resembling the temporal constraints in natural situations. Compared to a natural situation, however, information is also spatially constrained by a film picture being two-dimensional and the angle of observation, distance and framing being determined by the camera. As the constraint problem resembles that of decision-making in everyday life, an approach based on simple heuristics may prove fruitful.

The theory of **simple/frugal heuristics** has been formulated in response to the riddle of fast and most often accurate decision-making in real-world situations. (Gigerenzer et al. 1999) In everyday life, humans constantly make choices and decisions under conditions that do not allow for a complete information search and a conscious weighing of all available options.

In a simple heuristics context, we may formulate the hypothesis that a free-viewing situation resembles everyday choices, made on the basis of fast, simple "rules of thumb". A lot of data under constrained access has to be processed in a limited time with a limited computational capacity, to provide a coherent whole, a narrative that makes sense to the spectator.

A general task calls for fast overall processing by means of cognitive shortcuts and the spectator will rely on frugal heuristics. A visual search will focus on possible input for such heuristics, and will generate similar fixation patterns in most spectators, that is, the dispersion within a group will be low.

A specific task calls for a more systematic search for clues relating to a restricted aspect of a scene, and will yield a more varied visual search with dispersed fixation patterns. (Figure 3:5)

A more complete task-related rating test than the one in this study might lend support to a related hypothesis about evaluations: they should be less categorical in a specific task context than in that of free viewing or of a general task. If we assume that fast and frugal heuristics are used in free viewing, we may speculate that such reasoning will lead to categorical evaluations. Under a specific task, an investment in specific data search is made, the input to everyday simple heuristics is given less importance, hence a less categorical evaluation.

Film spectators are aware of **multiple layers of reality**: they evaluate the narrative for truthfulness and they relate emotionally to the characters, while they know that a film display consists of apparently moving colored spots on a screen, and that those colored spots depict actors dressed in costumes and speaking and behaving in ways invented by a script writer and a director. How would a model based on simple heuristics handle this fact? – Simple heuristics are explicitly said to exploit environment structure (Gigerenzer & Todd 1999, introduction), and fiction film is an easily identifiable special environment. Thus the spectator may make use of some expectations and heuristics not applicable to other situations: a person wearing black in a traditional Western may be assumed to be a villain, while people do not readily assume the same in everyday life.

And finally, the whole foundation of understanding of a film narrative lies in the generation of **causal** chains: the film presents information in the form of a story, with its internal points of view and temporal order, and the spectator uses the information to generate the mental model of the plot, the temporal and causal chain of events that the story depicts. Hertwig & Todd (2003) point out that simple heuristics facilitate detection of co-variation, a basis for the recognition of causality. Thus it is plausible to include a theory of simple heuristics into a model of the cognitive processing of film information. The results of the present study lend at least the beginnings of empirical support to such a model.

Obviously, suggesting a simple heuristics model of the cognitive processing of fiction film poses more questions than it answers. One of the major problems of the theory of simple heuristics is its attractiveness and *prima facie* plausibility. It facilitates the invention of "Just-So-stories", seemingly elegant but not empirically falsifiable explanations of everyday phenomena. The use of eye tracking together with psychometric and physiological measures to explore various aspects of film viewing may, however, advance film research as an empirical testing ground for simple heuristics hypotheses.

Sammanfattning

Summary in Swedish

En åskådares upplevelse av en film blir till genom en invecklad kognitiv process.

Fysiskt sett är filmen en serie rutor med färgfläckar och tillhörande reproducerade ljud. Åskådarens sinnen tolkar bildserien och ljuden som en rörlig avbildning av en tredimensionell verklighet. Hon vet dessutom att denna ”verklighet” är fiktiv: de rörliga bilderna avbildar skådespelare i kostymer och smink, som följer ett uppdiktat manus. Ändå är åskådaren känslomässigt engagerad i handlingen och karaktärerna och upplever t.ex. tillfredsställelse när filmens ”bov” får sitt ”rättmätiga straff”.

Att empiriskt undersöka filmupplevelsen som en kognitiv process har hittills varit svårt, eftersom forskningen har fått lita till åskådarnas egna beskrivningar. Filmforskare har dessutom inte prioriterat utvecklingen av experimentella metoder inom ämnet, därför att filmvetenskapen av historiska skäl ligger inom en annan vetenskaplig tradition än den renodlat empiriska.
Idag har datorteknikens framsteg öppnat nya möjligheter för att genomföra studier av ögonrörelser på rörliga bilder och utan att kräva orörlighet av försökspersonen. På så sätt kan forskare få fram objektiva mått på vad som händer med åskådarnas uppmärksamhet när de tittar på film. Med hjälp av matematiska modeller framtagna för att förbättra tekniker för komprimering av videomaterial kan man analysera ögonrörelsedata och undersöka samspelet mellan filmens form och åskådarens uppmärksamhet. Eftersom filmen i sin tur är en välkontrollerbar sorts verklighet, kan denna typ av filmstudier tillföra ny kunskap om människans kognition i allmänhet.

Två faktorer som man kan påverka i ett experiment med film och åskådare är a) egenskaperna hos de filmsekvenser som åskådaren tittar på, och b) de förutsättningar som åskådaren har ”i bakhuvudet”. Denna uppsats beskriver ett experiment där varje deltagare tittade på två filmsekvenser, en ur en traditionell berättande film (*Die Hard 1*, 1988) och en ur en film präglad av enkelt berättande och ett överflöd av visuella effekter (*Armageddon*, 1998). Inför filmtittandet fick personen dessutom en av tre möjliga uppgifter: 0, att titta neutralt; 1, att avgöra om hon/han skulle vilja se resten av filmen; 2, att bilda sig en åsikt om kvinnan i vardera sekvensen. Under filmens gång spelades deltagarnas ögonrörelser in. Efteråt fick varje deltagare besvara frågor om sina intryck av filmen och av de kvinnliga karaktärerna.

För *Armageddon*-sekvensen visade experimentet att deltagarnas ögonrörelsemönster var ganska likartade. Särskilt i avsnitt med mycket snabb klippning och ljusstarka flygande föremål samlades fixeringarna i mitten av bildrutan. Totalt sett märktes ändå en viss skillnad i ögonrörelserna mellan de deltagare som hade fått en uppgift och dem som skulle titta neutralt.
I *Die Hard* -sekvensen illustrerar kameraarbetet och klippningen dialogens rytm och karaktärernas egenskaper och beteende. I synnerhet de försökspersoner som skulle bilda sig en åsikt om kvinnan var inbördes ”oeniga” om de rätta ställena att titta på, och deras ögonrörelsemönster var påtagligt mera spridda än under de andra två villkoren.

Experimentet visade att det inte går att generalisera om hur filmtittandet påverkas av den uppgift som åskådaren får. Uppgiften får full genomslagskraft bara om filmens form tillåter att åskådare söker av bilderna på olika sätt. Då leder en specifikt formulerad uppgift till att enskilda åskådare söker information individuellt. I de övriga fallen tycks åskådarna istället använda snabba, generella beslutsregler för att bilda sig en helhetsuppfattning utifrån ett fåtal enkla ledtrådar.

Appendixes

Appendix I. Background and film habits questionnaire

1. Age:

2. Gender: F / M / HBT

3. In what language(s) do you communicate the most easily:

Swedish / English / Other:

4. Education in aesthetic subject(s) (film studies/art/literature/music)?

How long (in total)? none / one semester / 1-2 years / more

Subject(s):

5. Are you or have you been a member/active participant in a film art club (Cinemathèque, Film Studio)?

never / one semester / 1-2 years / more

6. Do you actively seek out independent, non-commercial or otherwise alternative/quality films?

never / occasionally / often / (almost) always

7. How do you rate your interest in film? (more than one option if applicable)

never go to the cinema / catch a movie occasionally / see a lot of film for entertainment / keep in touch with what's new in the big cinemas / keep in touch with what's new in independent film / discerning film viewer with special interests / enjoy discussing film style and interpretation

8. Which of the following mainstream movies have you seen?
For the ones you have seen, give your personal opinion of it on a 0-5 scale (0=terrible, 5=very good) haven't seen/don't remember/0/1/2/3/4/5

Independence Day / Pretty Woman / Jaws / You've Got Mail / Die Hard / Love Story / Space Cowboys / Bridget Jones' Diary / Armageddon / Legally Blonde

Appendix IIa. Film rating questionnaire

Based on the sequence that you have seen, do you rate the **film** as:											
worth seeing											not worth seeing
dated											modern
dumb											smart
deep											shallow
boring											exciting
touching											indifferent
original											formulaic
ugly											beautiful
artistic											mass-produced
worthless											valuable
propagandistic											objective
interesting											trivial

Appendix IIb. Character rating questionnaire

Based on the sequence that you have seen, do you rate the **woman character** as:											
active											passive
repulsive											attractive
dumb											smart
deep											shallow
boring											interesting
strong											weak
good											evil
ugly											beautiful
in professional role											in private role
pitiable											admirable
dominant											submissive
dependent											independent

References

Books and articles

Anderson, Joseph D.: *The Reality of Illusion; An ecological approach to cognitive film theory*, Southern Illinois U. Press, 1996

Blackmon. Theodore T. et al.: 'Eye Movements While Viewing Dynamic and Static Stimuli', Proceedings - *19th International Conference - IEEE/EMBS* Oct. 30 - Nov. 2, 1997 Chicago, IL. USA

Bordwell, David & Thompson, K.: *Film Art: An Introduction with Film Viewer's Guide and Tutorial*, McGraw-Hill, 2003

Bordwell, David & Staiger, J.: *The Classical Hollywood Cinema: Film Style and Mode of Production to 1960*, Routledge, 1988

Bordwell, David: *Narration in the Fiction Film*, Routledge, 1987

Branigan, Edward: *Narrative Comprehension and Film*, Routledge, 1992

Damasio, Antonio: *Descartes' Error: Emotion, Reason and the Human Brain*, Avon Books, 1994

Deubel, Heiner & Schneider, Werner X.: "Saccade Target Selection and Object Recognition: Evidence for a Common Attentional Mechanism", *Vision Research*, Vol. 36, No. 12, pp. 1827-1837,1996

Duchowski, Andrew T.: "A Breadth-First Survey of Eye Tracking Applications", *Behavior Research Methods, Instruments, and Computers*, 2002:1

Elsaesser, Thomas & Buckland, W.: *Studying Contemporary American Film*, Arnold Publishers, 2002

Gigerenzer, Gerd, Todd, P.M. & The ABC Research Group: *Simple Heuristics That Make Us Smart*, Oxford U. Press, 1999

Goldstein, R.B., Peli, E., Lerner, S. & Luo, G.: "Eye Movements while Watching a Video: Comparisons Across Viewer Groups", Vision Science Society 2004

Grodal, Torben, *Moving Pictures*, Oxford U. Press, 1997

Hedling, Erik: "Independence Day" i Hedling, E. & Andersson, L.G.: *Filmanalys*, Lund 1999

Heiman, Gary W: *Basic statistics for the behavioral sciences*, Houghton Mifflin 2006

Henderson, John M.: "Human gaze control during real-world scene perception", *Trends in Cognitive Sciences*, vol.7 no.11, November 2004

Hertwig, R & Todd, P.M.: "More is not Always Better: the Benefits of Cognitive Limits", *Thinking: Psychological Perspectives on Reasoning, Judgment and Decision Making* , Edit. David Hardman and Laura Macchi, John Wiley & Sons, Ltd., 2003

Johnson, Addie & Proctor, R.W.: *Attention – Theory and Practice*, Sage Publications, 2004

Kline, Paul: *The Handbook of Psychological Testing*, Routledge, 1993

Mulvey, Laura: "Visual Pleasure and Narrative Cinema", *Screen 16*, 1975

May, Jon, Dean, M.P. & Barnard, P.J.: "Using Film Cutting Techniques in Interface Design"; *Human–Computer Interaction*, 2003, Volume 18, pp. 325–372

Nyström, Marcus, Holmqvist, K. & Novak, M.: "A quality function for off-line foveated video compression"; Lund University 2005

Stam, Robert, *Film Theory. An Introduction*, Malden, Mass. 2000

Yarbus, A.L.: *Eye Movements and Vision*, New York: Plenum, 1967

d'Ydevalle, Géry, Praet, C. Verfaillie, K. & Van Rensbergen, J.: "Watching Subtitled Television. Automatic Reading Behavior", *Communication Research*, vol. 18 No. 5, October 1991

Wolfe, Jeremy M. & Horowitz, T. S.: "What attributes guide the deployment of visual attention and how do they do it?", *Nature Reviews / Neuroscience*, vol. 5, June 2004

Wyatt, Justin: High Concept: *Movies and Marketing in Hollywood*, U. of Texas Press 1994

Unpublished and Internet material

Björklund, Elisabet & Jansson, L.: "Möjligheten till tvärvetenskapligt samarbete – om vissa problem inom den filmvetenskapliga forskningen såsom de yttrar sig i behandlingen av åskådare", unpublished third-term paper in film studies, Lund, 2006

Ebert, Roger: "Armageddon", http://www.suntimes.com/ebert/ebert_reviews/1998/07/070101.html (originally published in The Chicago Sun Times)

Germeys, Filip: "Film Perception", 200? (undated page, accessed and printed April 2006) http://ppw.kuleuven.be/labexppsy/top/filipweb/research2.htm

Glenstrup, Arne John & Engell-Nielsen, T.: "Psychological and Physiological Aspects of Eye Gaze Tracking", Eye Controlled Media: Present and Future State, Institute of Computer Science, University of Copenhagen, 1995 http://www.diku.dk/~panic/eyegaze/article.html

The Internet Movie Database, http://www.imdb.com

Persson, Tomas: "Tre sätt att se på bilder hos människor och djur", LUCS mini conference, Lund, December 19th, 2005

Spanne, Janna G.: "Bigger, badder, louder. High Concept och ideologi – en explorativ studie", unpublished examination paper in film studies, Lund, 2004

About this paper:

Master's Thesis in Cognitive Science, May 2006

Lund University, Department of Philosophy

Author: Janna G. Spanne
janna.geggus-spanne.675@student.lu.se

Advisor: Kenneth Holmqvist, LUCS / LU Humanities Laboratory

Films

Die Hard

20th Century Fox/ Gordon Company/Silver Pictures, 1988

Director: John McTiernan
Written by: Jeb Stuart, Steven E. de Souza (screenplay), Roderick Thorp (novel)
Producers: Charles Gordon. Lawrence Gordon, Beau Marks, Joel Silver
Original music: Michael Kamen et al.
Cinematography: Jan de Bont
Film editing: John F. Link, Frank J. Urioste

Cast: Bruce Willis (John McClane), Reginald VelJohnson (Sgt. Al Powell), Bonnie Bedelia (Holly Gennero McClane), Alan Rickman (Hans Gruber) et al.

Armageddon

Touchstone Pictures/Jerry Bruckheimer Films/Valhalla Motion Pictures, 1998

Director: Michael Bay
Written by: Robert Roy Pool, Jonathan Hensleigh (story), Tony Gilroy, Shane Salerno (adaptation), Jonathan Hensleigh, J.J. Abrams (screenplay)
Producers: Kenny Bates, Michael Bay, Jerry Bruckheimer, Jonathan Hensleigh et al.
Original music: Jack Blades (songs "What Kind of Love Are You On" and "Remember Me"), Trevor Rabin, Steven Tyler (song "What Kind of Love Are You On"), Diane Warren (song "I Do not Want To Miss A Thing")
Cinematography: John Schwartzman
Film editing: Mark Goldblatt, Chris Lebenzon, Glen Scantlebury

Cast: Bruce Willis(Harry S. Stamper), Billy Bob Thornton (Dan Truman, NASA Administrator), Ben Affleck (A.J. Frost), Liv Tyler (Grace Stamper), Will Patton (Charles 'Chick' Chapple), Steve Buscemi (Rockhound), Owen Wilson (Oscar Choi, Geologist), Michael Clarke Duncan (Jayotis 'Bear' Kurleenbear), Peter Stormare (Lev Andropov, Russian Cosmonaut) et al.