An Attentional Theory of Continuity Editing

Tim J. Smith

Doctor of Philosophy

Institute for Communicating and Collaborative Systems

School of Informatics

University of Edinburgh

2005

Abstract

The intention of most film editing is to create the impression of continuous action ("continuity") by presenting discontinuous visual information. The techniques used to achieve this, the continuity editing rules, are well established yet there exists no understanding of their cognitive foundations. This thesis attempts to correct this oversight by proposing that "continuity" is actually what perceptual and developmental psychologists refer to as *existence constancy* (Michotte, 1955): "*the experience that objects persist through space and time despite the fact that their presence in the visual field may be discontinuous"* (Butterworth, 1991). The main conclusion of this thesis is that continuity editing ensures *existence constancy* by creating conditions under which a) the visual disruption created by the cut does not capture attention, b) *existence constancy* is assumed, and c) expectations associated with *existence constancy* are accommodated after the cut.

Continuity editing rules are shown to identify natural periods of attention withdrawal that can be used to hide cuts. A reaction time study shows that one such period, a saccadic eye movement, occurs when an object is occluded by the screen edge. This occlusion has the potential to create *existence constancy* across the cut. After the cut, the object only has to appear when and where it is expected for it to be perceived as continuing to exist. This spatiotemporal information is stored in a *visual index* (Pylyshyn, 1989). Changes to the object's features (stored in an *object file;* Kahneman, Treisman, & Gibbs, 1992), such as those caused by the cut, will go unnoticed. A duration estimation study shows that these spatiotemporal expectations distort due to the attention withdrawal. Continuity editing rules show evidence of accommodating these distortions to create perceived continuity from discontinuous visual information.

The outcome of this thesis is a scientific understanding of filmic continuity. This permits filmmakers greater awareness of the perceptual consequences of their editing decisions. It also informs cognitive scientists of the potential of film as an analogue for real-world perception that exposes the assumptions, limitations, and constraints imposed upon our perception of reality.

Acknowledgements

I would like to thank my supervisors Helen Pain and John Lee, for their support and guidance. Their acceptance of my ideas and tolerance of my frequent leaps across research areas allowed me to produce this thesis.

I would also like to thank Graeme Ritchie who started out on the road towards this thesis with me and provided much needed instruction on how to be an ideal Ph.D. student. If I ever do another one I'll be sure to follow your instructions.

I want to acknowledge the financial support of EPSRC, the Institute for Communicating and Collaborative Systems, and the Informatics Graduate School.

Thanks has to go out to Scott Nowson who accompanied me on this journey. He was always present when I needed distraction, support, and reassurance that it's ok to do something a little bit "out there".

My deepest gratitude to my parents, Alison and Steve Smith, who have supported me emotionally and financially throughout my continuous 22 years of education. Mum, thank you for always being there to comfort me and bring me down to earth. Dad, thanks for getting me hooked on horror films from a ridiculously early age. Who'd have thought where it would lead!

My final thanks go to the person without whom this Ph.D. wouldn't have happened; without whom the last four years of my life wouldn't have been so joyous; without whom I wouldn't be the man I am today; and with whom I hope to spend the rest of my life…..Andrea Greve.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Tim J. Smith)

Table of Contents

Chapter 1: Introduction

This thesis constitutes a theoretical and empirical investigation of continuity editing. Continuity editing is the dominant style of film editing in practice today. It is so prevalent that to refer to the techniques of constructing and composing a film is to refer to continuity editing and the associated continuity style of filmmaking. The influence of continuity editing does not end with film. Television brings the conventions of continuity editing into our homes where we devote a significant portion of our waking hours to its consumption[1]. The visual experience of film or television is incompatible with our experience of reality (Arnheim, 1957). The dynamic visual nature of film allows us to perceive its content without minimum effort but formalistic conventions of cuts, spontaneous changes in viewpoint, and omissions of time and space seem to suggest that the perceptual skills we use to interact with reality should be incompatible with film (Hochberg, 1986).

However, recent evidence from research into change blindness, inattentional blindness, transsaccadic memory, and object perception suggests that our perception of reality is not as rigid or reliable as we think. It has been proposed that this flexibility of perception allows film to be perceived without any need for specialised perceptual skills (Anderson, 1996; Cutting, 2004; Gibson, 1979). Developmental evidence of young infant's ability to perceive film seems to support this view (Comuntzis, 1987). However, no dedicated empirical study has attempted to identify how we perceive film or how continuity editing functions to make film compatible with our perceptual abilities. This thesis constitutes such a study.

---

[1] 13.25% of an adult American's waking hours are spent watching television (source of statistic: Bureau of Labor Statistics, United States Department of Labor, 2004; http://www.bls.gov/news.release/atus.t08.htm).

An empirical and theoretical investigation of film editing should be of interest to various camps. Cognitive science can benefit from such an investigation as, by identifying how film experience is compatible with our perceptual abilities, a greater understanding of these abilities and their flexibility can be identified. Film can be viewed as a deviant form of visual experience. Its level of deviancy is controlled by continuity editing. The need for such control and the institutionalised status of continuity editing implies that not all film is compatible with our perceptual systems. By understanding the difference between "continuity" and "discontinuity" in film, insights may be gained for the perception of similar "continuity" during real-world viewing such as across saccadic eye movements, blinks, head turns, and occlusions.

The second group who should benefit from this empirical investigation of film perception are film theorists. Film theory attempts to understand, analyse, and criticise film at many different levels. Most of these levels are too high to be concerned with the issue of editing. However, the introductory film literature is pre-occupied with the continuity editing style and discussions of how film is constructed. These discussions typically resort to prescriptions of how the continuity editing rules should be followed and rarely question the rules. The recent emergence of Cognitive Film Theory indicates that there is a rising interest in the application of cognitive science to film. This thesis provides such a link.

Finally, the group who have most to gain from an empirical and theoretical investigation into film perception and continuity editing are the film editors. As should become clear during the course of this thesis, film editors are remarkably sensitive individuals who have great insight into the perceptual and attentional behaviours of others. This insight has been partially formalised as the continuity editing rules but the extent of their insights resist formalisation. Editors can benefit from this investigation in two ways: 1) the theoretical dissection and empirical manipulation of continuity editing allows a finer definition of the processes involved in making cuts "invisible" and creating the "illusion of seeing a continuous stream of action." (Reisz & Millar, 1953; pg. 48), and 2) the complexity of explaining film

perception and inadequacy of current theories highlights just how clever and accomplished editors are for being able to control factors that cognitive scientists are only just beginning to appreciate.

The thesis will be structured around three main questions:

How does continuity editing

1. minimise awareness of a cut,
2. create the perception of "continuity" across a cut, and
3. ensure that "continuity" is not violated as a consequence of the cut?

The 'Background' chapter (2) will first outline the main continuity editing rules as a reference point for the rest of the thesis. This will be followed by a brief discussion of existing theories of film perception as well as establishing the debate concerning the origins of film perception: innate, learnt conventions, or compatible with normal cognitive development. The psychological effects of cuts that violate the rules of continuity editing will then be presented and associated with the phenomenon of attention capture.

The 'Hiding a Cut' chapter (3) will document a theoretical investigation into the first question: "How does continuity editing minimise awareness of a cut?" Four mechanisms will be proposed by which attention capture by a cut can be avoided: 1) Focus attention on a part of the visual scene that does not change; 2) Have an expected change capture attention at the same time; 3) Direct attention internally to the processing of recently extracted information; or 4) Suppress attention during the change. The theoretical basis of each of these mechanisms will be presented as well as any existing empirical evidence that specific continuity editing rules employ such mechanisms.

The 'Cuing a Cut' chapter (4) follows on from the previous chapter by performing an empirical investigation into mechanisms 2 and 4. A primitive type of edit created according to the continuity editing rules is identified: matched-exit/entrance. The

editing factors used to create this cut are manipulated and their effect on attention measured. The results of this study provide an insight into how attention is controlled by editing, limits awareness of the cut, and creates the potential for continuity perception.

The 'What is Continuity?' chapter (5) addresses the second main question: "How does continuity editing create the perception of "continuity" across a cut?" This chapter begins with a theoretical analysis and re-imagining of continuity errors. The insights gained from this analysis are then compared to three cognitive research areas: object, spatial, and temporal perception. Existing theories and their compatibility with film perception are discussed. The result of this chapter (5) is a theoretical account of the concept referred to film editors as "continuity".

The 'Accommodating Expectations' chapter (6) follows on from the last question by empirically investigating: "How continuity editing ensures 'continuity' is not violated as a consequence of the cut?". A single dimension of continuity is empirically tested: time perception. The effects of attention on our perception of duration are identified and the attentional components of cuts specified. A duration estimation experiment is then used to identify how editing distorts perceived temporal continuity and how the continuity editing rules accommodate these distortions.

The 'Discussion' chapter (7) summarises each of the previous chapters, identifies possible limitations to the current studies, and discusses possible extensions and further work.

Finally, the 'Conclusion' chapter (8) summarises the main questions, theoretical and empirical contributions, conclusions and "take home" message of the thesis.

The result of this thesis can be considered an attentional theory of continuity editing.

Chapter 2: Background

This chapter will begin by establishing the main continuity editing rules as found in the film literature. Existing theories of the psychology of continuity editing will be discussed as well as some indication of whether film perception requires the learning of conventions. The psychological effects of cuts that violate the rules of continuity editing will then be presented and associated with the phenomenon of attentional capture. Based on the insights gained from discussion of attentional capture and inattentional blindness, three questions will be presented that need to be answered by any cognitive theory of continuity editing. These three questions will guide the theoretical and empirical investigation of this thesis.

2.1 What is Continuity Editing?

Film editing is the "coordination of one shot with the next" (Bordwell & Thompson, 2001). A **shot** is a single, unbroken period of recording with a moving-picture camera (whatever the recording medium: digital, video, or celluloid). When the film director decides that the narrative intention of the scene can not be communicated by the current shot a "cut" must be made to a new vantage point. Traditionally this entailed cutting the celluloid film between two frames (a single image resulting from exposing the film negative for 1/24th of a second), selecting the piece of film which contains the appropriate new shot, and gluing the two together. When the subsequent sequence is then projected, the visual scene prior to the edit will be instantaneously replaced by the scene depicted in the new shot with no noticeable gap or transition. Even today, with new forms of moving-picture recording and projection (e.g. digital and video), the process of "cutting" from one shot to another is essentially the same: the old shot is suddenly replaced by the new shot.

An ordinary Hollywood film typically contains around a thousand edits[2]; an action based film may contain more than two thousand edits (Bordwell & Thompson, 2001). In a film with a typical duration of 90 minutes this means that there is an edit every 5.4 up to 2.7 seconds (on average). This frequency may be even higher during fast-paced action sequences, sometimes bringing the time between cuts to less than 1 second. As such, the sensorially brutal act of suddenly replacing the entire visual scene should constitute a significant part of a viewer's experience of a film. Yet viewers find it hard to recall specific details of the editing after viewing a film, instead they recall the events of the film as one continuous sequence (Messaris, 1994). Somehow the editing has become "invisible" (Reisz & Millar, 1953), both during the initial viewing of the film and recollection of it. The technique used to accomplish this effect is known as **"continuity editing"**.

Continuity editing is the most dominant style of editing used today. It is synonymous with the Hollywood Style of filmmaking, emerging during the early days of cinema and consolidated during Hollywood's heyday of the 1930's and '40s. At that time continuity editing could be seen as distinct from other experimental and emergent forms of editing which were dominant in other parts of the world, most noticeably Russia and East Asia. However, these alternative forms of editing were gradually subsumed by the Hollywood style until they were seen as deviations from the standard and "only", according to the film industry, way to successfully edit a film: continuity editing.

When attempting to construct a film in accordance with the continuity style a filmmaker will make use of a series of "rules". These "rules" suggest how a scene should be staged, filmed, and edited so that the viewer can comprehend the scene with minimum effort. Whilst they are commonly referred to as the "rules" of continuity editing, in reality the rules are abstract heuristics, "rules-of-thumb", that provide a template to which a filmmaker can compare an edit as a gauge of the

[2] An "edit" refers to all joins between two pieces of film. It does not have to refer to a "cut". A cut is a specific type of edit involving the immediate change from one shot to another. The transition between shots can be made gradual by using dissolves or fades (see 3.1).

resulting degree of "continuity". They do not specify exactly how to film and edit a particular scene, they simply function as short-cuts that allow the filmmaker to assume that, by following the "rules", they will not be creating a "discontinuous" edit. The task of fine tweaking the continuity of each edit still falls to the editor who has to rely upon experience, intuition, and empathy with the target audience.

The continuity style of filmmaking is founded upon a system of camera placement and editing guidelines known as the **180° System**[3]

2.1.1 180° System

The 180° System is built upon the idea of an "**axis-of-action**": a clearly identifiable line in the 3D space of the scene down which all action, character movement, glances, and dialogue occurs. For example, in a classic dialogue scene involving two characters (see Figure 2-1) the axis-of-action joins the conversational partners. The director identifies this line by introducing a new scene with an **Establishing Shot**: a Full or Long Shot (see Appendix A in section 9.1) filmed perpendicular to the axis-of-action and far enough away from the main characters to show their positions relative to each other and their surroundings (camera A in Figure 2-1). In classical Hollywood cinema this will usually show the entire set on which the scene will take place and clearly identify any entrance/exits to the location. This shot also positions the audience within the space and establishes whether a character appears on the left or right of the screen.

[3] Also known as the "180° Rule".

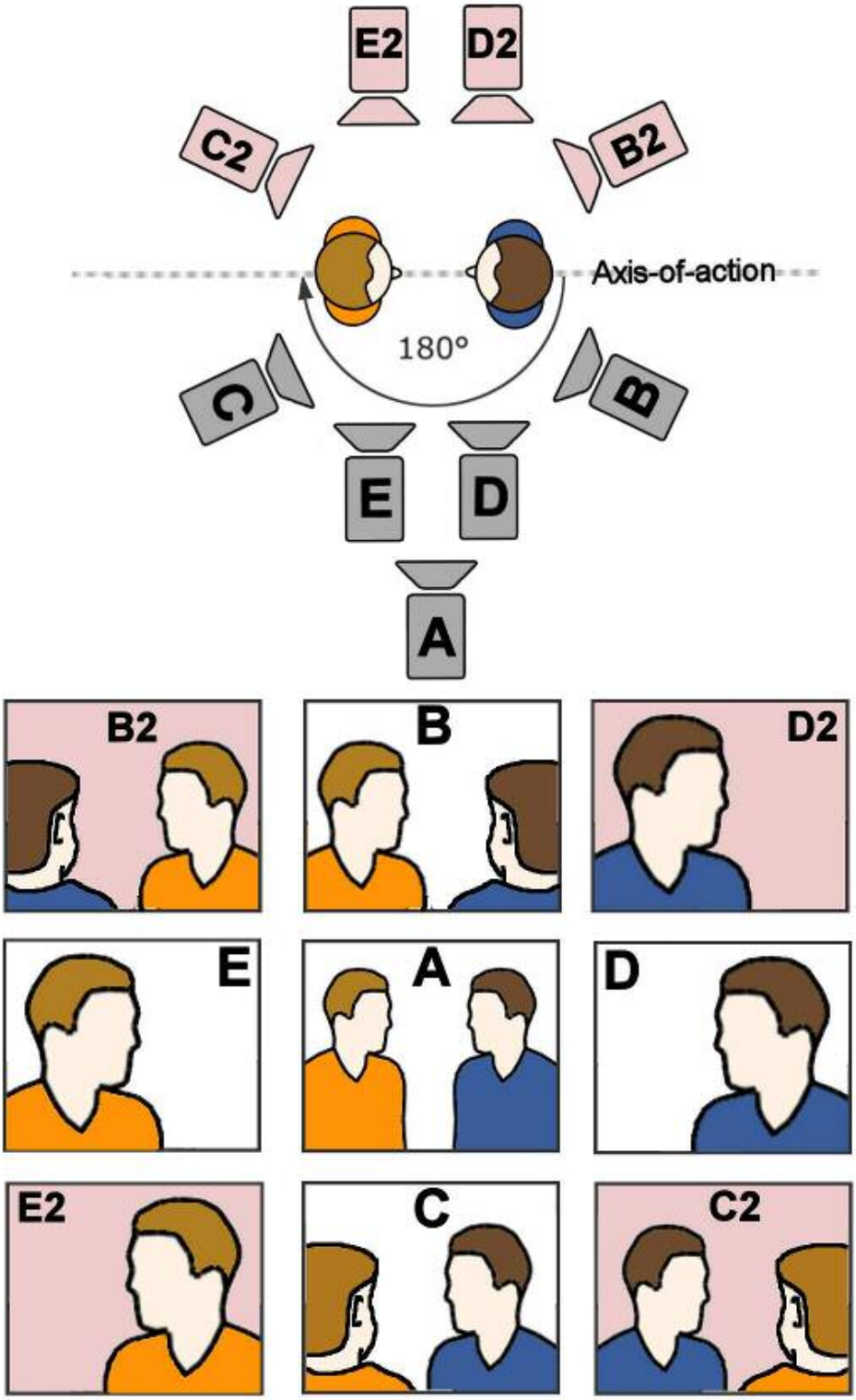


Figure 2-1: The 180° System of camera placement. Once the space of the scene has been established by camera A all other shots must be taken from the same side of the "axis of action". A cut across the line (cameras B2, C2, D2, or E2) would create a "discontinuity".

Once the axis-of-action and space of the scene have been established, the 180° rule states that **every subsequent shot must be filmed from the same side of the axis as the establishing shot** (paraphrased from Bordwell & Thompson, 2001; page 263). All camera positions along the 180° arc with the characters at its centre are permissible (Figure 2-1). If the camera were to cross the axis (cameras B2, C2, D2, and E2), the left-right relationships on the screen would be reversed (see the difference between B2 and B). This unexpected reversal is believed to "disorient the spectator" (Bordwell, Staiger, & Thompson, 1985). Violating the 180° rule in this way is known as "**crossing the line**".

2.1.2 Analytical Editing

Once the space of the scene has been established the director can begin to break it down to tell the story (referred to as **breakdown**). As Karel Reisz says "*the director's aim is to give an ideal picture of the scene, in each case placing his camera in such a position that it records most effectively the particular piece of action or detail which is dramatically significant.*" (Reisz & Millar, 1953). This is achieved by cutting to shots filmed from different positions within the 180° arc and at different distances from the actors. Close-Up shots (see Appendix A) permit the audience to see more detail and, in a dialogue scene, read the expressions on the character's faces. Drama emerges from the emotions, expressions, and thoughts portrayed in the character's face (Hitchcock, 1995). Therefore, the principle task of the director and editor is *"the organization of these oval shapes within the rectangle of the screen"* (Hitchcock, 1995).

As long as the camera stays on the same side of the axis-of-action as the establishing shot it is free to be positioned where ever the director desires (such as cameras A, B, C, D, and E in Figure 2-1). A common editing pattern is to start a scene as a Long Shot and then cut in to each character as they take turns speaking. If the camera is positioned perpendicular to the character (e.g. camera position D), they will have the entire screen to themselves. This is thought to attribute greater significance to the

actor (Bordwell & Thompson, 2001). The actor's eyeline will also follow the same path as in the establishing shot, e.g. looking from the right side of the screen off screen-left (compare shot A to shots B, C, D, and E). This **eyeline match** is the natural result of adhering to the 180° rule and is thought to reinforce the sense of space previously established (Messaris, 1994). Even though the second character is no longer on-screen his presence is implied by the current character's gaze (see how the actors appear to gaze at each other across shots E and D). By establishing eyelines in the establishing shot (A) cuts can later be made between shots that do not contain all actors. The intersecting eyelines are believed to result in the accurate comprehension of spatial relationships (Bordwell & Thompson, 2001).

If one of the characters were to move out of their Close-Up shot and change the spatial relationships of the scene, then the space of the scene would have to be re-established by cutting to a **Re-Establishing Shot** (e.g. cut back to shot A). This shot is usually identical to the original establishing shot so as to communicate how the space of the scene has changed. Once the character has stopped moving, the new axis-of-action joining the two characters becomes clear and the analytical breakdown of the scene can continue based on this new axis.

When an editor decides to cut from one camera positioned within the 180° arc to another camera, they must ensure that there was more than 30° between the two cameras. This is known as the **30° Rule**. If the cameras are too close and pointed at the same object the cut may make the objects on the screen appear to "jump" (Reisz & Millar, 1953). 30° is just a rule-of-thumb; the real implication of the rule is for the editor to ensure that the two shots joined by the cut are significantly different (Anderson, 1996). The "significance" of a change between shots is hard to define but it is usually achieved by using different shot sizes, changing the object the camera is pointed at, or introducing a new object into the second shot.

2.1.3 Reverse-Angle Shots

If editors only choose to cut between shots filmed perpendicular to the axis-of-action (cameras E and D) the lateral relationship of characters would be clear but the scene would appear flat. The scene would lack any depth as all the action and dialogue is directed out of the side of the frame, perpendicular to the audience. This is believed to lessen the impact the action has on the audience (Katz, 1991). The more commonly used sequence of shots is referred to as **Reverse-Angle Shots.** These position the camera at, or near to, one end of the axis-of-action and point it at the other end (shots C and B in Figure 2-1). The camera is typically positioned either in place of the listening character or just behind their shoulder (an **Over-The-Shoulder shot**). Then, when the conversational turn passes to the other character, a cut is made to a camera positioned in exactly the same position relative to the other character (this shot combination is known as Shot/Reverse Shot). An Over-The-Shoulder (OTS) shot shows both characters on screen at the same time: the shoulder and back of the listener's head dominate one side of the screen whilst the speaker's entire face can be seen in the other side of the screen. As long as the chosen shoulders are within the 180° arc the characters will remain on the correct side of the screen (see the reversal effect in shots C2 and B2 in Figure 2-1).

The use of Over-The-Shoulder and Reverse-Angle shots is probably the most recognisable product of the 180° system. When the Shot/Reverse Shot pairing of camera positions is combined with an analytical breakdown of the scene (establishment, breakdown, reestablishment) it is believed to create an apparently safe and rigid sequence of camera positions guaranteeing the viewer's appreciation of the spatial relationships between shots (Bordwell & Thompson, 2001; Dmytryk, 1986; Katz, 1991; Reisz & Millar, 1953).

2.1.4 Matched-Exit/Entrance

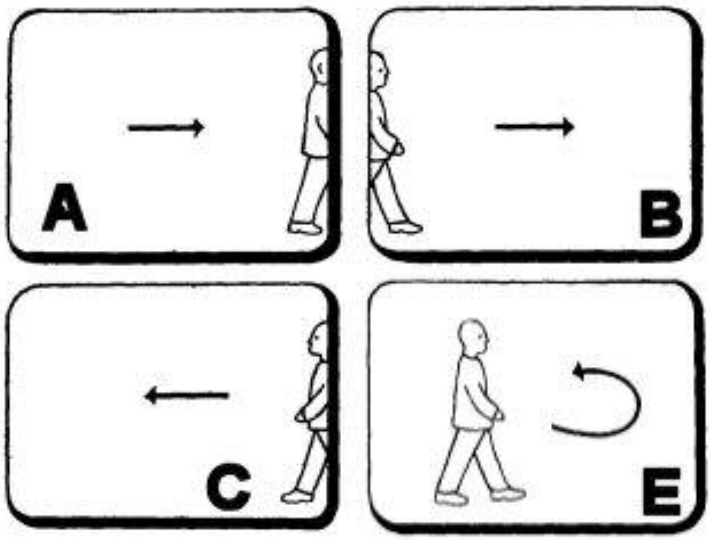


Figure 2-2: Directional continuity across a Matched-Exit/Entrance Cut. A cut from shot A to shot B is acceptable and will create what is referred to as a matched-exit/entrance cut. A cut from shot A to C is seen as a creating a spatial "discontinuity" unless shot E is placed between them.

The preservation of spatial relationships is not only a concern of Over-The-Shoulder and Reverse-Angle shots it applies to all shots depicting action. If an actor or object is seen moving left-to-right across the screen in one shot, the next shot of their action should also present them moving left-to-right (Bordwell & Thompson, 2001; Dmytryk, 1986; Katz, 1991; Reisz & Millar, 1953). This is referred to as a "**directional match**". It can be conceived as an axis of action extending beyond the confines of a single scene. Two shots of an action could be spatially adjacent within the same scene, such as two shots either side of a door, or omit thousands of miles, such as a shot of an actor boarding a plane left-to-right, the plane flying left-to-right, and landing left-to-right (Dmytryk, 1986)[4].

One technique used to ensure that viewers experience the "illusion of seeing a continuous piece of action" (Reisz & Millar, 1953) is to use a character's departure from the screen to cue a cut to another shot of that actor re-entering the screen. This

[4] This is quite an extreme example but smaller distance examples, such as a character leaving their apartment and travelling to work across numerous cuts are common.

type of cut will be referred to as a **matched-exit/entrance** cut[5]. An example can be seen in Figure 2-2. A cut from shot A to shot B will create the impression of continuous motion, whilst a cut from shot A to shot C is believed to confuse and disorient the viewer (Dmytryk, 1986). This confusion could be removed by inserting shot E in which the actor is seen changing direction between shots A and C.

2.1.5 Match-Action

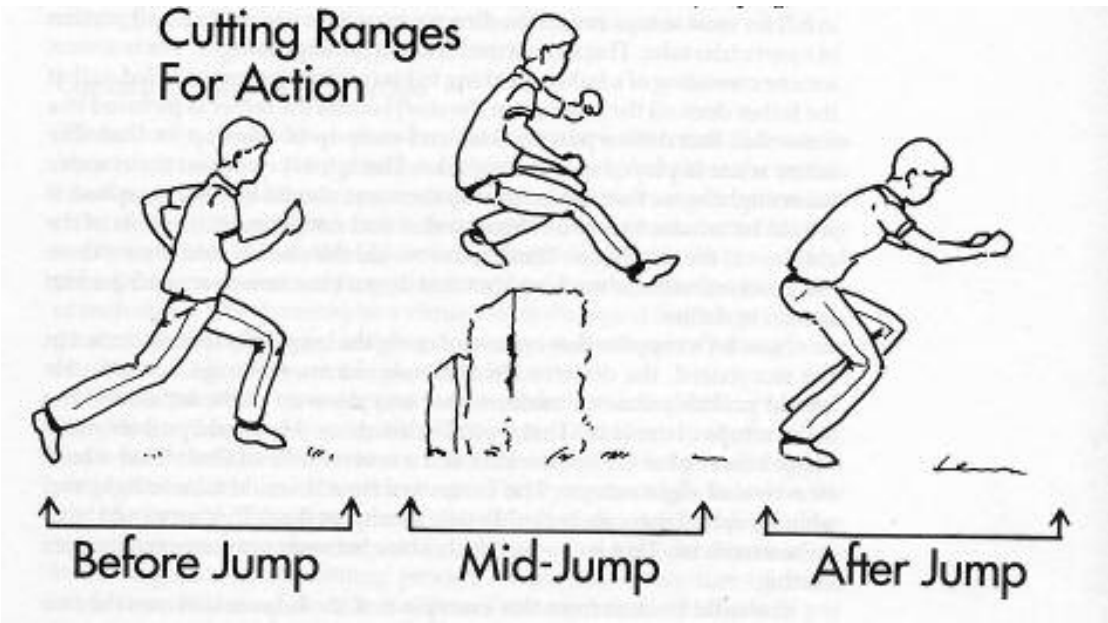


Figure 2-3: An optimal match-action cut would be placed between "Before Jump" and "Mid-Jump". Taken from (Katz, 1991).

As well as advising which shots to cut to and how to position cameras to create the least disruptive cuts, the continuity editing rules also suggest when to cut. Various rules exist for choosing the best time to cut (a range of these will be discussed in section 3.2) but by far the most prominent is the **match-action rule**. When an editor wishes to cut between two shots of the same actor, the least disruptive time to cut is

[5] This name was devised by the author of this thesis. This type of cut is referred to in the editing literature in various ways, such as "frame cut" (Bordwell et al., 1985), but there exists no consensual name. Matched-exit/entrance was chosen as it represents the cut's status as a match-action cut whilst specifying that the action is a screen exit followed by a screen entrance.

"at the point of greatest action" (Anderson, 1996). This is normally seen as a sudden change in action or motion such as an actor standing up, moving an arm, or making a head-turn. In terms of knowing when to cut in relation to this action "the common practice in the continuity style would locate the cut somewhere into the action…. This tends to hide the cut and make the transition to a new shot invisible. The exact point of the cut is dependent on the subject and the editor's sense of movement." (Katz, 1991; page 154). Following this advice, a match-action cut would be located somewhere just after the beginning of an action (Figure 2-3).

The benefit of using a match-action across a cut is believed to be due to our expectations about the action, "we are thus unlikely to notice the cut, because our expectations lead us to want to see what happens next." (Bordwell & Thompson, 2001; page 267). These expectations are believed to be critical to the construction of "continuity" and to make the editing "invisible": "It takes a practiced eye to spot a smooth match on action; so powerful is our desire to follow the action flowing across the cut we ignore the cut itself." (Bordwell & Thompson, 2001; page 267). This belief that continuity editing matches the viewer's expectation is common and will be discussed in the following section.

2.2 Existing theories of Continuity Editing

The rules and guidelines presented in the previous section are the key techniques prescribed by the continuity editing style for ensuring that the cuts are "invisible" and that "continuity" is perceived. Other rules exist depending on the complexity and detail of the editing literature referenced and editors will have their own rules-of-thumb that they consider part of the continuity style. However, an insight into the techniques of continuity editing can be gained from the selection presented here.

The next step towards motivating an empirical investigation of the cognitive foundations of continuity editing is to identify existing theories that attempt to explain how continuity editing works. Finding such theories amongst the wealth of theoretical movements associated with film can be very difficult. Most film theories

are not concerned with film at the level of the cut. Film is usually regarded as an artefact of a director (Auteur Theory), a psycho-social display (Psychoanalytic Theory), a political document (Marxist and Socialist Theory), or a text (Semiotic Theory). The only theoretical movement related to film that is interested with similar issues to this thesis is Cognitive Film Theory[6]. This is a relatively young theoretical movement which has yet to gain full support from the rest of the film community. The theorists associated with Cognitive Film Theory endeavour to find a bridge between the tradition of film theory they are based in and the current theories of cognitive science. This is a noble endeavour which, for it to be a success, requires cognitive scientists to participate in the endeavour and attempt to contribute to the bridge from the scientific side.

However, even within the area of cognitive film theory there are very few detailed explorations regarding cognition of film editing. To find insights in this area a broader range of literature must be surveyed.

Since the earliest days of film (circa 1916), editing has been interpreted as an analogue for real-world shifts of attention (Münsterberg, 1970). This view is elegantly summarised by one of the most accomplished early film directors, D. W. Griffths:

> "Looking at real things, the human vision fastens itself upon a quick succession of small comprehensible incidents, and we from our eventual impressions, like a mosaic, out of such detail… The director counterfeits the operation of the eye with his lens." (Griffith, 1926, quoted in Jesionowski, 1982; page 46)

A cut is seen as analogous with a shift in attention and the shots either side of that cut, periods during which the eyes are static (i.e. fixations). This apparent analogical relationship between film and attentional shifts in the real-world is seen, by some

---

[6] A summary of the intentions of Cognitive Film Theory can be found at http://www.geocities.com/david_bordwell/cognitive.htm

theorists, as being sufficient explanation for our ability to perceive edited film as continuous (Münsterberg, 1970). However, the visual experience of film and our experience of real-world attention shifts are not directly compatible (Arnheim, 1957). There are at least five features of film that make the visual experience of film distinct from reality (adapted from (Hochberg & Brooks, 1978b):

1. Film can provide motion information about a three-dimensional space that is not shared by the viewer (e.g. depicted camera movements whilst the viewer is static).
2. Film can represent visual scenes much larger or smaller than the size of the screen upon which it is presented (e.g. a Close Up Shot).
3. Film can instantaneously change the visual scene across a cut. Such changes are outside of the viewer's control (i.e. exogenous[7]).
4. Film can represent scenes and events in a piecemeal fashion, by juxtaposing views of objects that are not spatially, or even temporally adjacent.
5. Redundant sections of actions, periods of time, or extents of space can be elided (i.e. removed), and series of events can thereby be reduced to their minimal communicative features.

Any theory of film perception must account for these distinctions. The theories that have been developed can be identified as belonging to two different schools of psychology: ecological perception and constructivism.

The ecological theory of film perception suggested that the incompatibilities between film and reality (such a those presented above) do not obstruct a viewers ability to perceive a film directly i.e. without the need for specifically developed perceptual skills (Gibson, 1979). The changing visual field presented within a shot is similar enough to real-world experience that the information can be perceived. For example, movements of the camera create the same changes in the visual field as a similarly

[7] "Exogenous" refers to a locus of control external to a person e.g. when a loud noise involuntarily attracts attention. "Endogenous" control refers to the opposite: a voluntary act of control by a person e.g. choosing to move their eyes to an object.

moving viewer. These movements induce perceived motion in a viewer. The cognitive dissonance resulting from the conflict between the visual information and viewer's awareness of their actual status in space is dismissed in favour of successful comprehension of the depicted events (Gibson, 1979). Essentially, film perception is driven by direct perception of the authentic visual experience within shots and the viewer's conceptual understanding of the depicted events enables the visual discontinuities of cuts to be bridged (Gibson, 1979).

The use of ecological perception to explain film viewing is still popular today (Anderson, 1996; Cutting, 2004; Tan, 1995). The ecological view has been extended to incorporate recent discoveries concerning the flexibility of our perceptual system (these discoveries will be discussed in section Chapter 5:). The main assumption is that viewers are tolerant to a sudden change in viewpoint because the change presents a new visual scene which is of interest. The incompatibility of the visual experience of cuts with that of the real-world is not important to the viewer (Cutting, 2004). What is important is that they can continue to follow the narrative. This would appear to place the viewer in agreement with continuity editing which has been interpreted as placing the preservation of "narrative continuity" above all other things (Bordwell & Thompson, 2001).

This explanation of the perceptual function of continuity editing as preserving the narrative is similar to the concept of *visual momentum* proposed by Julian Hochberg (Hochberg & Brooks, 1978a; Hochberg & Brooks, 1978b). *Visual momentum* refers to the "*impetus to obtain sensory information and to formulate and test a schema*[8]" (Hochberg & Brooks, 1978a; page 295). As previously hypothesised by the earliest film theorists (e.g. Münsterberg, 1970), continuity editing is seen as functioning by driving viewer's attention across cuts. A viewer pursues *perceptual enquiries,* periods of development and test of schematic maps, across cuts. As long as the rate of these *perceptual enquiries* is maintained by editing and the editing provides

[8] Template-like knowledge of events.

information compatible with the *enquiries*, the viewer will be able to overcome the visual incompatibilities between film and reality (Hochberg & Brooks, 1978b).

These theories of continuity editing provide insight into how cuts may be bridged by conceptual relationships (e.g. narrative continuity, perceptual enquiries) but they do not deal with the low-level perceptual experience of a cut. The visual incompatibilities between real-world shifts of attention and their analogue in film should mean that the perception of film is effortful and requires adaptation. This view was shared by Hochberg:

> "Some of the filmic devices.... are purely conventions and are arbitrary. But some probably draw on innate characteristics of the nervous system, and some are learned but rest on a great deal of perceptual habit that must be established outside the cinema and have much greater strength and stability than would be expected with conventions that had been learned merely through experience with cinema." (Hochberg, 1986; page 52)

Hochberg was acknowledging that direct perception could not explain all film perception as there appeared to be a need either for the learning of uniquely cinematic conventions or of perceptual behaviours developed through interaction with other compatible visual experiences. These different sources of the perceptual abilities required to perceive film can be divided into three categories:

- **Innate**: aspects of film that are directly compatible with the human visual system before maturation i.e. from birth;
- **Developed**: the abilities necessary for perceiving certain aspects of film emerge as a natural byproduct of normal perceptual development;
- **Learnt**: aspects of film that are "pure" convention and require the viewer to be repeatedly exposed to the convention before they are able to successfully perceive/comprehend the film.

The ecological view of film perception only considers the first two of these categories and dismisses the need for the learning of conventions. A constructivist theory of film perception, such as that proposed by Hochberg, acknowledges that there are many levels of cognition required to successfully process film and some of

these may be dependent on learnt skills. For example, the cinematic convention of using fades to indicate a long passage of time between shots compared with dissolves for short elipses seems to be a convention that has no real-world correlate[9]. An ecological theory of film perception cannot explain how such transitions would be interpreted as different omissions of time without referring to higher-order cognitive processing. By comparison, a constructivist theory appreciates the levels of processing required to perceive a visual scene. By viewing all perception in this way, the learnt perceptual skills required to process arbitrary cinematic conventions can be seen as just another level of the cognitive processes required to perceive film.

Identifying which cinematic constructions require learning and which are compatible with normal perceptual experience is methodologically difficult as the conventions of continuity editing are universal. They exist in all visual medium and we are bombarded by them throughout our lives. Distinguishing between our ability to perceive filmic constructions based on specialised learnt skills or as a function of normal perceptual development is difficult in our society as such a study would require subjects that were without this exposure to film and had already fully perceptually developed.

Developmental studies have shown that a child's allocation of attention to television increases fourfold between the ages of 12 and 48 months and continues to increase right into the teenage years (Anderson & Levin, 1976). If viewers were able to perceive film directly due to the compatibility between the filmic constructions and normal shifts of attention, we wouldm't expect to see any signs of change in their ability to perceive film during childhood. This evidence that development does occur has been explained as being due to both natural cognitive development and the

[9] It is possible that our visual system is trained to associate a gradual decrease in light levels with the setting of the sun and the onset of night time, a period which is usually accompanied by a long period of missing time during sleep. By comparison, short dissolves between different viewpoints may be visually similar to the eye blinks associated with microsleeps, minute periods of time when a person falls asleep and wakes up immediately. However, drawing such analogies with real-world behaviours as a way of motivating cinematic conventions is rather weak.

learning of the cinematic conventions. The important components related to natural development seem to be the increased control of attention allocation, comprehension of events, and development of the concepts required to understand the visual information (Anderson & Levin, 1976).

In order to separate the contributions of cognitive ability and learnt conventions from a child's ability to perceive film, their level of cognitive ability needs to be identified. In experiments by Comuntzis-Page, children between the ages of 3 and 7 were shown films edited according to the 180° Rule containing Reverse-Angle Over-The- Shoulder shots (Comuntzis, 1987; Comuntzis-Page, 1991). The children were then tested for their ability to understand the spatial relationships of the events depicted across the shots. The principal assumption of the 180° System is that it results in clear comprehension of spatial relationships (Bordwell & Thompson, 2001). Comuntzis-Page's experiments provided evidence for this assumption by showing that her subjects were able to comprehend spatial relationships between the shots (Comuntzis, 1987; Comuntzis-Page, 1991). She also showed that this ability was dependent on the child's ability to adopt another person's perspectives within the real-world and independent from the child's exposure to television and film. This *perspective-taking* skill is developed during a child's first few years of life (Piaget, 1954). This evidence suggests that our ability to perceive the 180° System and (at least some of) its constituent techniques is due to normal perceptual ability.

Similar evidence that some of the conventions of film are based in our normal perceptual ability has been provided by a social-anthropological study. Renee Hobbs and her associates presented two versions of a film to people from a remote part of Kenya (Hobbs & Frost, 1988). The subjects of this study had almost no experience of television, film, or mass-media in general. One of the films presented a familiar event in a single-continuous Long Shot. The other film was edited according to the continuity editing rules. When asked to recall the content of the film, there was no difference in recall between the two types of film. It was concluded that, even for first or second-time viewers, the conventions of continuity editing do not obstruct a viewer's perception of the depicted events (Hobbs & Frost, 1988).

However, this does not discount the possibility that some of the conventions used in continuity film are arbitrary and require learning. The most famous study that attempted this separation was Worth and Adair's social-anthropological study of the Navajo (Worth & Adair, 1972). Several young Navajo, with apparently limited exposure to film, were trained in the use of film and editing equipment. They were then encouraged to construct films using these devices without the experimenters instructing them on accepted filmic conventions such as continuity editing. The resulting films were then analysed to see if any peculiarities in their construction could be identified. Worth and Adair identified several sequences in the resulting films where their Navajo filmmakers violated rules of continuity editing. One sequence widely discussed (see Messaris, 1994), cut an action into multiple shots without matching action, time or space across the cuts. When the filmmakers were asked about these discontinuities they saw nothing wrong with them. They believed that the clear context and progression of the action was sufficient to allow the viewer to "fill in" the rest (Worth & Adair, 1972). Worth and Adair identified these violations of the rules of continuity editing as indicating differences between the cultural conventions of the Navajo and those prescribed by Hollywood (the main source of the continuity style). This would specify match-action editing as being a cinematic convention that viewers must learn.

However, Worth and Adair's studies must be judged with caution as it is methodologically questionable. First, the Navajos which were used as filmmakers originated from a reservation but most had broad experience of life outside of the reservation, such as college. Second, the assumption that the use of camera and editing equipment can be learnt without the implication of some of the principles of continuity editing seems flawed. For example, showing how film is cut and spliced together to make new sequences of images implies that such a technique is desirable in film. Such practical insight into the possibility of cutting out parts of a visual sequence could explain the Navajo filmmaker's eventual use of temporal ellipsis across match-action (i.e. cutting out time during an action). These methodological

issues make it hard to attribute the difference between the Worth and Adair Navajo films and continuity films as being due to cultural differences and learnt conventions.

A third technique[10] for distinguishing between the innate, learnt (e.g. conventions), or developed basis of our ability to perceive continuity films is to look for compatibilities between known perceptual behaviours and those required by film. By analysing the apparent differences between the visual experience of film and reality and then investigating whether normal perceptual abilities are sensitive to such differences (as suggested by Anderson, 1996; Cutting, 2004; Tan, 1995) allows us to identify whether the basis of film perception is innate, learnt, or developed. This technique will be employed during the rest of this thesis.

Before an investigation of the cognitive foundations of continuity editing can begin an understanding of the concept referred to as "continuity" is required. Editor's definitions of "continuity" have already been seen (see section 2.1). However, for the definition of "continuity" to motivate a cognitive investigation the definition must be framed in the psychological experience of film. Such a definition will be sought in the next section.

2.3 The psychological effects of discontinuities

One of the main assumptions of continuity editing is that it prescribes ways of "joining together two shots in such a way that the transition does not create a noticeable jerk and the spectator's illusion of seeing a continuous piece of action is not interrupted." (Reisz & Millar, 1953; pg 216). As such, "continuity" is defined as an absence of experience (e.g. the "jerk"). Therefore, to begin understanding the psychological experience of "continuity" the experience of "discontinuities" should first be identified. Then by comparing the visual experience of a "discontinuity" to a

[10] The other two being developmental and social-anthropological studies.

"continuity" cut a direct connection between visual and psychological experience can be created.

Finding evidence for the psychological effects of editing discontinuities requires a broad survey of a range of research areas. The most significant contributions do not come from psychological disciplines such as visual or perceptual psychology as may be expected but from such areas as communication research, media psychology, computational vision, and human-computer interaction. By surveying these areas a range of psychological and physiological effects can be found that are associated with editing discontinuities:

- When subjects are instructed to press a button as soon as they see a cut, they respond faster to discontinuity cuts than continuity cuts (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990).
- If subjects have to perform a secondary task whilst watching a film, their responses to the secondary task will be slower after discontinuities (Geiger & Reeves, 1993; Lang, Geiger, Strickwerda, & Sumner, 1993).
- Both of these effects can be attributed to attention being directed towards the source of the discontinuity. This has been observed in EEG recordings (indicating focussed attention; Reeves et al., 1985).
- An increase in attention and the need to extract as much visual information from the scene as possible also results in an increase in the frequency of saccadic eye movements 200-400ms after the discontinuity (d'Ydewalle, Desmet, & Van Rensbergen, 1998; Hochberg & Brooks, 1978a; May, Dean, & Barnard, 2003).
- This is preceded by a short period (150-300ms) of cognitive overload during which no visual information is processed. The length of this period increases with discontinuity (Geiger & Reeves, 1993).
- The increase in attention also results in improved accuracy (Frith & Robson, 1975; Lang, 1991; Seddon, 2003) and faster recognition memory for information originally presented after a discontinuity (Carroll & Bever, 1976).

- Recognition accuracy increases as the degree of temporal discontinuity increases across a match-action cut (Seddon, 2003)[11].
- But, large discontinuities such as those occurring between scenes damage recognition memory (Bolls, Hibbs, & Lang, 1995; Bolls, Potter, & Lang, 1996; Lang et al., 1993).
- Whilst immediate recognition may improve, recall of the entire film sequence deteriorates when discontinuities are present (Frith & Robson, 1975; Kraft, 1987).
- Physiological studies have also shown that discontinuities are followed by a deceleration of the heart rate (Lang, 1990; Lang et al., 1993; Reeves et al., 1985) and an increase in physiological and self-reported arousal (Lang, Zhou, Schwartz, Bolls, & Potter, 2000).

Considered in combination, these effects indicate that discontinuities result in attention being directed towards the cut (d'Ydewalle et al., 1998; Geiger & Reeves, 1993; Hochberg & Brooks, 1978a; Lang et al., 1993; May et al., 2003; Reeves et al., 1985), leading to heightened awareness of the editing (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990), and a deterioration of the viewer's ability to comprehend and remember the film's content (Bolls et al., 1995; Bolls et al., 1996; Frith & Robson, 1975; Kraft, 1987; Lang et al., 1993). This would be catastrophic if the film was intended for educational or informative use. Even if the film was only intended as entertainment, increasing viewer's awareness of the editing detaches them from the action, characters, and world depicted in the film and, most importantly, undermines the narrative. As all film form is subservient to narrative (Bordwell & Thompson, 2001), it appears that the objective of continuity editing is to avoid these negative effects. To understand why cuts violating the continuity editing rules create these effects whilst cuts adhering to the rules do not, an examination of the known trigger conditions for these effects is required.

[11] Mark Seddon performed this MSc project under the supervision of the author of this thesis.

2.3.1 Orienting Response

The psychological and physiological responses identified in the previous section are not unique to editing. They can occur any time an unexpected sensory stimulus suddenly appears. This set of responses are collectively referred to as the **Orienting Response (OR)** (Lynn, 1966; Pavlov, 1927). An OR immobilises the body whilst increasing the senses' ability to acquire information (Lynn, 1966). The complete response set includes dilation of the blood vessels to the head (increasing blood flow), decrease in the alpha frequency of the EEG (interpreted as the focusing of brain activity), deceleration of the heart, increases in skin conductance and temperature, and general constriction of the blood vessels to the major muscle groups (inhibiting mobility; Lang, 2000). These responses are all associated with an increase in attention and cognitive resources available for the processing of sensory information (Lang, Simons, & Balaban, 1997).

Every cut has the potential to trigger an orienting response (Geiger & Reeves, 1993; Lang et al., 1993; Singer, 1980). The deciding factors are whether an OR has recently occurred (an OR lasts ~8-9 seconds[12] and not until after this time period can another OR be triggered), and whether the change in sensory stimuli is significant enough to capture attention. The cognitive results of an OR are also dependent on the size of stimulus change. An OR always causes an increase in cognitive resources which, if the change in stimulus does not require a lot of cognitive resources (e.g. a cut between two shots of the same action) results in improved encoding (Frith & Robson, 1975; Lang, 1991; Seddon, 2003). However, if processing of the new stimulus is cognitively difficult (e.g. a cut between scenes/locations) these extra resources will lead to cognitive overload and encoding will deteriorate (Bolls et al., 1995; Bolls et al., 1996; Lang et al., 1993).

If the degree of discontinuity becomes too large, as in a sudden explosion, the observer's sensory system will respond with a **defensive response (DR).** This has

[12] An OR is typically measured in heart beats. An average OR lasts 10 heart beats which at an average resting rate of 72 beats per minute is equal to 8.33 seconds.

the opposite effect to an OR, diverting cognitive resources (and blood flow) from the sensory system, averting the sensory organs from the source and readying the body for action (Andreassi, 1995). This is the so-called "Fight or Flight" response.

This continuum of automatic responses to visual discontinuities indicates that the human visual system is sensitive to varying degrees of continuity and has innate ways of dealing with them. There is obvious evolutionary advantage to a continuum of responses. If a slight but sudden movement is seen in the periphery of vision the immediate question the viewer needs answering is "what is it?" (Pavlov, 1927). An OR provides the tools required to answer this question: orienting of sense organs, focussed attention and increase cognitive resources. The same can be said about the visual experience of viewing a cut. Whilst not critical for survival, the successful comprehension of the cause of the change e.g. the new shot is important for the current task: following the film's plot. For every cut there should be a degree of discontinuity at which the increase in attention and cognitive resources compliment the existing cognitive load and result in improved encoding of the visual information. Once this level of discontinuity is exceeded (or the cognitive resources already employed in processing the signal become too high) cognitive overload will result (Bolls et al., 1995; Bolls et al., 1996; Lang et al., 1993) and possibly even a defensive response[13]. The important point to note about DRs is that whilst they may produce a momentary increase in arousal (readying the body for "flight") they become habituated very rapidly (Lang, 1990). If editing repeatedly elicits DRs eventually the viewer will stop responding to them.

So what is it that elicits orienting responses? The human visual system is clearly sensitive to certain changes of visual stimuli but what exactly are these? And, of critical importance to this thesis, what decides if the viewer becomes aware of these changes, and by implication, the editing? The answers to these questions lie in recent findings from attention capture and inattentional blindness.

---

[13] Defensive responses can be seen as the "shock" or "startling" sensation experienced during an action or horror sequence. Good directors of these genres implicitly know how to trigger DRs in a deliberate fashion, using them sparingly for maximum effect.

2.3.2 What captures attention?

Attention capture is the involuntary redirection of attention towards a sensory event (Folk & Gibson, 2001). Visual attention may be captured overtly, manifest as eye movements towards the source of capture (e.g. Brockmole & Henderson, in press; Theeuwes, Kramer, Hahn, & Irwin, 1998), or covertly (cognitive resources withdrawn from a primary task resulting in a decrease in performance or increase in reaction times (e.g. Folk, Remington, & Johnston, 1992; Theeuwes, 1994). When the capture is overt this does not, as might be thought, automatically indicate that the viewer is aware of the cause of the capture. Eye tracking studies have shown that ocular capture occurs even when the viewer is unaware of the source of the capture (Brockmole & Henderson, in press; Theeuwes et al., 1998). The distinction between whether a viewer is aware (explicit capture) or unaware (implicit capture) of the sensory event is a much more complicated than overt/covert (see Simons, 2000 for an overview of these distinctions).

Investigations into what visual features capture attention have produced a wide range of different results. The list of features that have been shown to capture attention includes:

- **abrupt onsets of visual objects** (Boot, Brockmole, & Simons, in press; Brockmole & Henderson, in press; Folk, Remington, & Wright, 1994; Franconeri & Simons, 2003; Theeuwes, 1994; Theeuwes et al., 1998; Yantis & Jonides, 1984),
- **abrupt disappearances** (Brockmole & Henderson, in press),
- **motion onset** (Abrams & Christ, 2003; Franconeri & Simons, 2003),
- **apparent motion** (Folk et al., 1994),
- **looming stimuli** (Franconeri & Simons, 2003),
- **contrast/luminance changes** (Enns, Austen, Di Lollo, Rauschenberger, & Yantis, 2001), and
- **unique and changing colours** (Boot et al., in press; Folk et al., 1994; Theeuwes, 1994).

These features are widely seen in film. In fact, certain visual events, such as sudden appearances, disappearances, apparent motion or change in colour occur more frequently in film than they do in reality. It is physically impossible for a real-world object to suddenly appear, disappear, or change visual properties without going through some period of transition (Gibson, Kaplan, Reynolds, & Wheeler, 1969; Michotte, 1991). This abnormality of the visual events could be the precise reason why they capture attention. However, in film any change of the visual scene is possible either by cutting to a new shot or using special effects to transform parts of the scene. Attention capture should be just as (if not more) prevalent when watching film than when viewing the real-world.

The key distinction that needs to be made is between capture by the content of the film and capture by the formal elements such as editing. A cut constitutes an abrupt onset of an entire visual scene. The attention capture experiments cited above indicate that this type of visual event produces the most reliable capture (Boot et al., in press; Brockmole & Henderson, in press; Folk et al., 1994; Franconeri & Simons, 2003; Franconeri, Simons, & Junge, in press; Theeuwes, 1994; Theeuwes et al., 1998; Yantis & Jonides, 1984). However, increased evidence of orienting to discontinuities compared with continuity cuts (see section 2.3) indicates that not all cuts capture attention to the same degree and viewers are more aware of attention capture by discontinuities (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990).

2.3.3 Inattentional Blindness

When a sudden salient sensory event (such as the abrupt onsets caused by a continuity cut) fails to capture attention it is described as **inattentional blindness** (Mack & Rock, 1998). The most famous examples of inattentional blindness are Simons & Chabris' Umbrella Woman and Gorrilla studies (Simons & Chabris, 1999)[14]. The Umbrella Woman study presented two simultaneous semi-transparent

[14] The Umbrella Woman study is a replication and extension of an earlier study (Neisser, 1979).

visual scenes in the same display[15]. In one scene a group of people wearing white play basketball. In the other scene, a group in black are also playing basketball but 45 seconds into the game a woman carrying a black umbrella walks through their game (see Figure 2-4; a). When subjects are instructed to count the number of times the team in white pass the ball 57% fail to notice the umbrella woman (Simons & Chabris, 1999).

Even more impressive is the Gorilla study. Simons & Chabris (1999) constructed a similar film depicting two teams playing basketball but this time they were both filmed at the same time (not two films overlaid; see Figure 2-4; b). In this film, when a man wearing a black gorilla suit walked through the game mid way through 73% of viewers failed to notice him (Simons & Chabris, 1999). When shown the film again all subjects notice the gorilla and are stunned that they failed to notice him the first time (Simons & Chabris, 1999).

Figure 2-4: Simon and Chabris' (1999) inattentional blindness study. Top row shows the two superimposed films from the Umbrella woman experiment. Bottom row shows a series of frames from Gorilla Study.

[15] This is a version of the selective looking paradigm, a visual version of the dichotic listening task (Becklen & Cervone, 1983; Neisser, 1979; Neisser & Becklen, 1975).

These examples of inattentional blindness highlight the importance of viewing task for the detection of unexpected sensory events (Simons, 2000). When subjects were instructed to count the passes of the white group only 8% noticed the black gorilla but this increased to 46% when subjects followed the black team (Simons & Chabris, 1999). The subject's use of the colour black to direct their attention within the scene meant that they were also attending to the gorilla. The task dictated how they should focus their attention and filter out irrelevant visual features. This specificity has been described as a viewer's **attentional set** (Folk, Remington, & Johnstone, 1992).

The main technique used for highlighting the effect of attentional set is the **pre-cueing paradigm** (Folk & Remington, 1998; Folk et al., 1994; Folk et al., 1992; Gibson & Kelsey, 1998). This paradigm asks subjects to search for a target that differs from all other objects in the display according to a specific feature (e.g. the colour green). If before the display appears the location of the target is pre-cued using the same feature (e.g. a green position marker), subjects will be quicker to identify the target even though they are informed that the pre-cue is equally as likely to appear anywhere in the display (Folk & Remington, 1998; Folk et al., 1992). If the pre-cue differs to the target's defining feature (the pre-cue has a sudden onset) it will have no effect on performance (Folk et al., 1992).

It appears that the attentional set can occur at many levels. It can be location based or object based (Yantis & Jonides, 1990), dynamic or static (Folk et al., 1994), unique items or a specific conjunction of object features (Folk & Remington, 1998). If the distracting object shares the features at exactly the right level as that specified by the attentional set then capture should occur. Any deviation from this and the likelihood of attentional capture decreases (Most, Scholl, Simons, & Clifford, 2005; Simons, 2000).

This evidence suggests that some bottom-up properties such as salience or sudden onsets, influence the likelihood that an unexpected object will capture attention, but the most important factor appears to be the attentional set adopted by the individual (Simons, 2000). In the absence of an attentional set abrupt onsets appear to be the

only reliable feature that will capture attention (Franconeri et al., in press). This has been explained as due to a default attentional set that has survival benefits (Gibson & Kelsey, 1998). In the real-world, objects that suddenly change colour, brightness, or shape rarely pose any danger to the observer (if the occur naturally at all). By comparison, the sudden appearance of an object or its sudden movement could indicate that a predator is about to attack. Including these features in a default attentional set that allows the feature to capture our attention would make us more capable of surviving in hostile environments. Also, limiting the incidence of attentional capture through the use of an attentional set also ensures that when attention is focussed it is not involuntarily captured. These two processes, voluntary focussing and involuntary capture, and their coordination is key to visual attention (Allport, 1989).

Applying this idea of attentional sets to film viewing it becomes clear that the orienting response potentially triggered by every cut (Geiger & Reeves, 1993; Lang et al., 1993; Singer, 1980) can be attributed to this default attention to abrupt onsets (Franconeri et al., in press). However, it appears that this level of attention capture is generally implicit as continuity cuts do not lead to awareness of the editing (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990).Whilst there does not currently exist a clear understanding of why certain visual events lead to implicit and others explicit capture there does seem to be an indication that relevance to attentional set is a factor (Most et al., 2005). If a visual event is relevant to a viewer's current attentional set there is a higher probability that they will become aware of the event. However, this effect is modulated by the availability of attention. If attention is engaged by another task the likelihood of capture is reduced (Simons, 2000).

This interpretation of inattentional blindness suggests that for a viewer to be unaware of a cut the continuity editing rules must ensure that the visual information does not change in a way that is significant to the viewer's current attentional set or that insufficient attention is available for awareness. Suggestions of how the continuity editing rules could manipulate these factors will be presented in Chapter 3:.

2.4 Main Questions

This background section has established the continuity editing rules, summarised existing theories of "continuity", and identified the psychological effects of "discontinuities". The incompatibilities between the visual experience of film and reality have been identified. These incompatibilities are believed to be "bridged" by viewer's *perceptual enquiries* (Hochberg & Brooks, 1978a; Hochberg & Brooks, 1978b) and primary interest in the narrative (Bordwell & Thompson, 2001; Cutting, 2004). The result is believed to be a clear comprehension of space and action across a cut (Bordwell & Thompson, 2001; Reisz & Millar, 1953). The perceptual abilities required to perceive film are believed to be either innate, learnt from exposure to film, or developed as part of normal cognitive development (Hochberg, 1986). Evidence for the compatibility between developed perceptual abilities and the perception of film constructed according to the continuity editing rules has been provided by developmental (Anderson & Levin, 1976; Comuntzis, 1987; Comuntzis-Page, 1991) and social-anthropological studies (Hobbs & Frost, 1988). However, it is also believed that continuity editing takes advantage of innate *perceptual enquiries* and attentional shifts (Hochberg & Brooks, 1978a; Hochberg & Brooks, 1978b). This evidence limits but does not exclude the possibility that some editing conventions are learnt.

Continuity editing has two main objectives: to ensure that the viewer is not aware of the editing, and that they perceive continuous action across the cut (Reisz & Millar, 1953). It has been shown that continuity editing appears to result in the perception of spatial continuity across cuts (Comuntzis, 1987; Comuntzis-Page, 1991; Hobbs & Frost, 1988) but how this is possible when the visual information is spatially discontinuous is not known. Empirical investigations of the psychological effects of cuts have shown that attention is captured by both continuity and discontinuity cuts (Lang, 2000) but that only discontinuities lead to cognitive dissonance and awareness of the editing. Understanding how the continuity editing rules control the psychological effects of a cut and allow the perception of "continuity" will be the main objective of this thesis. These issues will be investigated in three stages according to three questions:

How does continuity editing

1. minimise awareness of a cut,
2. create the perception of "continuity" across a cut, and
3. ensure that "continuity" is not violated as a consequence of the cut?

Chapter 3: Hiding a Cut

As identified in the previous chapter (2), the visual disruption caused by a cut has the potential to capture attention and make the viewer aware of the editing (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990). The degree to which this capture occurs differs between continuity and discontinuity cuts (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990). As the discontinuity in visual stimulation is just as large across both continuity and discontinuity cuts some other factor must be limiting the degree of attention capture. There are four main ways that attention capture by the cut can be avoided[16]:

1. Focus attention on a part of the visual scene that doesn't change (**Inattentional Blindness**);
2. Have an expected change capture attention at the same time (make the change part of viewer's **Attentional Set**);
3. Direct attention internally to the processing of recently extracted information (the resulting effect is known as the **Attentional Blink**);
4. Suppress attention during the change (**Blink** and **Saccadic Suppression**).

Each of these mechanisms might allow a cut to occur without the cut itself attracting viewer's attention. As this is the intention of the continuity editing rules there should be evidence that some, if not all, of these mechanisms are utilised by continuity editing. Defining each of the mechanisms and finding existing evidence that

[16] These categories of capture avoidance mechanisms have been identified during the development of this thesis. The attention phenomena associated with each category are widely established but they have not previously been brought together in this way to create a system of capture avoidance. It is not known if this list is exhaustive as there does not currently exist an absolute understanding of attentional mechanisms such as capture. However, the phenomena listed are as up-to-date as possible and could always be supplemented later as understanding of visual attention expands.

continuity editing creates viewing conditions under which these mechanisms occur will be the focus of this chapter.

3.1 Focus on a constant

Attention is captured by visual transients. These transients are sudden changes of the visual scene such as abrupt onsets or changes of objects, motion, size, colour or contrast, and apparent rotations and relocations (see 3.4.2.2 for a summary). These transients capture attention to varying degrees depending on what other transients exist at the same time, how much attention is available for capture, and whether they match the viewer's attentional set (Simons, 2000). These factors can be manipulated to limit the probability that a transient captures attention. Evidence that continuity editing manipulates these factors (existing transients, available attention and compatibility with attentional set) will be presented in the next sections of this chapter. However, there is also another technique that can be used to limit attentional capture: limit the saliency of the transients. If there are no transients there can be no attention capture evoked by the visual scene. In the context of film editing, if a cut does not create any visual transients that capture attention then the cut itself cannot be said to have captured attention. Without attention capture there can be no sudden awareness of change other than that deduced by referencing memory for the visual scene before the change (see (Simons & Levin, 1997) and the area of Change Blindness for further discussion). The cut itself and its effects on the visual scene could be said to be "invisible". As this is exactly the goal of continuity editing as stated by editors (Bordwell & Thompson, 2001; Dmytryk, 1986; Katz, 1991; Reisz & Millar, 1953) we would expect to find evidence that continuity editing sometimes limits these transients.

3.1.1 Dissolve

Figure 3-1: A dissolve from a shot of a painting to a shot of the sea. Camera moves towards the picture (top left) then a dissolve begins (top centre and right) to a shot of the sea (bottom left, centre) until the shot of the sea completely replaces the original shot (La Sindrome di Stendhal, Dario Argento, 1996)[17].

The most obvious technique used to minimise the visual transients caused by transitioning from one shot to another is to dissolve or fade between the shots instead of cutting. A cut presents the new shot immediately after the old shot. This sudden change between visual scenes is what creates the visual transients. To remove the transients this transition needs to be more gradual. A dissolve produces this effect by gradually fading out the first shot whilst fading in the second (see Figure 3-1). A variant of a straight dissolve is a fade. This is created by placing an intermediate set of blank frames between the two shots. Typically these frames are coloured black so the fade is known as a fade-to-black but occasionally other colours are used such as a fade-to-white to imitate looking at a bright light or symbolise transcendence. A fade typically takes longer than a straight dissolve and indicates a clear juncture between the two shots. This difference in actual duration of the transition is used by editors to symbolise different degrees of temporal ellipsis between the shots. The conventionalised meaning of a dissolve is that a short period of time has passed between the shots whilst a fade through black indicates a long period of time (Lindgren, 1948).

[17] A video of this dissolve is available from http://classes.yale.edu/film/videos/stendhal-dissall.wmv

A dissolve limits the attentional capture caused by the cut as the objects of the new scene do not appear suddenly. Instead the viewer's attention would be distributed across the shot based on the content, not the transients caused by the cut. However, the absence of attentional capture does not mean that viewers are not aware of the change of shot. In fact the extended duration of the transition makes it more likely that viewers become aware of the changing visual scene.

It is interesting to note that in the earliest days of film editing (pre 1910) a straight cut from one shot to another was thought to be visually disruptive, uncomfortable for the viewer, and to be avoided at all cost (Bottomore, 1990). A lot of the earliest filmmakers employed fades or dissolves instead and maintained a preference for this practice throughout their careers (e.g. Cecil Hepworth; see Bottomore, 1990). By 1918, straight cuts had become more common but some filmmakers still believed that shot transitions "without warning and without intermediate change" meant that "the eye suffered a shock" (Croy, 1918; page 184). However, by this time editors were beginning to understand that under some viewing conditions cuts could be acceptable. This can be seen in Croy's conditional "without warning". By the time Croy made this statement, the basic principles of continuity editing, such as 180° Rule, reverse angle cutting, point-of-view shots, and, most significantly, matching action were well established. Whether because of these new conventions, or because viewers had become more familiar with the film medium, the cut became the accepted form of transition between shots and the belief that they resulted in a "shock" was retained only for editing discontinuities. Dissolves remained only to signify a passage of time or a symbolic connection between shots[18].

[18] However, if you perform a frame-by-frame analysis of a modern feature film you will still quite often find minute dissolves between shots. A couple of frames of dissolve does not appear to be enough for the audience to register (the transition looks like a cut) but editors must believe that even a slight dissolve such as this can ease the transition between shots. It might also be used as a way of hiding a discontinuity, i.e. a quick-fix.

The problem with using dissolves as a way to hide a change in shot is that wherever the viewer is looking eventually the visual information they are fixating will change. Dissolves are traditionally created in the photographic laboratory by exposing a single piece of film to two different shots. The degree of exposure of the shots is balanced so that gradually the second shot dominates the first. There is no optical transformation of one shot into another, just superimposition. The objects of the first shot will gradually disappear and be replaced by new objects. If we compare this visual experience to that of the inattentional blindness experiments it is clear that the absence of a visual constant across the dissolve makes it impossible for viewers to focus their attention without disruption. For example, in Simons & Chabris' Gorilla experiment viewers were instructed to follow the basketball players dressed in white. The presentation of these players was constant during the experiment even though they would occasionally be occluded by the basketball players in black and the unexpected gorilla (Simons & Chabris, 1999). The almost constant nature of the white players meant that viewers were able to adopt a "white" attentional set and focus their attention in a continuous fashion even when the stimuli was presented as two semi-transparent superimposed films (Simons & Chabris, 1999). By comparison, in a traditional dissolve there is no visual constant that isn't affected by the dissolve. For inattentional blindness to be observed across a dissolve (or a cut, for that matter) the part of the scene to which the viewer was focussing their attention would need to stay constant whilst the peripheral details change in a way that does not capture any remaining attention.

3.1.2 Graphical Match

There is an existing technique in continuity editing that achieves this level of constancy: graphical match. A graphical match involves the composition of the two adjacent shots in such a way that some or all of their graphical elements are similar e.g. general object shape, colour, shadows, patterns, etc. The graphically matched shots are usually joined by a dissolve so that the matched objects are seen to slowly transform. The two shots could be joined by a cut but the close similarity might

create the impression of a Jump Cut[19] as the focal object undergoes an unexpected apparent rotation, relocation, or deformation. This could be avoided by making the viewer expect the cut (how this is achieved will be discussed in the next section) or by changing the location of the focal object across the cut (i.e. eliminating the graphical match). There is a large potential for accidental graphical matches across most cuts considering that most shots depict a single focal object (e.g. a person) framed centrally (May et al., 2003; Tosi, Mecacci, & Pasquali, 1997). The fact that the majority of these cuts are not identified as Jump Cuts highlights the power of continuity editing to obscure these apparent transformations (the techniques employed to achieve this will be discussed in the next section).

Graphical matches with dissolves are designed to maintain the viewers' attention on a part of the visual scene (usually at the centre of the frame) whilst the rest gradually changes. Two examples of graphic match taken from the trailer for the recent remake of War of the Worlds (Spielberg, 2005) can be seen in Figure 3-2. The first sequence dissolves from a shot of space with a red planet moving towards earth, to a busy city street with a red traffic light. The second sequence uses a camera movement to match the Arc de Triomphe to an Arabian gateway. Both of these matches match the visual features of the most salient object within the scene[20] to maintain the viewers' focus across the dissolve.

---

[19] A sudden cut between two shots of the same object taken from slightly different camera positions. Typically an indiscernible amount of time is omitted between the shots leading to an apparent "jump" in the image as the object suddenly changes location or orientation.

[20] The red planet is salient as it is moving and red. Both of these properties are highly salient relative to the background. In the second sequence, the archway is the most salient as it is the brightest region of the screen and it is located in the centre of the frame.

Figure 3-2: Two graphical matches (one top line, one bottom, both running left to right) from the trailer for War of the Worlds (Spielberg, 2005). Both involve moving centres of attention which are (roughly) matched for their size, shape, colour, and position.

It is unlikely that viewers are unaware of the change of shot occurring across these two graphical matches as the focal objects noticeably change during the transition and occupy only a small proportion of the screen. The position of the red spot is not an exact match across the shots and the archways share so few visual features that the match is more on the general arch shape and its movement. However, an invisible transition was not the point of this dissolve. The dissolve produced a subtle bridge between shots that encourages the viewer to interpret the connection between the shots (e.g. we are all one world). Most graphical matches are used in this symbolic way.

Using graphical matches to hide a change of shot is much more difficult as it requires the focal object to be matched exactly during the dissolve. If the viewer is fixating an object the object's visual features, at least as much as is projected on to the viewers fovea (2° of a visual angle), cannot change unexpectedly during the transition. If they do, the viewer will become aware of the manipulations of the editor. Such close control of visual features is very difficult to achieve using traditional cinematographic techniques and, given that the focal object will usually be a person, very hard to act. Even if the focal object is precisely matched across the dissolve the viewer's point of fixation would have to be held on the matched object whilst the background changed. If their eyes wander over to the background they would be able to see the dissolve. To ensure that the matched object holds their attention a highly

salient object, such as a character's eyes or face (Yarbus, 1967) is commonly used. This is usually supplemented by a movement of the face of some sort, such as a blink, change in gaze or head turn which captures attention prior to the change in shot. This technique moves into the realm of match-action editing which will be discussed in the next section (see 3.2).

If the editor cannot guarantee that the viewer's eyes will not wander onto the changing background the only solution is to remove the background completely. This is achieved by zooming-in or moving the camera towards an object (e.g. a person's face or inanimate object) so that the object fills the screen. A cut can then be made to the same object in a different location and the change in background will not be registered as it cannot be seen on the screen. The camera is then zoomed or moved back and the new scene is gradually revealed. This technique is very effective at hiding the join between shots as it removes all evidence of the cut. However, it is a technique that is rarely used and usually only to induce a state of wonder and surprise in the viewer as the new scene is gradually revealed. The experience this type of transition induces in the viewer is different to most films shot for continuity in that it is more akin to a magic trick. The transition dupes the viewer into expecting one thing and then surprises them when something else happens (e.g. a sudden change in location). By comparison, a continuity cut is accompanied by no shock at the sudden change or at what is presented. It is as if the cut provides exactly what the viewer expected[21].

[21] This idea will be expanded later in section 3.2.

3.1.3 Digital Composites, CGI, and MoCo

Figure 3-3: A motion-control composite shot of a continuous camera movement around a car. The car rotates whilst the location and characters fade in and out (copyright www.mrmoco.com) [22].

Traditional graphical matches were limited in terms of the control that could be achieved over the graphical elements of the shots. Objects could be matched precisely across shots by careful composition, lighting, and filming of the shots but there always needed to be a point at which one shot ended and another began. However recent advances in digital compositing of shots, computer graphics (CGI), and motion control cameras (MoCo) now means that a finished shot can be an amalgamation of multiple filmed elements or computer graphics all filmed by a moving camera. For example, a single object can be filmed using exactly the same camera movement in multiple locations. When the resulting shots are composited together (combining visual elements from various shots into one shot) the resulting film is seen to depict a constant object filmed by a continuously moving camera whilst peripheral details change. An example of such a motion control shot can be seen in the shot sequence Figure 3-3[23]. These techniques blur the distinction between shots and raise interesting questions about the application of continuity editing rules to a film form without shots. Currently these techniques are most often seen in advertisements and music videos, both of which don't prioritise narrative in the same way films shot for continuity do. However, these techniques are beginning to enter

[22] A video of the moco shot can be viewed here: http://www.mocoforum.com/videos/octavia_ad.mpg .

[23] Motion control refers to the computer controlled camera needed to film such a shot. The camera movement has to be reproduced exactly in several locations and this would be almost impossible if a human were controlling the camera. The solution is to mount the camera on a motorised arm and program its movements into a computer. These movements can then be reproduced exactly as many times as is necessary.

into mainstream use in feature films such as The Hulk (Ang Lee, 2003) and Eternal Sunshine of the Spotless Mind (Michel Gondry, 2004). These techniques are beginning to undermine the foundations of the continuity editing rules and over the next few years the old conventions will have to adapt to accommodate the new graphical potential of film. From a cognitive perspective, what will be interesting will be the experimental dissection of the editing conventions to see which are essential for film as a medium of communication and which can be abandoned or modified without major sacrifice.

3.1.4 Summary

This discussion of the potential application of inattentional blindness to film editing has highlighted the difficulties involved its direct application. Inattentional blindness, as found in most experimental studies, requires a part of the visual scene to remain constant during the period of presentation. The viewer can then focus their attention to this part of the scene and fail to have their attention captured by visual transients occurring in their periphery. However, in film it is highly unlikely that the visual object fixated by a viewer prior to a cut appears in exactly the same screen location or in exactly the same visual form after the cut. This would mean that the object to which viewers were focussing their attention would bear signs of the cut, increasing the potential for awareness of the cut. Even if the focal object could be matched across a cut, the viewer's attention would have to be focussed on the object to such a degree that peripheral visual transients would not capture attention. How this is achieved will be discussed in the next section. If attention is not focussed, the visual transients created by the cut would either have to be obscured by dissolving between the shots or removed by filling the screen with the matched object. Both of these techniques are occasionally used as part of the continuity editing style but they are not as common as straight cuts between non-graphically matched shots. It appears that, whilst graphic matches are sometimes used to hide the transition between shots, the creation of most "invisible" cuts requires more than just limiting visual transients.

3.2 Expecting a visual change

Expectation is an important factor in deciding if a sensory event captures attention. When the viewer has no expectations attentional capture will occur "by default" in response to a range of sensory events, such as sudden changes in saliency or abrupt onsets or movements (Simons, 2000). However, the potential for capture associated with these events is modulated by endogenous control of attentional resources (Ruz & Lupiánez, 2002)[24]. If attention is engaged by focussing on an existing stimulus or expecting a certain type of new stimulus the likelihood that another type of will capture attention decreases (Simons, 2000). By adopting an attentional set the viewer is effectively filtering out all extraneous information and focussing their attention on what they believe to be important to the task (Most et al., 2005). However, if this attentional set contains precise predictions about the type of visual events that will occur, when they do not occur, or occur in an unexpected form this violation of expectation will also capture attention (Berlyne, 1971). This relationship between viewer expectation and the probability that a sensory event will capture attention provides a mechanism that can be utilised by film editors to either hide or highlight their manipulations.

A saccadic eye movement can only be triggered by one sensory change at a time (Theeuwes et al., 1998). When a cut occurs, even though the entire visual scene has changed, a viewer can only saccade to one of the newly presented objects. If the viewer did not expect the cut and its occurrence resulted in the abrupt onset of an unexpected object, the viewer's attention might be captured by the new object. As this capture is the result of the cut, not the viewer's attentional set, the possibility that the viewer would become aware of the editor's artificial manipulation would increase. By comparison, a cut to a shot that the viewer expected to see should lead to attentional capture in agreement with their attentional set. In this situation, the viewer should be able to resolve the new shot with the previous shot based on their expectations.

---

[24] Endogenous control means that it is based on decisions made by the individual not due to some external factor.

This ability of editing to satisfy viewer's expectations has been described by some film theorists as the "*fundamental psychological justification for editing*" (Lindgren, 1948; page 54):

> "[editing] reproduces the mental process…. in which one image follows another as our attention is drawn from this point to that in our surroundings. In so far as the film is photographic and reproduces movement, it can give us a life-like semblance of what we see; in so far as it employs editing, it can exactly reproduce the manner in which we normally see it." (Lindgren, 1948; page 54).

When viewing reality our saccadic eye movements present our perceptual system with a succession of views, all of which are presented in response to some form of perceptual inquiry (Hochberg & Brooks, 1978a). For example, the perceptual question "What is that man looking at?" is answered by a saccadic eye movement to the target of his gaze. At this point the question is answered by the object now occupying the centre of the viewer's attention. This perceptual question was endogenously answered [25]but a similar question could also be answered by the answer itself capturing attention (i.e. exogenous control). The "snap" of a twig off to your side whilst you walk through the woods elicits involuntary orienting to the source of the sound. The initial "snap" poses the perceptual question, "What made that sound?" which is answered by the eyes being captured by the cause. The same pattern of perceptual question and answering (referred to as Q&A for short; Katz, 1991) occurs whilst watching film. The main differences lie in the extent of reorienting and the locus of control: the eyes can never be directed beyond the screen edge and whilst the viewer may *want* an answer they can never *get* the answer unless the editor *gives* it to them. These differences may change the perceptual consequences of a filmic Q&A sequence compared with reality (this will be discussed in more detail in chapter 5) but in terms of the distribution of attention it is very similar.

---

[25] i.e. the viewer redirected their attention to answer the question.

Describing continuity editing as an analogue for the shifts of attention an observer would perform in a similar situation provides a mechanism by which the logic of the editing can match viewers' expectations. However, it does not necessarily provide a mechanism by which the cut itself fails to capture attention. In fact, as the reorienting of attention (that would normally happen) has been made redundant by editing, the viewer's attention may be highly susceptible to capture by the visual disruptions of the cut. If attention is captured by a feature of the scene that is of no importance to the logic of the cut the viewer might be more likely to become aware of the cut. To disprove this claim evidence would need to be presented showing that attention across a continuity cut moves directly to the answer of the perceptual inquiry without being captured by peripheral details. It would also need to be shown that continuity editing rules either identify time points at which viewers have a perceptual question that they expect to be answered or that the film itself constructs the question. Such evidence will be presented in the next section.

3.2.1 Attracting Attention with Motion

> "Excluding cuts made at the beginnings and ends of sequences and self-contained scenes, cuts to reactions or responses, and cuts involving exchanges of dialogue, the cutter should look for some movement by the actor who holds the viewer's attention, and he should use that movement to trigger his cut from one scene to another. A broad action, will offer the easier cut, but even a slight movement of some part of the player's body can serve to initiate a cut which will be "smooth", or invisible…. The important consideration here is that there be just enough movement to catch the viewer's attention." (Dmytryk, 1986, page 435-436)

This quotation from Edward Dmytyrk (film editor and director) shows his incredible insight into the psychology of attention capture. By instructing editors to use movement within the scene as a way of hiding a cut he is identifying one of the main visual events that has been shown to attract attention by default (Abrams & Christ, 2003; Franconeri & Simons, 2003). Dmytryk's faith in the power of movement to hide a cut has been so influential over the editing community that his insight is now

considered as editing gospel (see Pepperman, 2004). The classic name for the editing technique Dmytryk is referring too is Match-Action editing. Match-Action editing (see chapter 2) is one of the main techniques used to decide when to cut. Typically the examples used when describing a match-on-action cut use large actions e.g. a person sitting down, throwing a punch, or picking up a glass. However, as indicated in Dmytryk's quote, small movements such as head turns, shifts in gaze, even changes in facial expressions and blinks (Murch, 2001; Pepperman, 2004) can still attract viewer's attention and be used to hide the cut. These small actions are harder to use as the editor has less confidence that they will attract all viewers at the same time but if the shot has been composed in a suitable manner (e.g. the blink of an actor's eye cannot be used to attract attention when the actor is shot in Long Shot as the eye is not salient enough: a Close-Up must be used[26]) then they can be just as useful as larger actions. One of the defining qualities of a good editor is their ability to predict where the majority of the audience will be looking at any moment in a film. They do this by developing *"the uncanny facility to have your brain 'watch and note' your [own] eyes' automatic responses"* (Pepperman, 2004, page 11).

There are two ways movement can be used to hide a cut: to **attract** and to **direct** attention. The first, uses motion simply as a way of capturing attention, the benefit of which will be discussed in more detail later (see section 3.4.2). The latter, directing attention, constructs a perceptual inquiry in the mind of the viewer which the editor can then use to lead the viewer's attention across the cut. To direct attention a certain type of attention capturing visual event is needed. These will be referred to as **deictic cues**[27].

---

[26] Therefore, the type of action used to attract attention must change depending on the shot size and complexity of image. Evidence of this consideration can also be seen in the editing literature Katz (Katz, 1991; Pepperman, 2004).

[27] They are deictic as they gain their meaning depending on their context in a similar way to deictic words such as "I", "he", "there", or "now". In other words, they are referential.

3.2.2 Deictic cues and the power of the gaze

Deictic cues are a subset of associative cues: visual properties of a shot that connect it in some way to the following shot (Gregory, 1961). Associative cues can include environmental factors such as location, lighting, sound effects or even music flowing across the cut. Associative cues tell the viewer that the two shots occur in the same location and at the same time (more on spatiotemporal continuity later, chapter 5). However, some associative cues are actions within the scene that form causal or logical connections between shots. These will be referred to as **deictic cues**.

Figure 3-4: Two point/glance shots from *Blade Runner* (Ridley Scott, 1982)

The classic example of a deictic cue is the **point/glance shot**: a Medium or Full Close-Up depicting an actor's face looking off screen (Branigan, 1984; see Figure 3-4). The key element of this composition is the actor's eyeline. The eyeline automatically establishes a question in the mind of the viewer: *"What is the actor looking at?"*. This leads the viewer to expect the answer to the question in the form of a shot of the target of the actor's gaze (known as the point/object shot; Branigan, 1984). If the viewer does not know in advance what the character might be looking at they will adopt an attentional set for 'new objects'. When the new object is then presented their attention will move to it with minimum effort and the validation of their expectation will allow them to logically connect the two shots together. If the space of the scene had already been established and the viewer knew that by looking off screen right the character must have been looking at a particular character, they will expect the next shot to depict that character. If it does then their existing mental representation of the space is reinforced. If the shot depicts something else then their mental representation deteriorates and the viewer will be confused.

This ability to use an actor's gaze to create expectations about what they are looking at and use this to construct cohesive 3D representations of space (see chapter 5 for more details) is not an arbitrary skill learnt specifically for film viewing, it is a key element of human social behaviour. To an observer, another person's eyes are the most significant part of their body. The eyes will almost always be the first visual feature to be fixated when presented with an image of another human being even if that person is surrounded by a highly detailed background (Yarbus, 1967). This is believed to be due to the important role gaze plays in social interaction (Kleinke, 1986). Gaze has been shown to be used to regulate turn-taking in conversation: express intimacy, and exercise social control (Kleinke, 1986). When involved in a conversation, a listener will predominantly look at the speaker's eyes and mouth, with only occasional glances to their nose, ears and hair line (Yarbus 1967). When the speaker is addressing another person other than the viewer, the viewer's eyes will alternate between the eyes of the speaker and listener, looking for signs of the speaker's intention as well as the listener's responses (Klin, Jones, Schultz, Volkmar, & Cohen, 2002a). The accurate perception of other people's gaze is seen as critical to our ability to interpret other people's intentions, attribute the mental state of 'seeing', and structure social interactions (Baron-Cohen, 1995). Without the ability to read gaze these essential social skills disappear. This is most tragically seen in autism (Baron-Cohen, 1995).

This preference for seeking out eyes and following their gaze is developed during the first few months of life (Hood, 1998). The structure of the eye, a white sclera either side of a dark pupil, has been shown to give the human eye a strong salience making it "pop-out" of visual scenes (Bruce & Young, 1998). It has also been shown that the structure of the eye allows the direction of gaze to be perceived very simply and quickly (see Langton, Watt, & Bruce, 2000 for discussion). This prominence is associated with the eyes' function as a highly effective attentional cue. As was seen in the previous chapter (see section 2.3.3), the speed with which attention shifts to an object can be increased by previously cueing the object (Posner, 1980). These cues usually involve a sudden flash or onset of colour or object at the location being cued.

These visual transients capture attention even if the viewer is informed that the cue does not predict the location of the subsequent target (Posner, 1980). These cues can be said to *pull* attention to an object or location (Langton et al., 2000). By comparison, if the viewer fixates a human face and the cue takes the form of a shift in gaze, the viewer's attention is *pushed* towards the target of the gaze (Friesen & Kingstone, 1998; Hood, 1998; Jonides, 1981). Gaze is the only type of cue known to send a viewer's attention in a particular direction. Cues such as arrows presented at fixation have exogenous effect on attention other than to *point* in a direction where the viewer might choose to look (Jonides, 1981).

This potential for eyes to first *attract*, by "popping-out" of the visual scene, and then *direct* attention is critical for the use of gaze as a deictic cue in editing. If a viewer is fixating the eyes of an actor, when those eyes suddenly shift and point across the screen, the viewer's attention will be involuntarily pushed in the same direction (Driver et al., 1999; Friesen & Kingstone, 1998). The viewer's eyes will not move, as there is not yet a target for them to move their eyes to, but their attention will covertly shift in the direction of the gaze. This shift in attention combined with the viewer's ability to read intentionality into another person's gaze (Baron-Cohen, 1995) leads the viewer to expect a target for the gaze. The viewer adopts the attentional set of 'New Object'.

In real-world vision, the viewer would then use their cued attention to either locate an object in the periphery of their vision or move their head to locate an object out of view. They would then perform a saccadic eye movement to the first object they found that aligned with the gaze. In film, the same projection of the gaze through visual space will occur but it will stop as soon as it reaches the screen edge. If the target of the gaze is found within the screen a saccadic eye movement will be initiated (see left column, Figure 3-5). If no valid target exists the editor will have to provide one by cutting to the point/object shot. The object depicted in the point/object shot can either be located along the path of the actor's gaze, requiring a saccade to fixate (see middle column, Figure 3-5), or be collocated with the viewer's current point of fixation (see right column, Figure 3-5). In the latter case no saccadic

eye movement is required but attention will still be captured by the sudden onset of the expected object. The perceptual consequences of these types of cuts will be discussed in detail later (see chapter 5).

Figure 3-5: Shot pairs (paired top to bottom) occurring across cuts taken from *Blade Runner* (Ridley Scott, 1982). Two-shots (left column) in which both characters are visible in both shots (the shadow down the side of the frame is the other character; the viewer knows this from the establishing shot). Medium Close-Ups (centre column) with matching eye-lines. Collocated centres of attention (right). Red spots indicate most likely fixation positions in each shot. Red arrows indicate saccadic eye movements between these fixations.

An appreciation of the time taken for a viewer to shift their attention can also be seen in editing insights provided by Dmytryk:

> "The viewer, as a rule, will not accept the 'fact' of a look until he sees the actor's eyes focus, or 'freeze' on something off-screen. At that point he, too, will look off, following the actor's gaze. By the time his own eyes have refocused, the actor's point-of-view (POV) shot should occupy the screen. To make the cut, then, we fix the frame in which the actor's eyes have 'frozen', add three or four frames more to give the viewer time to react and move his eyes as he follows the actor's look, at which point the cut is made." (Dmytryk, 1986, page 444)

This three to four frame (125-167ms at 24 frames per second) wait before the cut matches the time taken to perform a saccadic eye movement (average 150-200ms Palmer, 1999). This similarity further reinforces the validity of Dmytryk's intuition about attentional capture.

3.2.3 Pull, push, and point across a cut

Other events can be used as deictic cues but none have the ability to *push* attention in the same reflexive manner as gaze. Instead they either have to be located where attention is to be drawn to (i.e. *pull* attention) or indicate a direction in which the viewer could choose to distribute their attention (i.e. *point* attention). A visual example of a *pull* cue is the sudden head turn by a character not currently fixated (Block, 2001; Dmytryk, 1986; Pepperman, 2004; Reisz & Millar, 1953). This is often seen to attract attention in film (Faraday & Sutcliffe, 1997) and can be used to hide a cut to a closer shot of the character. However, it is limited to directing attention within the space of the scene already on screen. By comparison, the *push* of a glance breaks out of the frame by directing attention into a space not currently depicted. This creates the perceptual question: "What are they looking at?". By comparison, a head turn poses no question as the target is already known (i.e. the character who attracted attention has already been seen)[28].

The best example of a *pull* cue that poses a perceptual question is a spatialized sound. A sudden sound generated so that it sounds as if it is coming from a space beyond the screen edge (e.g. with a surround sound system or in a cinema) automatically attracts attention (Pashler, 1998). This can be a very useful technique for eliciting a perceptual question in the viewer (e.g. "What was that noise?") but should be used sparingly as there is the chance that the sound's attention capturing powers might be too effective and actually direct the viewer's attention away from the screen. The best way to use spatialized sound is to direct viewer's attention to a character just out of shot. If a character speaks just out of shot, a quick cut can be made that relocates them onto the screen before the viewer has had chance to saccade

[28] The head turns referred to here are performed by characters that are not currently fixated. Therefore the head turn will capture attention, moving the viewer's eyes to the turning head. If the character performing the head turn was being fixated at the time and the head turn was seen as indicating a change in gaze direction then the head turn might *push* attention in the direction of the character's gaze.

off the screen. This technique of using sound to lead attention across a cut is a very powerful editing tool (Dmytryk, 1986)[29].

The other types of deictic cues are *point* cues. The most obvious example of a *point* cue is an outstretched arm or finger pointing to something off-screen (Dmytryk, 1986). Experimental studies using arrow cues show that pointing cues viewed at fixation do not involuntarily direct attention in the way that gaze does (Jonides, 1981). However, there is evidence that pointing does influence where a viewer's eyes will move to (Klin et al., 2002a) and where they believe the pointer's attention to be focussed (Langton & Bruce, 2000; Langton, O'Malley, & Bruce, 1996). For the target of the pointing finger to be attributed the most chance of capturing a viewer's attention it should also be the target of the pointer's gaze, head and body orientation (Langton et al., 2000). Whilst the successful comprehension of a pointing gesture is of less significance than gaze, an inability to successful appreciate the intention of a pointing gesture can be seen as another indication of social incapacity. This is also clearly seen in autism (Klin et al., 2002a).

In film, pointing gestures whether made by a character, action (e.g. a punch), or an object (e.g. a gun) are used to attract attention, create a perceptual question (e.g. "What is he pointing at?"), and form a spatial and causal link across cuts. Their ability to attract attention is principally due to their accompanied motion (e.g. a finger is usually pointed, a punch thrown, and a gun drawn). Editors understand the

---

[29] A quick survey of the editing literature will lead you to the conclusion that sound is a hugely important component of editing (Dmytryk, 1986; Murch, 2001; Pepperman, 2004; Reisz & Millar, 1953). In fact, some editors believe the key to good editing to lie in the harmonious (or disharmonious) marriage of sound and image (Murch, 2001). However, synchronised sound developed after the majority of the visual conventions of continuity editing had already entered mainstream use (Bordwell et al., 1985). As such, it is the belief of the author of this thesis that a survey of continuity editing is permitted to discuss only the visual components as these were the original foundations of the medium. Also, audio events effect factors such as attention and the perception of time and rhythm (which will become important later, see chapter 6) in different ways to visual events. As such a discussion of continuity editing must first consider the visual foundations of film before discussing their interaction with sound.

power of changes in motion to attract attention and, as indicated by Dmytryk's quote (see 3.2.1) use it across all cuts where they want continuity except for those already using sound to lead attention (Block, 2001; Dmytryk, 1986; Pepperman, 2004; Reisz & Millar, 1953). Capturing attention by using motion is useful for hiding edits but it does not do so by creating a perceptual question answered by a cut. As such, further discussion of this will be left for a dedicated section (see 3.3.1). Also, the spatial and causal connectives created by the *point* cues (and *pull* cues such as spatialized sound) are of most importance to the successful creation of a cohesive causal and spatial perception of events. Discussion of this will also be saved for a later chapter (see chapter 5).

3.2.4 Summary

Expectation provides an attentional and conceptual bridge across changes in viewpoint. In reality, a viewer's perceptual inquiry is answered by a reorienting of their senses. In film, the same change of viewpoint is accomplished by editing. The importance of associative, and specifically deictic cues in the continuity style indicates that film editors are aware of the need to answer a viewer's perceptual inquiry. If a viewer expects a certain type of visual event they will adopt the appropriate attentional set for this event. When the visual scene then changes it will be this visual event that captures their attention. Unlike unexpected attentional capture, capture by an expected stimulus allows the viewer to form a conceptual bridge across the viewpoints e.g. a question and answer, cause and effect, or a simple spatiotemporal relationship. These expectations occur, not as artificial constructs of the 'cinematic language' but as natural by-products of social/cultural perception. Changes in gaze, head and body orientation as well as directional cues such as pointing, motion, or projected events (e.g. a gun shot), direct attention both across the screen and across the cut via the creation of perceptual questions. If this question is answered by the new shot in the expected fashion (e.g. the target object is in the expected location on the screen) attention will move seamlessly from the old focal object to the new. Evidence for this seamless transference of attention across a cut will be presented in section 3.4.2.2.

What is currently not known about matching expectation across cuts is what kind of perceptual representation this results in. A viewer's expectations may be validated by the content of the new shot but the way that it was presented does not match any ecologically valid way of acquiring the information. For example, in reality when following a character's gaze to find its target a viewer would have to perform a saccadic eye movement and possibly even a head turn to locate the target. In film, the target would probably be collocated with the gaze across a cut requiring no saccade and no head turn. The film presentation roughly matches the information projected on to the viewer's retina after the target located but their experience of how it was acquired is completely different. It is assumed that the use of deictic cues leads to the accurate perception of the spatial relationship between shots (Bordwell et al., 1985; Dmytryk, 1986) but this cannot be the case if information about how the viewpoints were acquired is incorporated in the representation. This issue will be discussed in more detail in chapters 5 and 6.

3.3 Directing attention internally

Attention refers to the phenomenon by which some sensory inputs are processed faster or deeper than others, and thus become more readily available for action, memory, or thought (Egeth & Yantis, 1997; Posner, 1994). The usual interpretation of attention only refers to the first stage of this process: extracting the sensory information from the environment. As has been previously discussed in relation to the phenomenon of inattentional blindness (see 2.3.3), when attention is focussed on one type of sensory input, other salient sensory events fail to capture attention (Mack & Rock, 1998). This focussing of attention is not just limited to the extraction of the sensory input, it can also occur during encoding, storage, and, ultimately, conscious awareness of the input (Lamme, 2003). This broad application of attention from the external information to the internal conscious percept is usually viewed as a constant path of attention, starting external and ending internal (see Levin & Varakin, 2004 for a discussion). If attention is focussed on an external source of sensory information it is believed to remain focussed on that source throughout its processing. However, one visual phenomenon has begun to question this view of attention as "anchored" to the world: the **attentional blink**.

The attentional blink refers to the fact that perception of a second visual object is greatly reduced if it is presented within half a second of another visual target (Raymond, Shapiro, & Arnell, 1992). The phenomenon is usually studied using a **rapid serial visual presentation (RSVP)** visual search task in which subjects are shown a very rapid (~10 images per second) sequence of images at fixation and they are instructed to report targets of a specific type (Forster, 1979). If each image is presented for 100ms in isolation they can be accurately perceived (Lawrence, 1971). But if a second target image is presented within 500ms of a first target image the viewer has no awareness of the second image (Raymond et al., 1992). The phenomenon has been named the 'attentional blink' as it appears as if there is a period during which no attention is available for processing of a second target, similar to the way blinking keeps visual information from being processed whilst the eyes are closed (Raymond et al., 1992). The key difference being that, in an

attentional blink the 'blindness' occurs whilst the eyes are typically open and fixating the target (Palmer, 1999).

The attentional blink has been shown to specifically affect awareness of the second stimuli (Shapiro, Arnell, & Raymond, 1997). Whilst not reaching the level of awareness that would allow the subject to consciously respond to it, the second target is processed to the level of semantic interpretation (Shapiro et al., 1997). The second target has been shown to function as a semantic prime (Shapiro, Driver, Ward, & Sorensen, 1996) and trigger neural activation associated with the processing of meaning (Luck, Vogel, & Shapiro, 1996). This distinction between attending to a sensory input and becoming aware of it has also been seen in change blindness, specifically those indicating implicit change detection (Hollingworth & Henderson, 2002a) or implicit memory (Angelone, Levin, & Simons, 2003; Williams, Henderson, & Zacks, in press; see chapter 5 for more details).

This absence of awareness is exactly what is required for a viewer to be initially unaware of the visual disruption caused by a cut. The attentional blink suggests that if a viewer is dedicating attention to the processing of a visual event they will have insufficient attention left to have their attention captured by the visual transients of the cut. After the attentional blink is over the subject may be able to detect the change by comparing the new visual scene to their memory of the old scene but, due to the absence of attentional capture, this comparison will not be automatic (see section 5.2.1 for explanation of this process). Evidence that viewers are 'less aware' of continuity cuts compared to discontinuities has already been seen (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990). This suggests that continuity cuts might occur within 500ms (12 frames at 24fps) of an attention demanding visual event. To find evidence for this attentional blink hypothesis, similar changes in attention would have to be shown to occur during film perception and match what editor's would identify as valid 'edit points'.

3.3.1 Event Segmentation

Evidence that the attentional blink occurs during film perception can be found in recent work on Event Segmentation and Disruption Blindness. Event Segmentation refers to our tendency to 'parse' continuous visual actions into discrete events both intentionally and during normal viewing (Newtson, 1973; Newtson, Engquist, & Bois, 1977; Newtson & Enqguist, 1976; Zacks et al., 2001; Zacks, 2004; Zacks & Tversky, 2001). Disruption Blindness refers to our inability to recall whole-field disruptions (such as blank frames) inserted into a film depicting such events (Baldwin, Baird, Saylor, & Clark, 2001; Levin & Varakin, 2004; Newtson & Enqguist, 1976; Saylor & Baldwin, 2005). As this section will attempt to show, the time point at which one event is perceived as ending and another beginning are similar to the points identified by film editors as valid edit points. Editor's intuition that placing a cut at these points will make the cut invisible to viewers is supported by evidence from disruption blindness.

The ability to identify structure within events is an important part of perceptual behaviour. Our knowledge of event structures influences how we read, remember, and plan (Zacks & Tversky, 2001). This ability to decompose continuous activities into discrete events develops in infancy (Wynn, 1996) and, by adulthood, has developed to a level of consistency that if asked to repeatedly segment the same activity into its constituent events the same events will be identified (Newtson, Engquist, & Bois, 1976). Viewers can segment activities into events of different sizes (fine or coarse) and these are hierarchically related: groups of fine events corresponding to single coarse events (Zacks & Tversky, 2001). For example, if asked to segment the activity of "ironing a shirt", each stroke of the iron across the shirt, setting the iron down, picking the iron up, or lifting the shirt might be identified as individual fine events. By comparison, coarse events would probably be identified for groups of these events such as "ironing the left sleeve", "ironing the chest", or "folding the shirt". These events do not appear to be arbitrary as there is considerable agreement between the location of breakpoints across viewers (a breakpoint is the

point at which one event is identified as ending and another beginning; Newtson et al., 1977).

There is also recent evidence indicating that the same perceptual segmentation occurs during normal viewing when the viewer is not instructed to perform segmentation (Zacks et al., 2001). Neuroimaging studies have shown that the same brain regions are active during active and passive segmentation (normal viewing conditions when the viewer is uninformed of the segmentation task; (Speer, Swallow K. M., & Zacks, 2003; Zacks et al., 2001). This indicates that the segmentation of continuous visual activities into discrete events is a natural part of visual perception (Zacks et al., 2001).

Darren Newtson first developed the event segmentation task as a method for investigating the differences and similarities between how different viewers perceived visual activities (Newtson, 1973). His initial experiments indicated that viewers were very capable at performing segmentation (Newtson, 1973) and seemed to base their choice of breakpoints on significant changes in depicted motion (Newtson et al., 1977). Given that the films used depicted an actor performing a simple task, the changes in motion corresponded to a large number of changes in the position of the actor's body or limbs (Newtson et al., 1977). This correspondence between movement and breakpoints was most significant for fine breakpoints (points between small events; (Newtson et al., 1977). This relationship between changes in visual motion and perceived breakpoints has also been found by a more recent study (Zacks, 2004). Zacks created abstract animations in which two objects either followed random paths or exhibited intentional behaviour. He found that changes in motion reliably predicted where viewers would identify both fine and coarse breakpoints. This relationship was stronger for fine breakpoints and weakened as the size of the events increased (i.e. as the breakpoints became coarser). Zacks also found that viewer's inferences about the intentionality of the behaviours weakened the influence of movement features, specifically with larger events (Zacks, 2004).

Neuroimaging evidence also supports the involvement of changes in motion in the identification of breakpoints (Speer et al., 2003; Zacks et al., 2001). Neural activation was observed in brain regions identified as the Medial Temporal complex (MT+) and Frontal Eye Field (FEF) both during active and passive segmentation (Zacks et al., 2001). The MT+ complex is a visual area containing cells known to be sensitive to direction and speed of visual motion. The FEF is known to be involved in guiding saccadic and smooth pursuit eye movements. Its involvement in automatic and intentional event segmentation is less clear as subsequent neuroimaging studies in which the FEF was more precisely located have shown weaker activation under passive segmentation conditions (Speer et al., 2003). There is also no recorded activation in the brain regions normally associated with controlling shifts of attention (Speer et al., 2003). This seems to suggest that eye movements are not critical to the perception of breakpoints although Speer et al (Speer et al., 2003) do not rule out the connection.

This absence of a clear relationship between breakpoints and overt shifts of attention suggests that the end of an event may not provide a point at which the cut can be hidden. The evidence of the attentional blink seemed to suggest that the end of a visual event such as that identified during the event segmentation task would be accompanied by an absence of attention. This absence would provide a period during which the visual scene could be changed by a cut without the viewer becoming aware. However, the weak activation observed in the brain region controlling eye movements (the FEF) could be an indication that eye movements accompany some, but not all breakpoints (Speer et al., 2003)[30].

It is possible that a relationship between breakpoints and eye movements might be detected using a different recording technique. Tentative evidence for this relationship has been recently shown in an eye tracking study (Smith, Whitwell, &

[30] As the neural activation patterns are averaging across all breakpoints, different activation patterns for individual breakpoints is lost. This combined with the already weak activation observed under passive viewing might obscure the contribution of the FEF to event segmentation.

Lee, in press; Whitwell, 2005)[31]. By replicating Zacks et al (2001) methodology but replacing fMRI recordings with eye tracking, a significant decrease in saccade frequency was found 260ms prior to fine passive breakpoints followed by a significant increase in saccade frequency 140ms after the breakpoint (Smith et al., in press; Whitwell, 2005). No such effect of saccade frequency was observed after coarse breakpoints. The increase in frequency of saccades after fine breakpoints has been interpreted as visual search performed in response to the onset of a new visual event (Smith et al., in press). The time delay between the breakpoint and the increase in saccade frequency (140ms) is compatible with the average time taken to perform a saccadic eye movement in response to a sensory event (typically 150-200ms; Palmer, 1999). The potential for these eye movements to limit viewer awareness of cuts will be discussed in more detail in a later section (see 3.4.2).

As well as overt shifts of attention there is also the possibility that covert shifts of attention[32] or the reallocation of attention to cognitive processes, as in the attentional blink could occur during event perception. To find evidence of these shifts in attention, viewers' sensitivity to events need to be tested.

3.3.2 The Significance of Event Boundaries

Once Newtson had developed the event segmentation technique for identifying the perceived structure of visual activities he naturally wanted to test if there was anything significant about these events or if they were arbitrary. He did so using two techniques: 1) removing breakpoints or non-breakpoints from a film and testing if viewers perceived the deletions, and 2) creating filmic summaries of an activity by preserving either breakpoints or non-breakpoints and comparing viewers' memory for the two types of summary. The first method highlighted the significance of the

---

[31] Martyn Whitwell conducted this study as part of his MSc. in Informatics at the University of Edinburgh. His project was conducted under the supervision of the author of this thesis. A copy of Smith, Whitwell, & Lee (2005) is attached to this thesis as Appendix I.

[32] A "covert" shift of attention occurs when a viewer increases their ability to process a peripheral region of the visual scene without moving their eyes.

breakpoints by showing that their deletion was detected more accurately than non-breakpoints (i.e. periods during an event; Newtson & Enqguist, 1976). Viewers' inability to detecting deletions during events was later described as **disruption blindness** (Levin & Varakin, 2004). Newtson & Engquist's second method, creating filmic summaries, also supported the significance of breakpoints. Summaries constructed from breakpoints were described more accurately, rated as more coherent and recalled better (Newtson & Enqguist, 1976). This advantage of breakpoint summaries has since been replicated and supplemented by the finding that breakpoint summaries are recalled just as accurately as the original film from which the summary was constructed (Schwan & Garsoffky, 2004). This significance of breakpoints for communicating the events involved in a visual activity can be seen as support for the editing convention of temporal ellipsis: deleting the middle-point of events, such as moving between locations, and just presenting the beginning and ends.

For breakpoints to be recalled more accurately than information during events, more attention must be allocated to the encoding and storage of breakpoints in memory (Carroll & Bever, 1976; Schwan, Garsoffky, & Hesse, 2000). This has been observed in a secondary task reaction time (STRT) experiment in which viewers had to recall the location of audio tones whilst viewing a visual event (Baird & Baldwin, 2001). Tones located at breakpoints were recalled more accurately than tones located immediately before the breakpoint (Baird & Baldwin, 2001). STRT is seen as a good indicator of the extra attention left over after the primary task has used all that it requires (Lang & Basil, 1998). As such, the results on this study indicate that attention is fully occupied with encoding the event leading up to the breakpoint (leaving no attention to encode the tone) but after the breakpoint is reached and the event stored in memory, attention becomes available to encode the tone (Baird & Baldwin, 2001). This conclusion is compatible with the finding that single blank or sepia frames presented at breakpoints are detected faster than when presented during an event (Saylor & Baldwin, 2005). Such sudden wholefield changes[33] would usually

[33] i.e. changes that involve the entire visual scene.

capture attention (Posner, 1980) but during the event attention is not available for capture as it is focussed on encoding the event.

Evidence for this fluctuation of attention during event perception has also been shown by pupilometry (Smith et al., in press; Whitwell, 2005). During the eye tracking study previously discussed (see footnote 31) a significant increase in the size of viewers' pupils was recorded leading up to breakpoints (from 740ms prior to coarse breakpoints and 1000ms prior to fine). This can be interpreted as indicating high cognitive load (Hess & Polt, 1964). Given the existing evidence that breakpoints are recognised better than non-breakpoints (Newtson & Enqguist, 1976), this cognitive load can be attributed to the encoding of the visual event in memory. As suggested by the attentional blink, such a period of high cognitive load can result in an inability to process a secondary visual stimuli to the level of awareness (Raymond et al., 1992). This may suggest that a viewer may not be aware of a cut timed to coincide with the period of high cognitive load at a breakpoint. This possibility will be discussed in the following section.

For a brief period after a breakpoint (120ms for coarse breakpoints and 400ms after fine) the pupils contract back to their average size, indicating a sudden decrease in cognitive load (Smith et al., in press; Whitwell, 2005). This ties in to the evidence previously cited indicating that attention is available for encoding of secondary stimulus immediately after breakpoints (Baird & Baldwin, 2001).

3.3.3 Positioning a cut relative to an event

The evidence from event segmentation and disruption blindness previously cited leads to several conclusions:

1. Breakpoints are important for the successful perception, recall, and recognition of events (Newtson & Enqguist, 1976; Schwan & Garsoffky, 2004; Zacks & Tversky, 2001);

2. Attention is occupied with the encoding and storage of an event leading up to its expected end (Baird & Baldwin, 2001; Carroll & Bever, 1976; Schwan et al., 2000; Smith et al., in press; Whitwell, 2005);
3. For a 400ms period after a breakpoint attention is susceptible to capture by visual disruptions (Baird & Baldwin, 2001; Newtson & Enqguist, 1976; Saylor & Baldwin, 2005; Smith et al., in press; Whitwell, 2005).

Is it possible, based on these conclusions to suggest where, relative to a breakpoint, the best place to position a cut would be so that the viewer is not aware of the cut? There are four possible locations of a cut: during an event (**During)**, leading up to the end of an event (**End**), at the breakpoint between events (**Breakpoint**), or at the beginning of a new event (**Beginning**). The suitability of all of these will be addressed in turn.

3.3.3.1 During

Cutting during an event means identifying time points not adjacent to breakpoints. As breakpoints have been seen to correlate with periods of increased motion (Newtson et al., 1977; Zacks, 2004) time points in the middle of events are highly likely to involve an absence of visual motion i.e. stasis. These periods are also characterised by their insignificance to the successful perception, recognition, and recall of the overall event (Newtson & Enqguist, 1976; Schwan & Garsoffky, 2004; Zacks & Tversky, 2001). This means that there is probably no reason for attention to be allocated to encoding and conscious perception during these periods (Levin & Varakin, 2004). When attention is unfocussed it is vulnerable to capture by visual disruptions such as the whole-field disruptions caused by a cut (Simons, 2000). This is exactly as has been observed in disruption blindness experiments: visual disruptions ranging from 200-600ms motion-fields or blanks, down to single sepia or blank frames inserted during periods of stasis are perceived more often than when presented during motion (Levin & Varakin, 2004; Saylor & Baldwin, 2005). Without attention being internally occupied the sudden changes to the visual scene caused by

the cut, such as apparent motion of objects, is highly likely to capture attention (Folk et al., 1994).

There are also perceptual consequences of a cut during a period of visual stasis. Lang, Geiger, and colleagues have shown that all cuts have the potential to trigger an orienting response (Geiger & Reeves, 1993; Lang et al., 1993). This increases cognitive resources which, if the viewer is occupied with a cognitive task, will improve performance on the task (Lang, 2000). If the viewer is not occupied, such as when watching an insignificant part of an event, these extra cognitive resources will not result in the improved encoding of the visual event (Schwan et al., 2000). This indicates that a cut during an insignificant part of an event is not perceived as indicating a breakpoint in the event (Schwan et al., 2000). Instead the extra cognitive resources lead to increased awareness of the cut (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990).

In conclusion, all available evidence seems to suggest that cutting during an event when the visual scene is static will not hide a cut.

3.3.3.2 End

Placing a cut at the end of an event is better than during the event as the occurrence of visual motion is much higher (Newtson et al., 1977; Zacks, 2004). Cutting during motion has its advantages as the primary motion will attract attention and lessen the chance that viewers will become aware of a secondary visual disruptions such as a cut (Levin & Varakin, 2004; Saylor & Baldwin, 2005). However, if this secondary disruption interrupts the primary motion by making it jump ahead in time or space or pause momentarily, viewers may become aware of it (Baldwin et al., 2001; Newtson & Enqguist, 1976). To protect against this the moving object would have to be collocated across the cut and follow a continuous path (i.e. be graphically matched across the cut).

Deciding exactly when to cut prior to the breakpoint is very important due to the period of cognitive overload associated with the orienting response (OR). After every OR there is at least a 150ms period during which visual information cannot be processed (Geiger & Reeves, 1993). This increases to 300ms when the visual disruption is large (Geiger & Reeves, 1993). After this period of cognitive overload, extra cognitive resources allow for the increased encoding of new visual information but any information presented during the period of overload is lost (Lang, 2000). If the cut occurred less than 300ms (~ 8 frames) before a breakpoint, the visual information associated with the breakpoint would not be perceived. Given the importance of this information for the successful perception, recognition, and recall of the overall event (Newtson & Enqguist, 1976; Schwan & Garsoffky, 2004; Zacks & Tversky, 2001) a cut occurring in this position could lead to an inability to accurately comprehend the event.

In conclusion, cutting immediately prior to a breakpoint may stop the viewer becoming aware of the cut (if object motion and location is preserved across the cut) but there is the risk that it will have negative effects on comprehension of the overall event.

3.3.3.3 Breakpoint

A breakpoint is a transition from one event to another. Whilst the time around a breakpoint may be related to a high degree of visual motion (Newtson et al., 1977; Zacks, 2004) the breakpoint itself is probably a momentary point of stasis as one motion ends and another begins. The absence of any movement within the event to attract the viewer's attention could explain why the visual transients associated with a cut are ore likely to capture attention during a breakpoint compared to periods either side (Baird & Baldwin, 2001; Saylor & Baldwin, 2005). Cognitive processing of the previous event should also be coming to an end, as indicated by pupil contraction (Smith et al., in press; Whitwell, 2005), meaning that any extra cognitive resources made available by the OR may not benefit encoding of the previous event. However, there is evidence that when cuts coincide with breakpoints viewers' recall of breakpoints improves significantly compared with presentation of the breakpoints

without a cut (Schwan et al., 2000). This indicates that the increase in cognitive resources[34] is used to encode the breakpoint. However, this does not rule out the possibility that the attention capture caused by the cut does not also lead to awareness of the cut.

There is also another potential drawback associated with cutting at breakpoints. If an OR is triggered during the moment of stasis between events, by the time perceptual sensitivity has returned 150-300ms later the key change from stasis to motion signifying the beginning of the new event might have been missed. This might create an intact perceptual representation of the event prior to the cut but an incomplete representation of the new event. Critically, the transition connecting the two events might not be perceived. The importance of this transition for the perception of spatiotemporal continuity will be discussed in chapter 5.

In conclusion, the momentary stasis associated with a breakpoint increases the likelihood that a cut captures attention. However, the extra cognitive resources created by the OR improve recall of the breakpoint. What is not known is if the viewer is also aware of the cut and if the immediate cognitive overload associated with the OR hinders their perception of the connection between the old and new events.

3.3.3.4 Beginning

Cutting at the beginning of a new event benefits from the cut being hidden in visual motion (see 3.3.3.2), encoding of the previous breakpoint already being completed (Smith et al., in press; Whitwell, 2005), and the transition between the old and the new event already perceived. The OR triggered by a cut in this location might not lead to improved processing of the previous breakpoint but this is already recalled accurately without the co-occurrence of a cut (Newtson & Enqguist, 1976; Schwan et al., 2000). If a cut occurs at the beginning of a new event the viewer will be occupied

[34] Associated with the orienting response triggered by the cut (Lang, 2000).

with processing the transition of between the old and new events. This will be what benefits from the increased resources triggered by the cut. Also, the period of visual motion not perceived due to cognitive overload could be "filled-in" based on the start of the motion perceived before the cut and the motion perceived after the overload[35].

Therefore, whilst cutting at a breakpoint improves recall of the breakpoint, cutting at the beginning of a new event ensures that the new and old events are fully processed whilst also hiding the visual disruption of the cut in motion.

3.3.4 Does this match editor's intuitions?

This section has shown evidence that attention is not a constant during event perception; it fluctuates depending on the significance of the visual information to the successful comprehension of the event. These fluctuations provide periods in which the attention capturing potential of cuts can be limited or utilised to improve cognitive processing in a way that increases memory for the events whilst limiting awareness of the cut. The key question is: do editors identify the same points during visual events as the optimal places to cut?

The technique of match-action editing has already been discussed (see 3.2.1 and 2.1.5) but the exact timing of the cut relative to the action was not described. If the definition of match-action editing found in editing textbooks is read in detail a consensus of exactly where to cut emerges. Reisz and Millar (1953) provide by far the most in-depth discussion of the practice and intention behind match-action editing. They begin by stating that *"By showing one specific movement in long shot and the other in mid-shot, the cut does not interrupt a continuous flowing movement, but is, so to speak, punctuating the whole action at the moment of rest"* (Reisz & Millar, 1953; pages 217-218). They suggest that a cut at this location, identified by Newtson as the breakpoint between events, is acceptable but that it creates the

[35] Although it has been argued that such "filling-in" is not required if the significant elements of an event, such as the breakpoint and transition to the new event are perceived (Dennett, 1991; Levin & Varakin, 2004).

"impression of……two distinct phases of the movement… seen in two distinct ways" (Reisz & Millar, 1953; page 218). This implies that the cut separates the two events. This can be attributed to the cognitive overload caused by the OR (see 3.3.3.3).

Cutting at the momentary pause between events is the easy way to create a match-action cut but it is not the optimal. If the editors intention is to make the cut as "smooth" and "dramatically appropriate" as possible they will cut just as the new action begins (Bordwell & Thompson, 2001; Dmytryk, 1986; Katz, 1991; Pepperman, 2004). This is believed to be "smooth" as "it will coincide with the moment of change from rest to activity" and it is "dramatically appropriate" as "the spectator will anticipate what is about to happen and will be ready to see the effect of the resolution in another shot" (Reisz & Millar, 1953; page 218). In other words, the visual disruption of the cut is hidden in visual motion and the cut to another viewpoint matches the viewer's perceptual inquiry. As a result the cut is not an obstruction to the viewer's perception of the event but an aid.

3.4 Suppress attention

So far during this chapter three techniques have been presented that enable an editor to hide a cut. By using a graphically matched dissolve or composite of multiple shots the visual transients associated with the cut can be removed. Without visual transients the cut cannot capture attention. However, this technique requires a lot of effort to get right and so other, easier techniques are more commonly used. These either direct attention across the cut (see 3.2) or internally to the processing of previously attended visual events (see 3.3). Either way, attention is occupied during the cut limiting the degree to which the visual transients capture attention. The attribution of these two techniques to two of the main rules of continuity editing, matching eyelines and action across cuts, indicates how integral they are to the continuity style. However, there is an even simpler way to hide a cut: cut when the viewer isn't looking.

Active control over a viewer's gaze lies at the heart of magic. This art of **misdirection** allows a magician to direct their audiences' attention away from the manipulations and switches at the heart of their magic (Lamont & Wiseman, 1999). Without seeing the secret switch the audience believes the magic transformations implied by the trick. A magician's mastery over another person's attention is a wonder to behold and only recently has the extent of this control begun to be understood (Kuhn G. & Tatler, 2005; Lamont & Wiseman, 1999). Magicians utilise the sensory events that reliably capture attention such as sudden appearances (e.g. a puff of smoke), or movements (e.g. a flamboyant wave of the wand) to direct their audience's attention both overtly and covertly away from their secret manipulations (Kuhn G. & Tatler, 2005). A film editor's use of the same captivating events to direct covert attention has already been discussed (see sections 3.2 and 3.3). However, if a cut is not a precise graphical match (which most are not) overt attention cannot be directed away from the editor's manipulations as it can in magic as the evidence fills the screen. However, unlike stage magic, the manipulations that need to be hidden in film do not take time to accomplish. A viewer's attention does not need to be sent to a safe location for a period of time whilst the manipulation is made as the

manipulation occurs almost instantaneously across a cut. All that is required is a perceptual "hole" the duration of about one frame in which the visual transients of the cut can be hidden. Luckily, the human perceptual system provides two "holes" that might do the job: blinks and saccades.

3.4.1 In the blink of an eye

A blink is a reflexive, rapid closure of the eyelid occurring 10 to 15 times a minute (Burr, 2005). We blink reflexively to protect our eyes from potential damage (e.g. during a sandstorm), and to moisten and oxygenate our corneas (Burr, 2005). A reflexive blink usually involves the simultaneous closure of both eyes but intentional blinking with one eye (i.e. a wink) and both eyes is also possible. Each blink lasts 100-150ms during which time the eyelids stop most light from entering the eyes. This creates a period of transient whole-field luminance change that would usually be highly salient (Bristow, Haynes, Sylvester, Frith, & Rees, 2005). However, we are not generally aware of the absence of light associated with a blink. To explain this absence of awareness a blink suppression mechanism has been proposed (Volkmann, Riggs, & Moore, 1980). This proposes that a signal associated with the blink motor command acts to suppress the activation of the brain regions responsible for processing the visual transients created by the closing eyelid. Recent neuroimaging experiments have found supporting evidence for this hypothesis (Bristow et al., 2005). The result of this suppression is a 100-150ms period during which no visual transients can capture attention, no matter how salient the transients are during normal vision.

The suppression also acts to limit the viewer's awareness of the blink itself (Bristow et al., 2005). This has been most elegantly shown in Change Blindness studies. If a viewer is instructed to memorise a photograph and a feature of the photograph spontaneously changes during a blink the viewer is highly unlikely to notice the change (O'Regan, Deubel, Clark, & Rensink, 2000). In fact, O'Regan et al (2000) found that even if the viewer was fixating the changed object before and after the blink they still failed to detect the change more than 40% of the time. It has been

proposed that this failure is due to an incomplete or impoverished representation of the visual scene retained across the blink (O'Regan, 1992). However, this theory has been invalidated by recent evidence showing that viewers exhibit viewing behaviour[36] indicating that they implicitly detect the changes even if this detection does not reach the level of awareness (Brockmole & Henderson, 2005; Brockmole & Henderson, in press; Henderson & Hollingworth, 1999; Henderson & Hollingworth, 2003; Currie, McConkie, Carlson-Radvansky, & Irwin, 2000). For implicit change detection to occur a reasonably detailed representation of the visual scene must be retained across blinks (and saccadic eye movements; see chapter 5).

Even if a viewer implicitly detects a change made during a blink they are still unaware of the change. This suggests that, in theory, a blink could be used to hide a cut. If an editor was able to predict when a viewer was going to blink they could time the cut so that it occurs during the 150ms when the viewer's eyes are closed. When the viewer then regains perceptual sensitivity after the blink the visual transients of the cut have passed and the shot has changed. Whether the viewer then became aware of the change in shot immediately or after a few seconds is unimportant. All that is important is that the editor has achieved their goal of an "invisible" cut by hiding the visual transients associated with the cut.

3.4.1.1 The Cut = A Cinematic Blink

Almost forty years before O'Regan and colleagues were exploiting the potential of blinks for hiding visual disruptions, the film director John Huston was discussing exactly the same technique in relation to film.

"All the things we have laboriously learnt to do with film, were already part of the physiological and psychological experience of man before film was invented…. Let me make an experiment – maybe you will understand better what I mean. Move your eyes, quickly, from an object on one side of the room to an object on the other side.

[36] Implicit change detection is usually indicated by an increased frequency of eye movements to the changed object.

In a film you would use a cut. Watch! There- you did exactly what I expected: in moving your head from one side of the room to the other, you briefly closed your eyes. Try it again, in the other direction. There! You see, you do it automatically. Once you know the distance between the two objects, you blink instinctively. That's a cut….. In the same way, almost all devices of film have a physiological counter-part" (John Huston quoted in; Bachmann, 1965)

This quotation takes the idea of film as an analogue for attention (previously seen in section 3.2) and expands on it by adding the observation that our eyes blink during an eye movement. This observation is correct, blinks do frequently accompany eye movements but not always (Fogarty & Stern, 1989). The actual cause of the perceptual separation of a saccadic eye movement into two static views of the visual scene is saccadic suppression. This will be discussed in more detail in the next section (3.4.2).

However, the idea that a blink could be used to hide a cut struck a chord with some editors. The most significant of these was the editor of Apocalypse Now (Francis Ford Coppola, 1979) and The English Patient (Anthony Minghella, 1996), Walter Murch. In his book: "*In The Blink of an Eye: A Perspective on Film Editing"* (Murch, 2001) Murch outlines his intuitions about how the blink represents the juncture between film and our experience of reality. Murch's thesis is not outlined in a systematic manner (this is not to be expected from such a discursive book on film editing) instead he recounts his intuitions about blinks and their relationship to film in an anecdotal fashion. However, by examining his discussion his main arguments can be extracted:

1. A blink functions as punctuation to cognitive events, i.e. thoughts (page 61-62).
2. Blink rate increases as cognitive activity increases (page 62).
3. A cut is the cinematic equivalent of a blink (page 62-63).
4. Therefore, the rate of cutting should match the rate of blinking were the viewer to experience the action in real-life (page 62).

5. We blink in synchrony with other people when we are sharing their thoughts (page 64-65).
6. When all members of the audience are engaged with the film they will blink in unison (page 71).
7. If all of these factors are taken into consideration, blinks can be used to hide cuts (pages 59 and 69).

This is an interpretation of Murch's ideas and it is not known whether he would agree with this interpretation or believe that these effects would occur across the majority of cuts. However, as a thought experiment it is useful to take these ideas as hypotheses and try to find some supporting evidence.

Hypotheses 1 and 2 suggest that blinks can be used as a measure of cognitive activity. Murch himself referred to the work of "Dr. John Stern of Washington University in St. Louis" to validate this claim and a look at Stern's publications does provide supporting evidence (Fogarty & Stern, 1989; Fukuda, Stern, Brown, & Russon, 2005; Nesthus & Stern, 2002). Stern and colleagues have presented evidence that blink duration and frequency increase as the complexity of cognitive processing required to perform a task increases (Fogarty & Stern, 1989; Fukuda et al., 2005). This relationship is believed to occur only with tasks that do not involve the acquisition of visual information. When a task requires attention to be concentrated on external visual stimuli, such as reading, blink rate decreases significantly (Bentivoglio et al., 1997; Holland & Tarlow, 1975). By comparison, when engaged in a conversation speakers have a tendency to blink in between phrases and at the end of sentences (Holland & Tarlow, 1975). Whether this is due to the blink acting as punctuation to the speaker's thought processes or as a form of social attention used to emphasise the speaker's point (e.g. in combination with a gaze shift towards the listener) is not currently known. However, Stern does believe that changes in blink duration and frequency can be used as a measure of cognitive activity (Fogarty & Stern, 1989; Fukuda et al., 2005).

Murch's third hypothesis equates cuts with the individual thoughts a viewer would have if they were observing the same action in real-life. This relates to the idea of film as an analogue for the viewer's attention and thought processes. Murch references another quote from John Huston to summarise this idea:

> "To me, the perfect film is as though it were unwinding behind your eyes, and your eyes were projecting it themselves, so that you were seeing what you wished to see. Film is like thought, it's the closest to thought process of any art." (John Huston quoted in Sweeney, 1973).

Whilst blinks do show a degree of dependency on cognitive activity they are not the most informative ocular activity. Pupil dilation, as previously discussed (see section 3.3.2) can be used as a direct measure of cognitive load (Hess & Polt, 1964) and saccadic eye movements provide a real-time measure of which visual information is being processed and how it is being related to other parts of the visual scene (Ballard, Hayhoe, Pook, & Rao, 1997). In fact, the most reliable association of blinks to cognitive activity is actually seen in their co-occurrence with saccadic eye movements. When viewers make large saccadic eye movements they are usually shifting from processing one part of the visual scene to processing another. The probability that a blink co-occurs with a saccade increases as the length of saccade increases (Fogarty & Stern, 1989). Therefore, Murch's hypothesis that *"A cut is the cinematic equivalent of a blink"* (Murch, 2001; page 62-63) might actually express his intuition that a cut should represent a significant change in viewpoint, analogous to a large saccadic eye movement, and will therefore co-occur with a blink. The blink is a by-product of the change in viewpoint which serves as an external manifestation of the internal saccadic suppression. It is this saccadic suppression that is actually responsible for making the viewer blind to the changing scene during the saccade, not suppression due to the blink. If this blink-saccade combination is actually what Murch is referring to then every cut would have to coincide with a saccade if the viewer were to be blind to the cut. Whether this occurs will be discussed in the next section.

Murch's fourth hypothesis builds on this idea that a blink is an index of cognitive activity and suggests that a scene's cutting rate should reflect the blinking rate a viewer would have if they were to engage in the same activity in reality. This seems like a valid claim if the blinks Murch is referring to are those that occur during large saccades. If a viewer needed to make numerous large saccades whilst observing a scene then, if the filmmaker' aim was to replicate this experience in the film they would also need to use the same number of shots to capture the scene. Each large saccade will probably be accompanied by a blink making it analogous to the cut between shots. This relationship between blinks and cuts is useful to an editor as it emphasises the viewer's comprehension of the scene. If a shot is presented for less time than a viewer would choose to focus an analogous part of a real-world scene there is the chance that they will fail to fully comprehend the shot. Of course, a limited or skewed comprehension might be the intention of the editor but to accomplish this the editor must first understand how long the viewer needs to fully comprehend a shot. These time constraints are dependent on the ocular activity required by the shot (e.g. fixation or visual search) and the associated patterns of attention and suppression.

Murch's fifth and sixth hypotheses claim that we blink in synchrony with another person when we are deeply engaged with them. The survival benefits of such behaviour are easy to imagine. If we are confronted with an aggressor we do not want to provide them with a period of time during which we are unable to respond to an attack even if this is only 150ms. If we timed our blinks to coincide with that of an aggressor we could ensure that our eyes received the cleansing they require whilst not letting down our guard. There may also be a social advantage to mirroring another person's blinks. Our ability to automatically replicate other people's facial expressions is well known (e.g. Miller, 2005) and it is believed to be an important tool for expressing intimacy and developing social relationships. Mirroring another person's blinks might serve a similar purpose. However, a quick survey of the social intelligence and facial expression literature has produced no empirical evidence that such mirroring of blinks occurs. This could be because a targeted study has never

been performed meaning that Murch's insight might actually exceed that of social psychologists. Future studies must address this question.

However, there is evidence that blinks **do not** coincide with edit points. It has previously been found that the breakpoints between visual events (or just after) are valid edit points and are identified as such by editors (see section 3.3). These breakpoints have been associated with the encoding of the old event in memory (Newtson & Enqguist, 1976) and the acquisition of information about the new event (Smith et al., in press; Whitwell, 2005). These changes in cognitive activity resemble those identified by Murch as being associated with blinks. If blinks do coincide with edit points as hypothesised by Murch there should be a noticeable increase in the frequency of blinks around the time of breakpoints. In an eye tracking study of event segmentation behaviour (carried out in conjunction with the author of this thesis), no such relationship was found (Smith et al., in press; Whitwell, 2005). Breakpoints were accompanied by a contraction of the pupils (indicating a decrease in cognitive load) following by an increase in saccade frequency (indicating visual search) but there was no significant increase in blink frequency or duration. If all viewers blinked at the same time (as proposed in Murch's sixth hypothesis) this would show up as a clear increase in blink frequency.

However, the absence of blinks does not completely invalidate Murch's hypotheses as the stimuli used for this study were composed so that the actor's faces, expressions, and blinks were mostly indiscernible (Long Shots were used). If blink synchrony is a social behaviour then closer shots that place more emphasis on actors' faces or involve the viewer more (e.g. dialogue sequences rather than action) might lead to an increase in blink synchrony. This possibility must be explored in future studies.

Murch's final hypothesis ties all other hypotheses together and proposes that blinks can be used to hide cuts. This hypothesis can be assumed to be true based on the findings of Change Blindness. It is not assumed that viewers will fail to detect changes of all size, change blindness experiments usually change an individual

object where as a cut changes the entire visual scene, but the evidence does seem to indicate that even if the change is eventually detected it will not capture attention immediately after the blink. Therefore, even given the absence of current evidence that blinks do coincide with cuts, the potential for blinks to hide cuts does exist. If an editor tries to coincide cuts with blinks the chances of achieving an "invisible" cut will increase.

3.4.2 Saccadic Suppression

Whilst blinks occur 10 to 15 times a minute and are mostly reflexive, saccadic eye movements occur 2 to 3 times a second and are mostly deliberate. This makes the proportion of perceptual "holes" created by saccades greater and more predictable than blinks. If, given this knowledge, an editing system was to be designed that aimed to hide cuts by coinciding them with periods of perceptual suppression it would be much easier to design a system based on saccades than it would blinks. Because of this it is believed that a search for continuity editing rules that utilise saccades will be more fruitful than the previous search for blinks. Whether this proves to be the case will be discovered during this section.

The possibility of hiding cuts in saccadic eye movements has already been discussed in section 3.2. Two techniques were described: **attracting** and **directing** attention. The second technique uses deictic cues to construct a perceptual inquiry in the mind of the viewer which is then answered by a saccadic eye movement. This technique has been shown to be utilised by the 180° Rule, Directional Matches, and Point-Of-View shots triggered by changes in gaze. By leading the viewer to expect a change in the visual information the distracting effects of the cut are minimised. However, for this technique to work it relies on the viewer to adopt the right perceptual inquiry. This is assisted by the reflexive effects of the deictic cues but it is still not guaranteed.

An easier way of using saccades to hide a cut is to attract the viewer's eyes to a part of the screen and then cut whilst their eyes are moving. Similar to during blinks, our

ability to perceive visual information is suppressed during saccades. This is essential as, whilst the eye movement lasts for only a very brief period (about 30ms), during that time the eyes can reach speeds up to 900° per second (Goldberg, Eggers, & Gouras, 1991). As the majority of saccades occur with the eyes open, the light entering the eye sweeps across the retina at the same high speed as the eye movement. If the same rapid motion of a visual field is created artificially whilst the eyes remain still (e.g. by moving an image in front of the eyes)[37], the viewer perceives a "blurring" of the image and feel their balance disturbed (Ross, Morrone, Goldberg, & Burr, 2001). No such perception occurs when making a saccadic eye movement and, in fact, we are rarely aware that our eyes have even moved (Burr, 2004). This has been attributed to a process known as saccadic suppression.

Saccadic suppression affects our ability to perceive visual information over the entire course of a saccadic eye movement including the periods of preparation and recovery either side of the actual movement. Around 75ms prior to a saccadic eye movement our ability to accurately perceive the visual world begins to decrease (Diamond, Ross, & Morrone, 2000). This decrease is maximal at the point of motion onset and then recovers around 50ms after the saccade ends. In total, a voluntary saccade takes at least 150-200ms to plan, execute, and for perceptual sensitivity to return back to normal (Goldberg et al., 1991). Any saccades completed less than 200ms after the saccade target was first presented are interpreted as initiated in anticipation of the target (Rose, Feldman, Jankowski, & Caro, 2002).

The exact mechanisms by which saccadic suppression functions are not fully understood but there is evidence that it is not a complete obstruction of the visual system (Burr, Morrone, & Ross, 1994). Instead the suppression seems to be isolated to the blurring and fast motion created by the visual information falling on the retina during the saccade (Burr, Holt, Johnstone, & Ross, 1982; Burr et al., 1994; Ross et al., 2001). As these are the only visual components received by the retina during a saccade, as long as these are suppressed no visual information will be perceived. The

[37] A very rapid pan of a camera can also create this effect when the resulting film is projected. This kind of camera movement is referred to as a 'whip pan'.

continued processing of other visual components (such as fine static details) can be shown by stabilising an image on the retina during a saccade. As this image does not contain the suppression components it will still be perceived (Matin, 1974). The clearest example of this effect is the 'train illusion'. When looking out the window of a fast-moving train the world streaks across your retina creating an indistinct image but if you saccade against the direction of travel the image is momentarily stabilised on the retina and the world can be perceived (Ross et al., 2001; pg 114). During normal vision the visual information projected on to the retina after a saccade masks[38] any visual components not suppressed (Dodge, 1900; Woodworth, 1906; Campbell & Wurtz, 1978). If the visual information after the saccade is of lower quality than during the saccade (such as the blurred foreground seen out a train window) no masking will occur and the information presented during the saccade will be perceived. These two components of saccadic suppression, selective suppression of parts of the visual pathway and masking by the post-saccade visual scene are still only possible explanations of how saccadic suppression occurs. However, the fact that perceptual sensitivity decreases during a saccade is not disputed and can be relied upon every time a saccade is performed.

This reliability made saccades the ideal tool for hiding the manipulations of change blindness studies. As with blinks, viewers often fail to notice changes to visual scenes when they occur during saccadic eye movements (e.g. Currie et al., 2000; Henderson, 1997; Henderson & Hollingworth, 1999; Irwin, 1991). By making the change coincident with the saccade the visual transients associated with the change are hidden forcing the viewer to rely on memory to detect the change (Brockmole & Henderson, 2005; Brockmole & Henderson, in press; Henderson & Hollingworth, 1999; Henderson & Hollingworth, 2003). If the viewer is not explicitly instructed to look for a change they may not compare the new visual scene to their representation of the scene pre-saccade (Rensink, O'Regan, & Clark, 1997; Simons & Levin, 1997).

[38] Masking refers to the when two sources of visual information are presented in quick succession and (typically) the second source obstructs the processing of the first. For a comprehensive review see Breitmeyer (Breitmeyer, 1984).

Instead they appear to assume that the world is constant and adapt their representation to the new visual scene (Hollingworth & Henderson, 2002b).

This assumed continuity might be exactly what is required by film editors to change from one shot to another without obstructing their viewer's perception of the action. Discussion of exactly how this continuity is perceived will be reserved for chapter 5. For now what is important is the fact that saccadic eye movements can be used to limit viewers' awareness of a change to the visual scene. If this potential of saccades is being utilised in continuity editing there should be evidence of this within the editing literature and in the eye movements of film viewers. These will both be discussed in turn.

3.4.2.1 Furthering the attention analogy

The belief that a film edited for continuity presents an analogue of visual attention is widely held (see D. W. Griffith in Jesionowski, 1982, page 46; John Huston in Bachmann, 1965 and Sweeney, 1973; Dmytryk, 1986; Katz, 1991; Lindgren, 1948; Murch, 2001; Münsterberg, 1970; Pepperman, 2004; Reisz & Millar, 1953). However, only a few theorists thought to question the place of cuts within this analogy (John Huston in Bachmann, 1965, and Sweeney, 1973; Murch, 2001; Reisz & Millar, 1953) and even fewer were able to apply a true understanding of saccadic eye movements to answer this question (Block, 2001; Dmytryk, 1986; Pepperman, 2004).

> "… if a cut, lasting one twenty-fourth of a second can be made whilst a viewer consumes one-fifth of a second moving their eyes, the cut will pass unnoticed. The trick is to get the viewer, or an audience of viewers, to move their eyes, en masse, at the desired instant." (Dmytryk, 1986), pg 438)

As previously discussed in section 3.2.1, Edward Dmytryk exhibited a great insight into the properties of saccadic eye movements, saccadic suppression, and change

blindness several years before change blindness even existed as a research field. He understood that certain visual events, such as motion, automatically capture viewers' attention. He suggested how this could be used to lead attention across a cut, creating expectations about the content of the next shot and providing a perceptual 'hole' in which the visual transients of the cut could be hidden. The same insights were also independently developed by Richard Pepperman who expanded on the ways in which editors utilise attentional capture (or the "Quick-Eye Reflex" as he referred to it) to control viewers attention and ease the transition between cuts (Pepperman, 2004). The insights provided by Dmytyrk and Pepperman are not hypothetical techniques that could be used; they are intended as explanations of the techniques used by the best editors. They make reference to evidence of these techniques in existing films and suggest that, even if the editors were unable to express how or why they were doing it, editors were using saccadic suppression to hide cuts. However, without a direct record of viewers' eye movements there is no way of knowing if Dmytryk and Pepperman's assessments are valid.

3.4.2.2 Attentional Synchrony

According to Dmytryk, the main trick of using saccadic eye movements to hide cuts is to get all viewers to move their eyes at the same time (Dmytryk, 1986; pg 438). This seems like a difficult thing to achieve considering that the pattern of eye movements made whilst viewing static images varies between and within viewers depending on such factors as individual differences, task requirements, and familiarity with the image (see Figure 3-6 and Yarbus, 1967). In every image there will be regions that are more likely to be fixated by all viewers. Visual features such as human faces, unusual or out of place objects, objects presented in the centre of the image, areas of high detail or significance to the viewing task will all be fixated more frequently (Yarbus, 1967). However, all viewers are free to visit any region of the image whenever they choose. This lack of control would not provide Dmytryk with the synchrony of attention required to hide a cut.

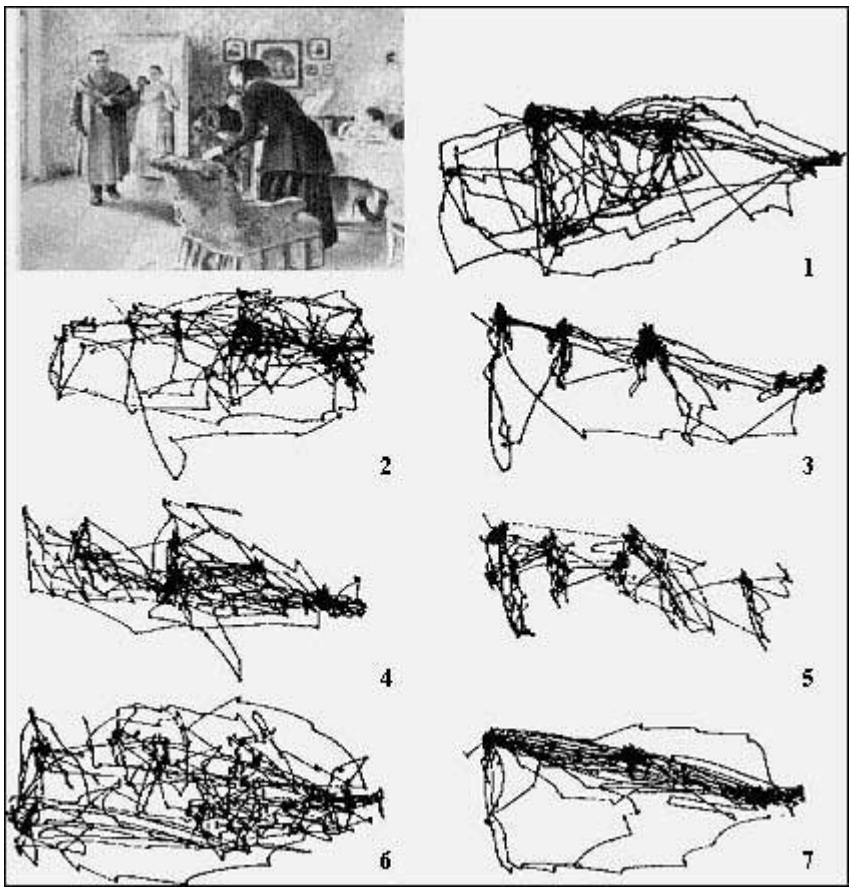


Figure 3-6: Seven records of eye movements by the same subject. Each record lasted 3 minutes. 1) Free examination. Before subsequent recordings, the subject was asked to: 2) estimate the material circumstances of the family; 3) give the ages of the people; 4) surmise what the family had been doing before 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 (from Yarbus, 1967).

Fortunately, moving images introduce the possibility of the image (or dynamic visual scene as it now is) capturing a viewer's attention instead of waiting to be fixated. The use of attention capturing visual features in film has already been established (see section 3.2.1), what is now of interest is whether such attention capture occurs across cuts.

This question has never been explicitly tested in an empirical study but several studies have performed similar enough investigations that by close examination of their results evidence of attention capture and attentional synchrony across cuts can be found. Studies recording the eye movements of viewers whilst they watch feature films have found that there is very little difference between where viewers look at

any point during a film (May et al., 2003; Stelmach, Tam, & Hearty, 1991; Tosi et al., 1997). The degree of agreement between viewers increases as the motion depicted in a scene increases and decreases (i.e. viewers look where they want) when the scene is more static (Tosi et al., 1997). Most fixations occur within a small region in the centre of the screen (a quarter of the width of the screen; Tosi et al., 1997). This central tendency of fixations has been explained as due to the convention of framing most shots so that they contain one significant object located at the centre of the screen (May et al., 2003). This tendency to centrally compose shots has also been associated with a lower frequency of eye movements and the impression of a film lacking in "visual momentum" i.e. being unexciting and "cinematically dead" (Block, 2001; Hochberg & Brooks, 1978a; Treuting, 2004).

Figure 3-7: Screen shots with superimposed gaze position from 19 viewers (pink and yellow spots). The frames number is presented under each image (30fps= 33.3ms per frame). Left-most image is the frame before the cut. Right-most is the frame at which all gaze positions finish moving after the cut.

A small exploratory eye tracking study was performed to generate illustrations of this supposed attentional synchrony[39]. This study presented a short film (2 minutes in

[39] This study is not reported in full as, due to time constraints, the study was limited in terms of the length and variety of films used (only one film 2 minutes in length was used) and the number of different viewing conditions that could be tested. Two different viewing conditions were used: silent

length) of a conversation edited by a professional editor[40] according to the continuity editing rules. Figure 3-7 shows two types of cuts: a cut across dialogue with matching eyelines (top) and cut to close-up with matching action (a head turn; bottom). Superimposed on to the frames are spots representing the gaze positions of 19 viewers[41]. The clear clustering of gaze positions around the actors' eyes can be seen. This is exactly as is expected given previous eye tracking evidence (Gullberg & Holmqvist, 1999; Klin, Jones, Schultz, Volkmar, & Cohen, 2002b; Yarbus, 1967). It appears that the larger the camera-subject distance (see Appendix A) the greater the degree of between-subject variability of gaze location (compare frame 1184 to all others). This is probably due to the increasing number of centres-of-interest as previously indicated by Hochberg & Brooks (Hochberg & Brooks, 1978a). The key effect illustrated by Figure 3-7 is the attentional synchrony occurring across cuts. The first shot of each sequence (frames 1115 and 1184) depict the gaze position immediately prior to the cut. The next frame occurs immediately after the cut (frames 1116 and 1185). Notice how the gaze position remains the same. The final shots depict the gaze position once all subjects have finished performing their first saccade after the cut (1125 and 1195). Notice how all subjects move their eyes to the same position. Any variability of pre-cut gaze position (e.g. frame 1184) has been eliminated by cutting to a new shot with a single centre-of-interest. What cannot be seen from this pattern of eye movements is if the saccades begin in anticipation of the cut as would be required for the cut to be hidden by saccadic suppression. The

---

(pink spots Figure 3-7) and audio (yellow spots). There were no differences between these groups in terms of where they looked and when they initiated saccades relative to a cut. This was highly unexpected and has been attributed to the low number of cuts tested and the infrequency of dialogue "bleeding" across cuts. For anything to be concluded about eye movements with and without sound a larger study would be required.

[40] Chris Learmonth, Cinesite UK

[41] A dedicated piece of software called Gazeatron was developed by the author of this thesis to allow gaze position from multiple viewers to be displayed at the same time. Gazeatron was developed in the Shockwave format (Macromedia, 2004). The screen shots presented in Figure 3-7 were generated by Gazeatron. The gaze positions are generated in real-time as the film plays which enables the experimenter to get a true insight into attentional synchrony. This function is not currently available in commercially available eye tracking software.

saccades, as illustrated here, take 300ms (top) and 267ms (bottom), which suggests that they were initiated in response to the cut. A larger, more controlled study is needed to look for anticipatory saccades as well as investigate the true pattern of attentional synchrony across cuts.

Luckily a study already exists that provides supporting evidence of anticipatory saccades across cuts (May et al., 2003). This study eye tracked viewers whilst they watched a popular feature film (*The Mask of Zorro;* Campbell, 1998) and then divided the data up into nine different types of cut. The length of saccadic eye movements occurring immediately after each cut was then calculated to see if the pattern of eye movements varied across cut types. May et al's results indicated that there was an increase in the length of saccadic eye movements between 120 and 440ms after the majority of cuts (~77%). This indicated that the centres of interest differed across most shots. For some cuts, saccades were initiated in response to the cut (cuts such as those depicting the outcome of an action initiated in the first shot). Therefore, the cut could not have been hidden by saccadic suppression. However, for about 75% of cuts there were signs that some saccades were initiated prior to the cuts[42]. These types of cuts included cuts during conversations, cuts to over-the-shoulder shots, point-of-view shots, and close-ups. In general, all cuts where the object of the new shot was already present in the previous shot or had been cued in someway.

The type of cut showing the clearest signs of saccades coincident with cuts were those occurring during conversations. The continuity editing rules (specifically the 180° Rule) specify that such cuts must maintain the location of each character on the screen across the cut. The benefit of preserving screen location has previously been seen in relation to the *push*ing of attention across the cut by gaze shifts (see section 3.2.2). Now with May et al's eye tracking results clear evidence of this benefit is seen. The social cues of the conversation, such as gaze shifts, head turns, speech and

[42] The peak in saccadic eye movements started 120ms after a cut. Given that most saccades take between 150 and 200ms to initiate and carry out (Goldberg et al., 1991), any saccade occurring less than 200ms after a cut is considered anticipatory.

hand gestures, direct the viewers gaze to a different character. The viewer initiates a saccadic eye movement and the editor cuts to a new shot of the same character at exactly the same time. By the time the viewer has recovered from their saccade 200ms later the character they saccaded to is still located in the same position but the shot has changed without them noticing. This suggest that editors construct their cuts to take advantage of viewers' overt shifts of attention as proposed by Dmytryk (Dmytryk, 1986).

3.4.3 Summary of Suppress Attention

As shown by Change Blindness studies (Currie et al., 2000; O'Regan et al., 2000), (Henderson, 1997; Henderson & Hollingworth, 1999; Irwin, 1991), blinks and saccadic eye movements both provide periods of decreased perceptual sensitivity during which visual transients associated with visual changes will not be perceived. These perceptual 'holes' are created by neural mechanisms that suppress the visual pathways responsible for processing the visual transients. It has been suggested that one reason why continuity editing results in "invisible" cuts is that cuts are made at the same time that a viewer is inclined to blink (Murch, 2001) or saccade (Dmytryk, 1986; Pepperman, 2004). Evidence that blinks coincide with cuts is hard to find. Blinks are believed to be an index of cognitive activity (Hess & Polt, 1964) as suggested by Murch but no evidence of blinks coincident with event boundaries (as would be expected given the explanation of match-action editing previously derived; see section 3.3) has been found (Smith et al., in press; Whitwell, 2005). However, there remains the possibility that synchronisation of a viewer's blinks with that of a film actor might provide a predictable mechanism for hiding a cut in blink suppression. This theory is yet to be empirically investigated.

By comparison, the use of saccadic suppression to hide cuts is compatible with the practicalities of continuity editing (Dmytryk, 1986; Pepperman, 2004), theoretically established (Currie et al., 2000; Henderson, 1997; Henderson & Hollingworth, 1999; Irwin, 1991), and empirically supported (May et al., 2003). The continuity convention of matching action across a cut has been shown to be a reliable way to

capture an audience's attention so that the visual transients of the cut are suppressed during the saccade (May et al., 2003; Stelmach et al., 1991; Tosi et al., 1997; Treuting, 2004). However, whilst there is evidence that saccadic suppression is used to hide some cuts a dedicated empirical study has not been performed that investigates exactly the triggers used by editors to initiate saccadic eye movements, the expectations viewers have about the post-saccadic scene, and the perceptual consequences of such a 'hidden' visual discontinuity.

3.5 Summary of Hiding a Cut

All cuts create visual transients that have the capacity to automatically attract attention and make the viewer aware of the cut. Cuts constructed according to the continuity editing rules have been shown to capture attention less than cuts identified as discontinuities (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990). This indicates that editors must be limiting the potential for attention capture when constructing continuity cuts. There are four techniques that could be used to achieve this: eliminating visual transients, choosing to cut when viewers expect a change, are processing previous visual events, or have their attention suppressed. The plausibility and evidence of each of these techniques was discussed in this chapter.

First, the plausibility of focussing attention on a part of the scene that was not affected by the cut was questioned. A straight cut from one shot to another would always contain visual transients unless a precise graphical match was used. Graphical matching shots is time consuming and complex and, therefore uncommon. It is also unclear if this results in less or more awareness of the cut. Another technique is to gradually composite multiple shots together into one shot, eliminating visual transients. This technique blurs the definition of a cut but is also currently an uncommon technique in feature films. Whether this technique becomes an established part of continuity editing will be seen over the next few years.

The large majority of transitions between shots used in feature films are cuts. Because of this any technique used to hide a cut must find a way to limit the effect of

the visual transients on attention. The first technique that has empirical support is using attentional cues to create a perceptual inquiry in the mind of the viewer which is answered by the cut. This ensures that attention is only captured by the answer to the inquiry rather than the unexpected transients of the cut. This technique is referred to as inattentional blindness (see Simons, 2000). On examination of the continuity editing rules, attentional cues seem to play a part in the 180° Rule, Match-Action, direction matches, and eyeline matches. Cues such as changes in an actor's gaze direct attention across the cut and create a perceptual inquiry that forms a conceptual bridge upon which a mental representation of the depicted action can be constructed. Empirical evidence of the form of this representation does not currently exist.

Another method of limiting attention capture is to ensure that attention is occupied with the processing of a previously attended visual event (known as the **attentional blink;** (Raymond et al., 1992). The visual events applicable to film viewing are temporal units of human behaviour consistently identified by viewers (Newtson, 1973; Zacks et al., 2001). It was shown that attention fluctuated during the course of these events as information was encoded and stored in memory (Baird & Baldwin, 2001; Schwan et al., 2000; Smith et al., in press; Whitwell, 2005). Visual disruptions presented during periods of high visual motion or during the processing of events were not perceived by viewers (**disruption blindness**; (Baldwin et al., 2001; Levin & Varakin, 2004). Based on this evidence it was suggested a cut positioned immediately after the breakpoint between two events would be resistant to attentional capture as the visual transients of the cut are masked by the depicted visual motion and attention is internally directed to the processing of the beginning of the new event. This suggestion matches exactly the ideal match-action edit point identified by film editors (e.g. Reisz & Millar, 1953).

The final section supplemented the previous two techniques by suggesting that as well as directing attention, deictic cues can be used to provide a period of perceptual insensitivity in which a cut can be hidden. First, the possibility of blink suppression was discussed. This has been proposed as a mechanism of continuity editing (Murch,

2001) but there exists no evidence that blinks coincide with cuts or event breakpoints. A better source of suppression was found during saccades. It was found that film viewers synchronised their saccadic eye movements based on the capturing effects of deictic cues such as changes in an actor's gaze (May et al., 2003; Stelmach et al., 1991; Tosi et al., 1997; Treuting, 2004). This suggests that editors exhibit a high degree of control over their viewer's overt attention. There is also tentative evidence that they use this control to initiate saccades and then hide cuts during the period of saccadic suppression (May et al., 2003).

The last three techniques can be seen as compatible components of the system used by editors to create "continuity". An editor presents an actor performing an action. They identify the point at which one action ends and another one begins as the best place to cut. Meanwhile, viewers have been automatically segmenting the action in order to process its constituent events and encode them in memory. When the first action ends the viewer's attention is occupied with encoding the previous event. This creates a potential perceptual hole in which the cut can be placed. However, the editor has chosen to show the beginning of the next action before cutting. This sudden onset of motion automatically attracts the viewer's attention and creates a new perceptual inquiry at the same time: "What is this new event?" The saccadic eye movement made to the new event suppresses the visual transients caused by the cut and when the saccade is over collocation of the matched action across the cut means that the viewer is able to continue answering their perceptual inquiry. Attention and perception have moved seamlessly across the cut.

This explanation of how editors hide a cut by manipulating their viewer's attention and perceptual expectations is currently hypothetical. During the course of this chapter existing evidence was presented that supports this explanation but the experiments cited are rarely dedicated to questions concerning continuity editing. Abstracting from existing evidence always introduces potential errors through misinterpretation. To ensure that the empirical evidence used to construct a cognitive theory of continuity editing is directly related to the experience of film viewing dedicated studies need to be performed. These studies will provide a direct bridge

between the practice and theory of film editing and the cognitive processes involved in film viewing. The first stage on constructing this bridge is to address the main questions of this thesis:

How does continuity editing

1. minimise awareness of a cut,
2. create the perception of "continuity" across a cut, and
3. ensure that "continuity" is not violated as a consequence of the cut?

The current chapter has addressed the first of these questions and touched on the second question. However, the hypothetical use of attentional processes to limit awareness of the cut proposed in this chapter now needs to be empirically tested. The account of an empirical investigation of questions 1 and 2 will be presented in the next chapter.

Chapter 4: Cuing a Cut

This chapter will address the issue of how a cut can be cued by primitive properties of a shot and how this cue will result in a period of perceptual "blindness" in which the cut can be hidden. The perceptual advantage of such cuing will also be discussed. An empirical study will be presented that investigates these issues.

4.1 Experiment 1: Introduction

In the previous chapter, various techniques were described that would allow a cut to occur without the viewer becoming aware of the visual disruption. Two of these techniques, Expecting a visual change (3.2) and Suppressing attention (3.4), proposed that changes in the depicted scene prior to a cut could be used to create perceptual expectations and attract attention. Various types of change were discussed that appear to be used by editors to serve this purpose such as actor's gaze shifts, sudden movements, pointing gestures, audio cues, etc. These changes were referred to as **deictic cues**. They can be subdivided into three categories: *pull* cues such as sudden movements, colour/contrast changes, object appearances, or spatialized sounds; *point* cues such as a pointed finger or gun; and *push* cues, the only known example of which is a gaze shift.

These types of cues all have different effects on attention and appear to be used by film editors in different ways. A *pull* cue involuntarily attracts attention usually resulting in a reflexive saccadic eye movement[43] to the subject of the cut (Langton et al., 2000). For such a cue to be used to hide a cut, the cut would need to occur as the eyes moved to the cue. This would suppress attention during the saccade and ensure that the viewer's attention was not captured by the visual transients caused by the cut (3.4). Evidence that *pull* cues are used in this way was seen in various eye tracking

[43] An involuntary movement of the eyes controlled by a sensory event external to the person.

studies (May et al., 2003). Viewers exhibit reflexive orienting to the *pull* cues. The source of the cues are collocated across the cut so that when attention returns after the saccade the shot has changed but the source of the cue is still in the same location. This convention of collocating[44] an object across a match-action cut (the technical name for such a cut) indicates that editors believe that such collocation will lead to what they refer to as "continuity". However, the object will show clear signs of the change in viewpoint occurring during the cut such as rotations or enlargements. The *pull* cue, unlike a *push* or *point* cue, does not prepare the viewer for such a change. The only perceptual expectation that a *pull* cue can be assumed to create is the expectation that, when attention returns after the saccade the source of the cue will still look the same and be located in the same position[45].

By comparison, a *point* cue can create both a reflexive shift of attention and a conceptual bridge across the cut. Under normal viewing conditions a *point* cue, such as an arrow, does not cause a reflexive shift of attention towards the target of the cue (Jonides, 1981). Attention will only be shifted to the arrow's target once the target has appeared. However, if a point cue is presented in a context where it is expected that the target of the cue will be significant, attention will be covertly[46] shifted to the location of the target before it appears (Yarrow, Whiteley, Rothwell, & Haggard, 2005). This covert shift of attention will speed a subsequent saccade to the target location. This reflexive attention shift can be used to hide a cut by ensuring that it is the target that captures attention after the cut, not the insignificant visual transients of the cut.

The expectation created by the *point* cue also provides the viewer with a conceptual bridge between the shots. The *point* cue directs attention to a new location or object. This creates a perceptual enquiry e.g. "what is it pointing at?". If a cut occurs as the covert attentional shift is seeking out the target of the cue, the new shot will answer the perceptual enquiry. The spatial relationship between the *point* cue and its target

---

[44] Maintaining the location of an object on the screen.

[45] The validity of this assumption will be examined in chapter 5.

[46] Without an eye movement.

may also allow the viewer to create a cohesive mental representation of the two shots as being spatially-adjacent within the same 3D scene (more in chapter 5).

Whilst a *point* cue can lead to the reflexive attentional shift required for a cut to be hidden it does not always do so. The viewer has to adopt the attentional stance that the target of the *point* cue will be significant (Yarrow et al., 2005). In other words, the viewer has to "pay attention" to the *point* cue. Only then will it create both an attentional and conceptual bridge across the cut. By comparison *push* cues, such as a gaze shift, establish these connections automatically (Langton & Bruce, 2000; Langton et al., 2000; Yarrow et al., 2005). When presented with a shot of an actor's face, viewers will mostly look at the actor's eyes (Yarbus, 1967). When their eyes shift, the viewer's attention will also shift reflexively in the same direction. This involuntary shift or attention will occur even if the viewer knows that the change in gaze is not significant (Langton & Bruce, 2000; Langton et al., 2000; Yarrow et al., 2005). This makes an actor's gaze shift a very reliable cue for controlling attention and explains why they are so commonly used to hide cuts (Dmytryk, 1986; Pepperman, 2004; Murch, 2001).

Whilst there exists theoretical motivation for using deictic cues to hide cuts (see 3.2) and some preliminary evidence that editors actually use them in this way (see 3.4), the precise mechanism by which they work is not fully understood. For example, it is not known if factors such as saliency of the cue affect the probability or timing of an attentional shift. It has already been suggested that the effectiveness of a cue, such as a gaze shift, decreases as the camera-subject distance increases i.e. the camera moves further away (3.4). There are also issues of the cue's significance. If a shot depicted three men and they all performed a head turn at the same time, would the main character be more likely to attract attention than the other characters due to his significance within the narrative? How would significance interact with salience? These are all interesting questions that need to be answered; but before they are addressed a more primitive understanding of the attentional cues used in editing needs to be developed. Do viewers have any default expectations that lead to

attentional cuing? How do these expectations distribute attention across a cut? What are the perceptual benefits of such expectations?

To address these questions a single form of attentional cue will be analysed in depth. This cue is commonly used within continuity editing yet simple enough to be empirically dissected and manipulated. If possible, it should be exempt from the combinatorial effects discussed above and universally effective (i.e. have the same effect on all viewers). Such a cue can be found in the primitive constituents of film, moving images within a frame, and emerges from their interaction: occlusion of a moving object by the screen edge.

4.1.1 Occlusion by the screen edge

Film is constructed from a series of cinematographic shots. These shots present a dynamic visual scene within the confines of a frame. The frame represents the limited space available on a strip of celluloid film for recording the light projected from a scene (or digital receptors in a video camera). The frame is both, a container for the recorded visual scene and a border separating the recorded visual scene from the real scene when the film is later presented to an audience. Whilst the frame limits how much of the recorded visual scene can be presented to the audience, it cannot limit the action within this scene as it has no physical presence within the recorded space[47]. Action occurs independently from the frame with objects moving in and out of shot. As an object within the recorded scene moves off screen, it gradually disappear from view behind the screen edge (see Figure 4-1). Visually this is very similar to the real-world perceptual experience of viewing a three dimensional space in which objects, either through their own motion, the motion of other objects, or our viewpoint, continually move in and out of view as they are *occluded* by opaque objects occurring between our eyes and their surfaces. The only difference being that

[47] Of course, action will usually be directed for a feature-film so that significant action does not move out of shot unnecessarily.

when viewing film the occluder is a factor of the presentation medium not an object within the same visual space as the moving object.

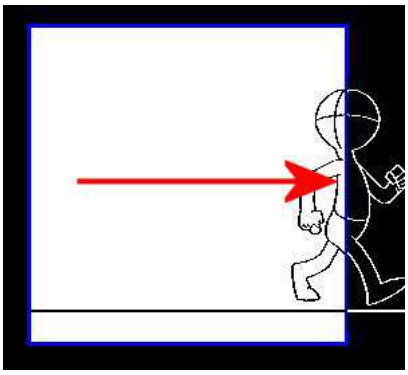


Figure 4-1: Occlusion of a moving character by the screen edge. White area represents the screen and the black is the space outside of the screen not presented to the audience. The portion of the character outlined in white will be occluded.

The screen edge is the only permanent feature of a projected film. The objects, locations, and events depicted within a film will change but the screen edge will remain constant throughout the film's presentation. All of the attentional cues previously discussed (see 3.2) can occur within a film but they are reliant on the conditions under which they are presented. By comparison the screen edge always remains the same and will occlude any object that moves out of shot. If occlusion could be shown to have a cuing effect on attention, occlusion by the screen edge could be interpreted as the most primitive attentional cue available in film.

4.1.2 Occlusion as a primitive editing cue

No evidence exists indicating that attention is attracted, directed, or changed in any way as an object is occluded by the screen edge. In fact, under normal occlusion conditions, i.e. when the occluder is a physical object within the same space as the moving object, it has been shown that attention can be directed to the occluder (Haimson & Behrmann, 2001; Moore, Yantis, & Vaughan, 1998) and continue to track the occluded object without any detrimental effects (Churchland, Chou, & Lisberger, 2003; Pylyshyn & Storm, 1988; Yi, Kim, & Chun, 2003). However, a screen edge is not a normal occluder. It does not share the same physical space as the occluded object and it has no fixed width. The absence of width is important as it

means that there is no opposite edge to that which occluded the object. When a viewer sees an object move behind a normal occluder they will continue to perceive the object (a phenomenon known as *existence constancy)* and formulate predictions about when and where it will reappear based on its motion prior to occlusion (known as *spatiotemporal continuity*; Bower, 1967; Gibson et al., 1969; Leslie, 1984; Michotte, 1955; Spelke, Kestenbaum, Simons, & Wein, 1995; more about this concepts in 4.1.3).

When the moving object is occluded by the screen edge, there is no space on the other side of the screen edge in which the object can reappear. Whether this incompatibility between occlusion by the screen edge and the perception of spatiotemporal continuity stops a viewer perceiving the object as it moves behind the screen edge will be discussed in more detail later (chapter 5). As for attention, there would appear to be no reason for the viewer to continue tracking the object during occlusion by the screen edge as there is no where for attention to go. Instead, the most useful place for attention to be focussed after an object has been occluded by the screen edge is back within the screen.

Attention is known to shift between informative parts of the visual scene (Yarbus, 1967). Once the available information from one part of the visual scene has been exhausted, a shift will be made to a new part. Occlusion of an object by the screen could be seen as constituting a removal of information from that object. Therefore, it would seem natural for attention to shift to another informative region of the screen after the previously fixated object has been occluded by the screen edge. Some indication that attention shifts between objects in a film in this way has been reported by eye tracking studies (Klin et al., 2002b; Stelmach et al., 1991; Treuting, 2004).

If attention is directed back onto the screen, occlusion of a fixated object by the screen edge could be interpreted as a *push* cue: an involuntary redirection of attention *away* from a visual event (Langton et al., 2000). Previously, the only *push* cue was thought to be a change in a viewed person's gaze direction (Yarrow et al., 2005). If occlusion by the screen edge can be shown to reflexively redirect attention in a

similar way then it may represent a primitive technique for hiding a cut. If the attentional shift takes the form of a saccadic eye movement then the cut can be hidden during saccadic suppression. If the attentional shift is covert, then the expectation, whatever form this takes, may focus attention and ensure that only the expected visual object captures attention across the cut. Establishing that such shifts of attention occur after occlusion by the screen edge will be the main goal of this chapter.

4.1.2.1 The history of occlusion by the screen edge

A good place to start looking for evidence that occlusion by the screen edge is a useful technique for hiding a cut is in the history of continuity editing conventions.

4.1.2.1.1 Forbidden Occlusion

The very first films (those made around 1900) consisted of just one shot depicting an entire episode as a distant tableaux[48] (Burch, 1990). If more than one shot was presented in sequence the spatiotemporal relations between shot usually remained unspecified (Bordwell, 1997). This was accentuated by the use of fade-to-blacks after each tableaux (Bottomore, 1990). Primitive films were highly artificial in their staging (resembling theatre in this respect) as all action had to remain within the confines of the frame (Burch, 1990). Characters would rarely be seen entering or leaving the shot and if they did leave, they were considered to have left the event depicted in the film (Nelmes, 2003).

[48] A composition in which the scene is presented in Long Shot so that all action occurs within the confines of the frame.

Parallel to the development of this "primitive" form of film (Burch, 1990) was the more innovative development of 'actuality' filmmaking: the filming of live events[49]. When recording real-life events the filmmaker could not direct the action to accommodate the limited view of the camera. These early actualities resemble more closely what we are used to seeing in film today: people, animals, and vehicles move in and out of shot with no respect for the frame edge and the composition of the shot.

This freedom of movement can be counter the filmmaker's artistic intentions for the scene and so techniques needed to be developed that allowed the filmmaker to accommodate the action with the camera. The simplest of these is the reframing of a shot through panning[50] and tilting[51] the camera. If a character had wandered off screen the camera could be slowly panned to bring them back into shot. However, slight reframing only works if the action occurs within the camera's field of view and does not take the character into a part of the scene hidden from the camera.

As these actuality films became popular, they began to attempt to record complex events such as football matches and Queen Victoria's Diamond Jubilee procession in 1897 (Salt, 1983). These events were often too big for one film to capture so multiple cameras would film the event simultaneously (Bottomore, 1990). This resulted in multiple views of an event that had to be sequenced in some way before they could be presented to an audience. The shots would be sold separately (i.e. they were not spliced together) but cinema exhibitors were advised to purchase multiple shots and show them together in an order that preserved the chronology of the event. These multi-shot actuality films proved very successful as they presented actions bigger than the confines of the frame. Filmmakers began to realise the potential of film to represent spatially and temporally more than was possible in the theatre (the

---

[49] The original street scenes filmed by the Lumiere Brothers and the factory gate films of Mitchell and Kenyon are fine examples of this. See examples of these films at http://www.screenonline.org.uk/film/id/1084507/index.html

[50] Left-Right rotation of the camera about its vertical axis.

[51] Up-Down rotation of the camera about its horizontal axis.

principal visual narrative medium up to that point; Bottomore, 1990). This potential was eventually realised in the *chase* genre.

4.1.2.1.2 Chasing attention across cuts

The success of the actuality films indicated that an event, such as a procession, could be shown across multiple shots and the viewer would still follow the action. This technique of dissecting an event into chronological units and presenting them sequentially was adopted by narrative filmmakers very early[52], but these early multi-shot films were often judged as confusing by audiences (Salt, 1983). The sudden cut from one tableaux to another was thought to be too disruptive to the viewer so the junctures between shots were emphasised using fades and dissolves (Bottomore, 1990).

It was not until action was used to cue the cut that cutting straight from one shot to another became acceptable (Salt, 1983). The first filmmaker to use this technique was James Williamson, a chemist and amateur filmmaking from Hove, England[53]. His films *Stop Thief!* (1901) and *Fire!* (1901) set the standard for all subsequent chase films. *Stop Thief!* (1901) depicted the theft of a joint of meat and the subsequent chase of the thief by a butcher and a pack of dogs (see Figure 4-2 on page 112). The action takes place across three different locations. A static camera is used to film the action in each location and a straight cut to the next shot occurs when the action moves out of shot i.e. is occluded by the screen edge. This continuation of action across multiple shots was what the actuality filmmakers had achieved just by recording natural events but Williamson's application of the technique indicated how it could be used to move narrative cinema beyond the frame.

---

[52] Such as Lumiere's *Vie et Passion de Jesus Christ* (13 shots, 1897) and Melies' *Cendrillon* (20 shots, 1899).

[53] See http://www.screenonline.org.uk/people/id/519902/ for a biography and film clips.

Figure 4-2: James Williamson's *Stop Thief!* (1901). A tramp steals a joint of meat from a butcher (top left) then runs out of shot. The tramp enters a new shot (top right) and is pursued by the butcher (bottom left). The final shot (bottom right) shows the tramp hiding in a barrel before being found by the butcher and his dogs.

As was discussed in chapter 2, matching an action across a cut is one of the main techniques described by the continuity editing rules as creating an "acceptable" cut. Any action, no matter how small, has the potential to hide a cut (Dmytryk, 1986; Pepperman, 2004). Over the first two decades of the Twentieth century, most of the conventions of the continuity style of filmmaking in use today were established [54](Bordwell et al., 1985; Burch, 1990). However, the first convention to be established was that used by Williamson in *Stop Thief!*: the cuing of a cut by a character's exit from the screen. The almost immediate development of this convention at the very beginning of the film form seems to indicate that there is something primitive about it. It cannot be viewed as an established convention as no other medium existing at that time was able to instantaneously transport the viewer to

[54] The obvious exception is the integral use of sound. Synchronised sound was not commonly used in film until 1928.

a new visual scene[55]. Something about an object's occlusion by the screen edge seems to make a cut to a new visual scene acceptable to the viewer. This convention seems to support the view that occlusion by the screen edge may function as an attentional cue. To investigate this possibility further the exact form of the convention will be identified.

4.1.2.2 Matched-Exit/Entrance Cut

The cut cued by an object's occlusion by the screen edge will be referred to as a **matched-exit/entrance**[56]. An object is depicted exiting the screen from one screen edge, a cut is made, and the same object is shown re-entering the screen from the opposite screen edge (see Figure 4-3). The direction of the object's motion must be the same across both shots. This constraint ensures that the matched-exit/entrance cut maintains the directional continuity enforced by the 180° Rule (see 2.1.1).

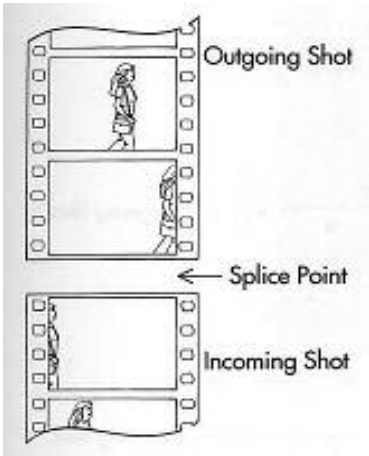


Figure 4-3: An illustration of how to create a matched-exit/entrance cut. Taken from Katz 1991; pg 155.

---

[55] Stage exits via the wings of a stage or doors in the set were an established part of theatre (Nelmes, 2003) but even if these exits indicated the end of a scene the new scene could not be presented instantly as it can in film.

[56] This name was devised by the author of this thesis. This type of cut is referred to in the editing literature in various ways, such as "frame cut" (Bordwell et al., 1985), but there exists no consensual name. Matched-exit/entrance was chosen as it represents the cut's status as a match-action cut whilst specifying that the action is a screen exit followed by a screen entrance.

Description of this type of cut can be found in many film theory and editing textbooks. For example, "a figure leaves the shot, and, as the body crosses the frame line, the cut reveals the figure entering a new shot, with the body still crossing the [opposite] frame line." (Bordwell et al., 1985; pg 51). This definition also suggests the correct timing of the cut: "as the body crosses the frame line". Matching the edit half-way through the screen departure seems to be an established convention (at least in the film literature[57]) and can be seen in more recent editing hand books such as Steven Katz's popular Film Directing: Shot by Shot: "it is common practice to make the cut while the subject is still partially within the frame." (Katz, 1991; pg 155; see Figure 4-3).

After the cut the object should re-enter the screen from the opposite edge. In the frame immediately after the cut the object should be located half on the screen (Bordwell et al., 1985; pg 51, and Katz, 1991; pg 155). It will then be seen to move back onto the screen at the beginning of the new shot. If the object does not appear on the screen immediately after the cut a discontinuity is believed to occur (Katz, 1991). Any expectations the viewer had about the new shot are erased and a new scene can begin. Leaving the screen empty either at the beginning of the new shot or at the end of the previous shot is a technique known as *Clearing the frame* (Figure 4-4; Katz, 1991).

[57] One of the difficulties of identifying the conventions actually employed by film editors is the prominence of Reisz and Millar's *The Technique of Film Editing* (Reisz & Millar, 1953) and Bordwell and Thompson's *Film Art: An Introduction* (Bordwell & Thompson, 2001). There is a tendency for modern editing handbooks to reproduce the editing conventions as they appear in these two books without acknowledging the original source. The general acceptance of the conventions cited in these two books has been taken as indicating that they are actually what are used by editors without this assumption being explicitly tested.

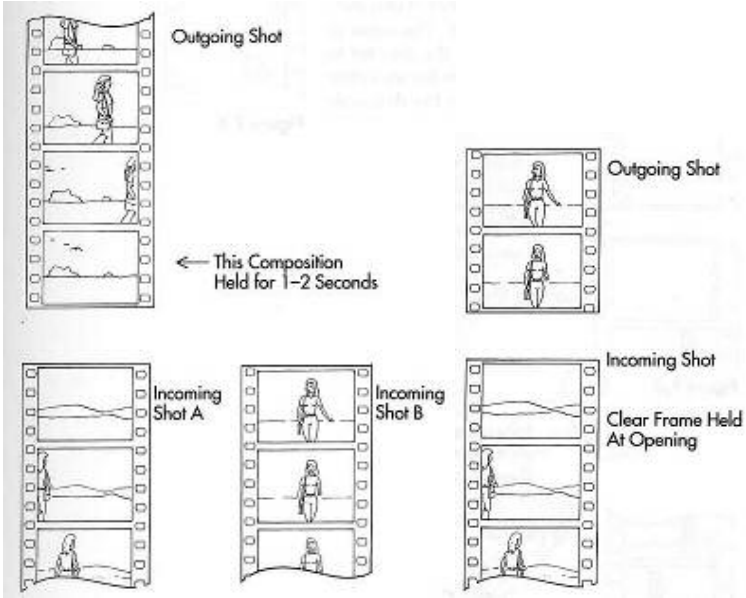


Figure 4-4: Two forms of "Clearing the frame" (taken from Katz, 1991). If the empty frame is held after the principal object has left the screen it is believed to enable the editor to make a usually discontinuous cut to a new shot in which action is not matched exactly ("Incoming shot B"). Holding an empty frame at the beginning of a new shot can either slow the pace of the film ("Incoming shot A") or weaken the impact of an unexpected cut (right film strip).

Clearing the frame is believed to be most disruptive to the viewer when the shot continues after the principal character has left the screen (Katz, 1991). As the shot is old the viewer has already had the opportunity to extract all useful information from it. By comparison, starting a new shot without a visible principal character may still be disruptive to the perceived continuity of action but the new scene will give the viewer something to look at (Katz, 1991). This explanation of the significance of a shot based on the number of places left to look at has previously been described as *visual momentum*: the rate at which attention shifts across a visual scene (Hochberg & Brooks, 1978a).

Attention is a recurring theme in descriptions of the matched-exit/entrance cut. Edward Dmytryk identified a "rule of thumb" used by film editors to find the right point to cut during a screen exit: "Good editing practice rules that the cut away from the first scene should occur at the point where the actor's eyes exit the frame" (Dmytryk, 1986; pg 437). Dmytryk believes the actor's eyes are important because this is the most probable point at which the viewer's attention is focussed (Dmytryk, 1986). The significance of a person's eyes is a long established fact in social

attention research (Yarbus, 1967). Dmytryk's recommendation that the cut should be timed to match the point at which the viewer's gaze crosses the screen edge suggests that at this point attention undergoes a shift that will enable the cut to be hidden. Belief in such an attentional shift seems to be shared by Katz who attributes the effect of *Clearing the Frame* to the detaching of viewer's attention from the principal object and erasure of any expectations the viewer had about where the object would appear in the next shot (see Figure 4-4; Katz, 1991). This shift of attention has been explicitly stated by Dmytryk:

> "At the point of the cut, two things happen: (1) the exiting actor's eye or face – usually the viewer's centre of interest - leave the screen and, as a result (2) the *viewer's* eyes, which have been following the actor's movement, encounter the darkness at the edge of the screen. These two actions cause a reaction: **The viewer's eyes swing back toward the centre of the screen, then continue to its right edge where they find the continuation of the actor's movement.** All this happens in a fraction of a second, not nearly long enough for the viewer to be aware of the passage of time, to be conscious of his eye movement, or to *notice the cut* has slipped by in the interim." (Dmytryk, 1986; page 438)

Therefore, the convention of using occlusion of an object by the screen edge to cue a matched-exit/entrance cut is well established. The conventions, almost immediate development during the early days of film and status as the founding technique of the continuity editing style, seems to suggest that it is in someway primitive. The mechanism by which it is assumed to function is an attentional shift accompanied by expectations of the content of the new shot (Dmytryk, 1986; Katz, 1991). Both of these have been shown to have the potential to hide a cut (see chapter 3). However, there is no empirical evidence that such an attentional shift exists or that the viewer develops expectations across a cut. The empirical study to be performed in this chapter will examine the distribution of attention across a matched-exit/entrance cut. Before this study takes place the conventions of 50% exit and 50% entrance described above as the best way to create a matched-exit/entrance cut should be examined. Does any evidence exist supporting the view that 50% occlusion before a

cut will lead to the construction of perceptual expectations compatible with the perception of "continuity" across the cut?

To begin investigating how a matched-exit/entrance cut might create the perception of "continuity" the perceptual consequences of normal occlusion should be first examined.

4.1.3 Visual Occlusion

Visual occlusion is a very powerful perceptual cue that leads to continued perception of the object during occlusion (Bower, 1967; Gibson et al., 1969; Leslie, 1984; Michotte, 1955; Spelke et al., 1995). The continued perception of an object in the absence of a visual referent is known as *existence constancy* (Michotte, 1955). The phenomenon was first investigated by Albert Michotte as the *tunnel effect*: the impression that a train continues to exist when it moves into a tunnel (Michotte, Thines, & Crabbe, 1991). Michotte interpreted this phenomenon as evidence that our perception of objects within the visual world is made up of both visual stimulation and perceptual reconstructions of absent visual information (Michotte et al., 1991). When the occluded object is static the perceptual "filling in" of the object is known as *amodal completion.* If the "filling in" occurs over time as an object moves behind an occluder it is referred to *amodal integration* (Michotte et al., 1991). The result of these processes is that the perceptual "whole is greater than the sum of its parts", an idea most commonly associated with the Gestalt Psychologists.

4.1.3.1 Gestalt Psychology and Amodal Completion

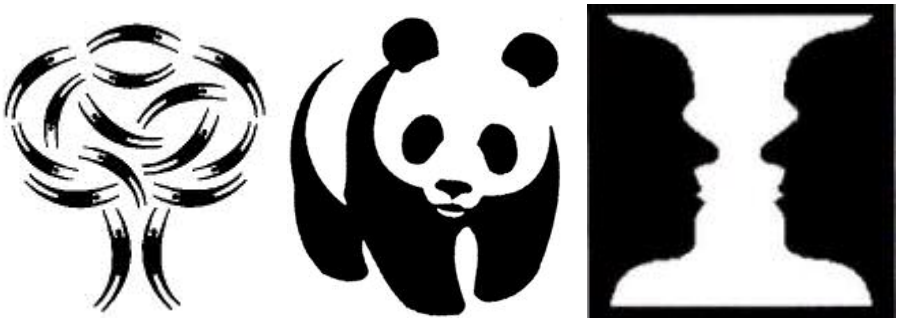


Figure 4-5: Proximity and Good Continuation create a tree from people (left), closure creates the WWF panda (centre), and switching between figure and ground interpretations of the vase-faces image creates two different perceptions (right).

The Gestalt psychologists used visual illusions such as the classic vase-faces example (see Figure 4-5, right) to indicate how visual perception is not based purely on sensory stimulation. They developed a number of principles of perceptual organisation which described how the human perceptual system prefers certain structures over others (see Koffka, 1935 and Figure 4-6). The principle that unified all other gestalt principles was known as the **Law of Prägnanz**: "Of several geometrically possible organisations the one that will actually occur will possess the best, simplest, and most stable shape" (Koffka, 1935), p138). Identifying the qualities that define the "best" shape has proven very difficult but the key properties seem to be **simplicity**, **regularity**, and **continuity**. When an object is partially occluded the rest of the object will be perceptually "filled-in" by extrapolating the simplest possible shape from the visible sections of the object. For example, in Figure 4-6 the partially occluded shape A will be perceived as a solid circle (B) rather than one of the other, equally probable shapes (C or D).

The Gestalt principle of '**good continuation'** (otherwise known as **'continuity'**) states that perceptual organisation will tend to preserve smooth continuity rather than abrupt changes (e.g. a tree is perceived from a collection of people in Figure 4-65; (Koffka, 1935). If an array of visual elements or visible parts of a semi-occluded object can be grouped together along a smooth continuous line then this is preferred over any grouping that deviates from this line (Kellman & Shipley, 1991).

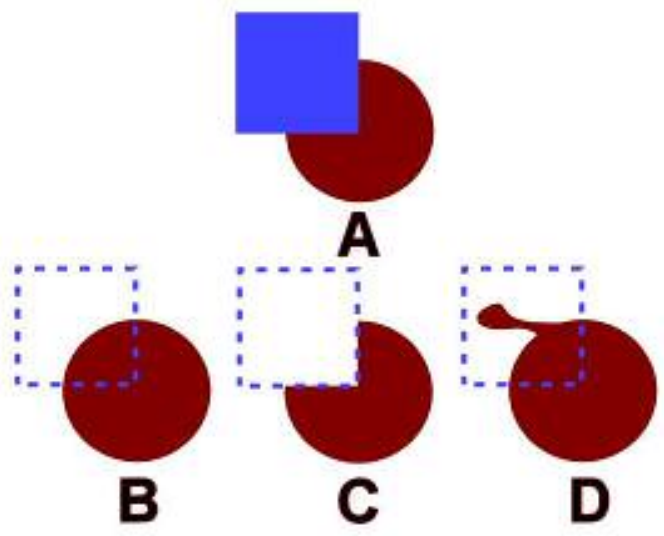


Figure 4-6: Example of amodal completion of a circle (A) occluded by a square. The circle is usually perceived as being completed behind the square (as in B). All other versions are equally probable but less likely to be perceived.

This principle of good continuation is also essential to *amodal integration*: the perception of a complete object as it moves behind an occluder (Michotte et al., 1991). As well as the perceptual completion of the object during its initial occlusion, *amodal integration* also requires the object to be perceived as moving behind the occluder. This requires the object's constructed perceptual representation to be projected through space and time based on the object's motion prior to occlusion. According to the principle of good continuation, the object would be perceived as following the simplest, most regular, and continuous path without any unpredictable discontinuities e.g. changes in direction, speed, or shape changes.

Evidence that viewers continue to perceive an object during occlusion (i.e. existence constancy) and expect the object to follow a continuous path (i.e. **spatiotemporal continuity**) has been shown by developmental studies (Spelke et al., 1995; Xu, 1999; Xu & Carey, 1996). When presented with an object that moves behind an occluder, infants will expect the object to trace a spatially continuous path (known as the **continuity constraint;** Hirsch, 1982) and continue moving at the same speed as prior to occlusion (the **smoothness constraint** Spelke et al., 1995). If either of these constraints are violated, such as when the object does not appear from the other side

of the occluder, the infant ceases to perceive the object as existing (Spelke et al., 1995).

Evidence that an occluded object continues to be perceived as it moves behind the occluder has been provided by neuroimaging (Baker, Keysers, Jellema, & Wicker, 2001; Olson, Gatenby, Leung, Skudlarski, & Gore, 2004) and behavioural studies (Scholl & Pylyshyn, 1999; Yi et al., 2003)[58]. Both types of study have shown that the perception of existence constancy is conditional on the type of visual transformation the object undergoes as it disappears from view. If the object suddenly disappears from view without an intermediate period of occlusion the object will not be perceived as continuing to exist. The exact visual transformations resulting in the perception of existence constancy have been investigated by Michotte (Michotte, 1955), Gibson and Kaplan (Gibson et al., 1969; Kaplan, 1969).

4.1.3.2 Creating existence constancy

Michotte and his students were the first to identify the specific stimulus conditions necessary for the continued perception of an object in the absence of sensory stimulation. These transformations are classified according to three qualities (cited in Bower, 1967):

1. **gradual vs. abrupt**, gradual changes have detectable intermediary stages, abrupt do not;
2. **wholefield vs. local**, the change effects the whole scene or an isolated object;
3. **perspectival vs. non-perspectival**, the change occurs due to occlusion by an edge within the 3D space or by some other means.

Michotte and his students identified two main types of transformations that led to existence constancy:

[58] The exact form of this perception and its implications for continuity perception across cuts will be discussed later (chapter 5).

1. **global, gradual, non-perspectival transformations.** These occur when the entire visual scene fades from view due to a change in light levels e.g. when dusk falls or indoor lights are dimmed[59];
2. **local, gradual, perspectival transformations**. These occur when an object moves behind another object in the visual scene (Michotte, 1955).

Michotte identified all other classes of transformations as being either physically impossible or leading to the object ceasing to exist. For example, a local, abrupt, non-perspectival transformation occurs when an object explodes or disintegrates. However, Gibson and Kaplan identified further conditions under which existence constancy occurred (Gibson et al., 1969; Kaplan, 1969):

3. **local, gradual, non-pespectival transformation**. An object will be perceived as continuing to exist when it moves to the horizon, decreasing in size and resolution until it is finally indiscernible at the vanishing point (Gibson et al., 1969).
4. **wholefield, gradual, pespectival transformation.** When our view of the entire visual scene is gradually obstructed either by a large dark occluder in the visual scene (such as a curtain being pulled across a window or stage) or closure of our eyelids (Gibson, 1979).

According to Gibson, the quality that connects all of these transformations and leads to the perception of existence constancy is their reversibility (Gibson et al., 1969). Any of the above transformations can be reversed and the part of the visual scene previously out of sight will come back into sight. If the reverse of an object's disappearance is not a valid way for it to come back into view then an object undergoing that transformation will not be perceived as continuing to exist. For example, if an object is seen to disintegrate or suddenly disappear from view, reversing the event will not create acceptable conditions under which an already existing object can come back into view or a new object can be created.

[59] This is analogous to a fade-to-black and could explain why early filmmakers such as Melies saw the fade-to-black as being easier on their audience.

4.1.3.3 Summary of Visual Occlusion

An object must disappear from view by gradually being occluded by an edge[60]. The object will then be perceived as continuing to exist during occlusion. The object will be expected to continue moving along the same path, at the same speed as prior to occlusion. These spatiotemporal expectations allow a viewer to predict when the object will reappear from behind the occluder. When the object reappears it must do so through the opposite visual transformation to the original occlusion. If the object reappears through the correct transformation and at the expected location and time it will be perceived as the same object that previously moved out of view.

These constraints are all very specific to the perception of occlusion within the real-world. The compatibility of these constraints with the conventions of matched-exit/entrance cuts will now be discussed.

4.1.4 Compatibility between matched-exit/entrance cuts and real-world occlusion

Applying the constraints discussed above to matched-exit/entrance cuts there appears to be an incompatibility between the perception of existence constancy and the film's inability to satisfy expectations of spatiotemporal continuity. The gradual occlusion of an object as it moves behind the screen edge satisfies the prerequisites for existence constancy identified above (transformation number 2). This should automatically result in the continued perception of the object during occlusion (Gibson et al., 1969; Kaplan, 1969; Michotte, 1955). However, it is not possible for the object to move behind the occluder as it would behind a real occluder. The occluder is not a part of the 3D space through which the object moves; it is a boundary between the projected scene and reality. A viewer cannot expect the object

[60] There are four other acceptable transformations but occlusion is the most common.

to reappear on the other side of the screen edge as the world of the object does not extend past the occluding edge. If the viewer did expect spatiotemporal continuity their expectation would be violated immediately as the object failed to emerge from the other side of the screen edge. This would destroy their perception of existence constancy (Hirsch, 1982).

The only way that the perception of existence constancy could be sustained across a matched-exit/entrance cut would be for the viewer's spatiotemporal expectations of the object's motion to be shifted to the opposite screen edge. If the opposite screen edge was expected to be the edge from which the object would appear, the object could satisfy spatiotemporal expectations, resulting in existence constancy. Shifting the path of the perceived motion in this way would seem to violate the continuity constraint (Hirsch, 1982). However, editors' confidence in the ability of matched-exit/entrance cuts to create perceived continuity suggests that such a shift might occur.

As well as spatiotemporal expectations being shifted to the opposite screen edge the object would also have to move back onto the screen from behind the screen edge. If the object's reappearance did not constitute the opposite visual transformation to its original disappearance, the object would be perceived as different to the object that was occluded (Gibson et al., 1969). This constraint seems to support the matched-exit/entrance convention of having the object move onto the shot after the cut (see section 4.1.2.2). Why the convention would be for the object to move back onto the screen from a 50% occluded position rather than the full 100% is not known.

Given that an object is expected to continue to exist and only disappear from view via one of the four transformations previously listed (Gibson et al., 1969; Kaplan, 1969; Michotte, 1955), if a film viewer expected a depicted object to continue to exist, it too would only be expected to disappear via one of these transformations. Therefore, it could be hypothesised that the screen edge provides an opportunity for the viewer to sustain their perception of the object whilst detaching their attention in preparation for the new shot. If the cut occurred without the object first being

occluded by the screen edge, the viewer would experience none of the visual cues required to perceive existence constancy across the cut. As the cut is unexpected it would form a perceptual "juncture" between the end of one object and the beginning of another.

As suggested by Dymtryk (see section 4.1.2.2), the expectation that an object will only disappear from view when it is fully occluded by the screen edge should be manifest as a withdrawal of attention coinciding with the object's occlusion. As the object moves towards the screen edge, the viewer prepares for its shift to the opposite screen edge by preparing a saccadic eye movement. If the cut occurs when expected the viewer's attention should be at a minimum due to saccadic suppression, making them blind to the visual disruption of the cut. This withdrawal of attention prior to the cut can be used as a sign that the viewer expected the cut.

This withdrawal of attention will be used in the current empirical study to detect if viewer's have particular preference for a cut on occlusion. If supporting evidence for this preference can be found it will be seen as indicating that occlusion by the screen edge is a primitive attentional *push* cue that has the potential to facilitate the perception of existence constancy across a cut.

4.2 Experiment 1: Design

The main question of this empirical study is:

Does a viewer expect a cut to occur as the focal-object[61] is fully occluded by the screen edge?

4.2.1 Hypotheses

This will be investigated by testing four hypotheses.

4.2.1.1 Occlusion expectation hypothesis

> Viewers expect a cut to occur when the focal-object is fully occluded by the screen edge.

To test this, subjects will be shown animations in which a cut occurs when a moving object is in a variety of positions relative to the outgoing screen edge. If they expect the cut to occur when the object is fully occluded they will begin withdrawing attention prior to this position so that the required saccade occurs at the same time as the cut. Attention will be seen to decrease during the 100ms prior to the cut and return to maximum around 100ms after the object has relocated to the opposite screen edge (Rose et al., 2002). If the cut is not expected, no attention withdrawal will be observed prior to the cut and recovery of full attention will not occur within 200ms of the cut. This will indicate that the saccade has been performed in response to, not anticipation of, the cut.

4.2.1.2 Occlusion advantage hypothesis

> Viewers will find it harder to adapt to cuts that occur before the focal-object is fully occluded.

---

[61] Focal-object = the object to which attention is focussed. Usually equivalent to the object being fixated.

The preference for full occlusion before the cut will remain even if the subject is trained to expect the cut to occur earlier. Subjects should find it easier to disengage their attention from the moving object when it is fully occluded compared to when it is still in view. The full-occlusion condition should also show a better increase in performance over repeated presentations.

4.2.1.3 Pursuit initiation hypothesis

> Viewers expect the focal-object to move back onto the screen from a fully-occluded position after the cut.

As well as expecting the moving object to be fully occluded before the cut, subjects should also show evidence that they expect the object to move back into view from an occluded position. According to Gibson, the only acceptable way in which an occluded object can come back into view is by moving back into view from an occluded position (Gibson et al., 1969). If the object suddenly reappeared it would not satisfy the requirements for existence constancy (Michotte, 1991; Spelke et al., 1995; Xu & Carey, 1996).

4.2.1.4 Clearance hypothesis

> Viewers expect the focal-object to begin re-entering the screen immediately after the cut. A time gap between the cut and the object's re-appearance will have adverse effects on attention.

If the frame is left empty at the end or beginning of a shot it is said to *Clear the frame* (Katz, 1991). The absence of the focal-object in the new shot makes it difficult for viewers to shift attention across the cut. This will be manifest as less attention directed to the focal-object when it eventually re-appears.

These hypotheses will be investigated using a reaction time experiment. Abstract animations will be used as an analogue for the matched-exit/entrance cut as these will allow the precise real-time control of the editing considerations in a way that would not be possible with film.

4.2.2 Methodology

4.2.2.1 Subjects

28 Edinburgh University students (16 male and 12 female) were recruited as subjects for this study. Their average age was 23.64 years (minimum 18, maximum 35). The subjects were screened for attentional disorders such as attentional defect hyperactivity disorder (ADHD). All subjects had normal or corrected to normal vision. Subjects were recruited from throughout the university and paid £5 for their time. Subjects signed a consent form prior to the experiment acknowledging their understanding of goals of the experiment, their voluntary participation, and their option to withdraw at any point. The experiment was designed according to the British Psychological Society's ethics guidelines.

4.2.2.2 Stimulus

The experimental stimuli consisted of a 2D animation analogous to the visual experience of a matched-exit/entrance cut. To make it easier for the viewer to focus their attention to the focal-object of the shot, the object was made to look like a "flying saucer" flying over a green grassy terrain. The "flying saucer" was viewed from above (a "Bird's Eye" view) and all motion was within the screen plane i.e. no change in depth relative to the viewer. The "flying saucer" flew silently over the terrain from the left of the screen to the right. Once the focal-object ("flying saucer") reached the right screen edge a cut was made to a new viewpoint which reframed the focal-object so that it was now on the left screen edge. Each trial of the experiment was made up of many cuts. To achieve the reframing of the focal-object a real camera would have had to instantaneously translate laterally along the path of motion of the focal-object (see Figure 4-7). This lateral movement is analogous to the camera movement implied by a matched-exit/entrance cut. As such the experience of viewing these simplified animations should not differ significantly from a similar cut joining shots depicting real 3D scenes.

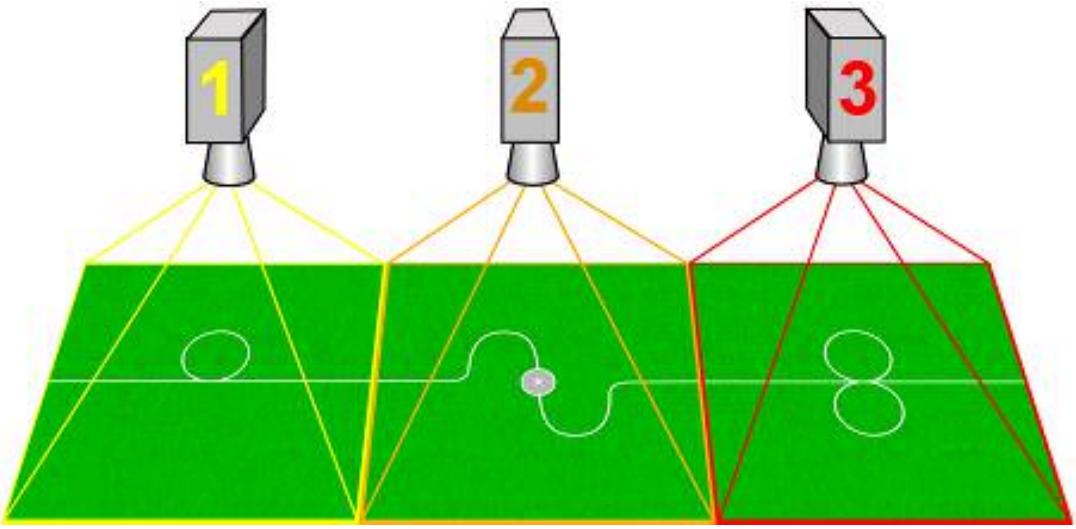


Figure 4-7: An illustration of how the motion (the white line) of the focal-object would be filmed in reality. Three successive cameras would film the continuous motion and then the resulting shots would be edited together. The resulting film would depict the focal-object always flying left to right on the screen but relocating from the right to left screen edge after a cut.

A variety of considerations were made to ensure the viewers tracked the focal-object during its entire screen time. First, the path taken by the focal-object on its journey from the left to the right screen edge was randomly selected from a variety of paths. It was decided that the focal-object would always exit and enter the screen along the centre line of the screen (see white line in Figure 4-7). This ensured that the viewers always knew where to perform a saccade to after the focal-object left the screen. This also ensured that the length of the saccade (and therefore duration) was as constant as possible.

However, it was also decided that the subjects should not be allowed to fall into a rhythmical viewing pattern as this would speed up their saccade initiation times and possibly slow reaction times as they withdraw attention from the screen over multiple repetitions (i.e. grew bored or fell asleep). To avoid this, a variety of smoothly curved routes traversing the centre of the screen were used to join the exit and entry points. This changed the duration of each shot.

4.2.2.3 Presentation Specifications and Apparatus

The focal-object had an on-screen diameter of 60 pixels and was presented against a square background 600 x 600 pixels. The entire movie occupied 840 x 840 pixels with a 120 pixel wide grey border marking the edge of the animation. This border allowed the focal-object to move completely off/on screen (i.e. to/from behind the grey border) before and after a cut occurred. It was intended that, whilst not the actual screen edge (i.e. the monitor casing), the viewer would regard the inside edge of this grey border as the effective screen edge as the depicted action would never move beyond it.

The focal-object moved at a constant speed of 10 pixels per frame (6.24° visual angle per second). The animation was presented on a 21" CRT monitor with a resolution of 1280x1024 and a refresh rate of 85Hz. This was connected to an Intel Xeon 2.4GHz PC with 1GB RAM and a Nvidia Quadro4 900 XGL display adapter. The PC operating system was Microsoft Windows XP.

Subjects viewed the animation from a distance of 61.5 cm. This was controlled by seating subjects at the required viewing distance with their heads rested against a solid chair back. The height of the seat was adjusted to make their eyeline perpendicular to the screen centre. At this viewing distance the screen subtended a visual angle of 15.46° and the focal-object a visual angle of 1.55°. The average saccade distance occurring as the focal-object relocated from the right screen edge to the left was the same as the screen width, 15.46°. The focal-object moved 0.26° of a visual angle every frame and took at least 2250ms to traverse the screen. The animation was presented at 24 frames per second (fps) giving each frame a duration of 41.67ms.

The experiment was implemented using Macromedia Director, an interactive animation development platform. The graphical programming functions of Director allowed the editing conditions (see 4.2.2.4) to be formally represented and manipulated during runtime. This level of graphical control would not have been

possible using an existing experimental package such as E-Prime[62]. Director was used to create a "stand-alone" experimental program that generated all stimuli, performed reaction time recordings, pre-processed the data, and output the data ready for statistical analysis. The accuracy in presentation and recording times was controlled using built-in fail-safes.

4.2.2.4 Independent Variables

The editing conditions were controlled by three independent variables each with three levels: **Exit Percent**, **Entry Percent**, and **Time Gap**.

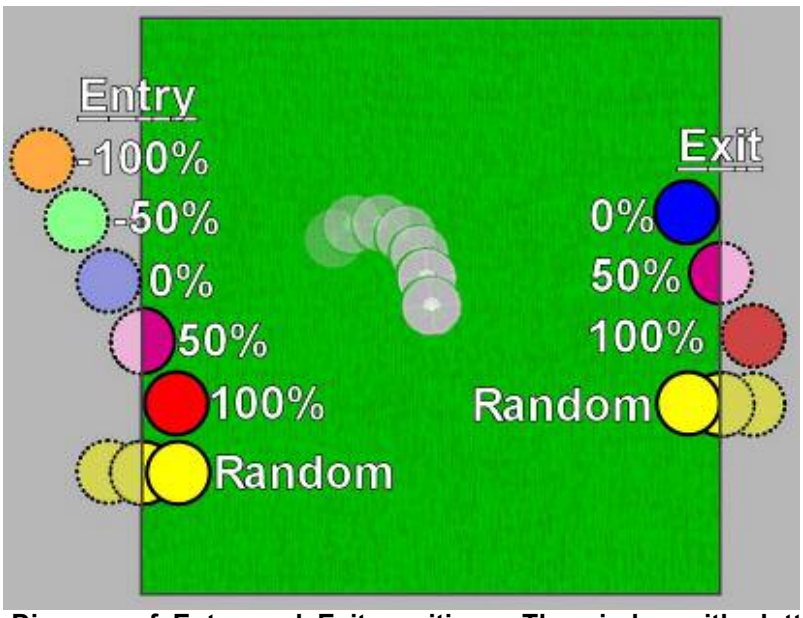


Figure 4-8: Diagram of Entry and Exit positions. The circles with dotted outlines indicate positions where the focal-object is partially or fully occluded by the screen edge.

4.2.2.4.1 Exit

There were four Exit groups: 0%, 50%, 100%, and Random. The percentage value referred to the amount of the focal-object off-screen in the frame immediately prior to the edit. An Exit value of 0% meant that all of the focal-object was on-screen but that it was touching the right screen edge. 50% meant only half of the focal-object could be seen whilst half was hidden behind the screen edge. 100% meant the focal-

[62] Psychology Software Tools Inc. (http://www.pstnet.com/products/E-Prime/default.htm)

object had just left the screen edge prior to the edit (see Figure 4-8). In the Random condition the Exit percentage for each edit was randomly chosen from these three options at the start of each shot. This made it unpredictable. It was decided that the degree of Exit would be represented as the percentage of the focal-object off-screen as this is how the matched-exit/entrance cut has been specified by editors (see 4.1.2.2).

Exit Percent was designed to test the *Occlusion Expectation* and *Advantage* hypotheses.

4.2.2.4.2 Entry

Entry percent had six levels, -100%, -50%, 0%, 50%, 100%, and Random, referring to the amount of the focal-object on-screen in the frame immediately after the cut (see left circles in Figure 4-8). An Entry value of 0% meant that the focal-object was completely off-screen immediately after the cut but touching the left screen edge. An Entry value of 50% or 100% meant that it was half or fully on, respectively. An Entry of -50% or -100% meant that the focal-object was fully occluded and positioned with a gap equivalent to half (-50%) or all (100%) of its width between its right edge and the screen edge. The Random condition meant that the degree of entry was unpredictable and changed between shots.

Entry Percent was designed to investigate two different hypotheses. The positive Entry and Random conditions were chosen to investigate the *pursuit initiation hypothesis* (i.e. full occlusion after the edit is optimum; see 4.2.1.2). The negative Entry conditions and 0% Entry were used in combination with Time Gap to investigate the *clearance hypothesis* (see 4.2.1.4).

4.2.2.4.3 Time Gap

The **Time Gap** variable had four levels: 0 frames, 3 frames, 6 frames, or Random. It refers to the amount of time between the cut and the point at which the focal-object re-enters the screen (i.e. reaches 0% Entry). Time is measured in frames as this is the

smallest unit of time in an animation. The animation was presented at 24 frames per second so each frame was equal to 41.67ms.

To create these time gaps the focal object was positioned at negative Entry immediately after the cut. The object then moved towards the screen edge (out of view), taking the allotted Time Gap to reach 0% Entry and begin re-entering the screen. The animation was designed so that the focal-object moved a distance equivalent to a sixth of its own width every frame. Therefore, to create a Time Gap of 3 frames the focal-object was positioned at -50% Entry immediately after the cut (see Figure 4-8). The focal-object then required 3 frames to move to 0% Entry. For a Time Gap of 6 frames the focal-object must initially be positioned at -100% Entry[63].

4.2.2.5 Experimental Conditions

The Entry and Time Gap variables were varied within subjects and the Exit variable between subjects. A complete crossing of these three variables within subjects was not possible due to the large number of experimental conditions that would have been required. By setting Exit as a between-subjects variable it could be used to investigate whether subjects showed better adaptation to 100% Exit compared to the other Exit values (*occlusion advantage hypothesis*).

At the start of each experiment, subjects were allocated to one of the four Exit groups (0%, 50%, 100%, and Random). For each subject, the Exit value would then always be the same for every cut during the entire experiment. The only exception was the Random Exit group for whom the Exit would always be Random i.e. unpredictable.

---

[63] The decision to have the focal-object move during the Time Gap rather than just pause the animation was made because 1) it was thought to be more authentic, when *Clearing the frame* the film is not paused, and 2) for technical reasons. If the animation had stopped there was an increased chance that the display duration of each frame would change from that achieved whilst moving the object.

The remaining variables, Entry and Time Gap, were mixed to create nine experimental conditions (see Table 4-1 page 133). Four of these conditions (2, 4, 6, and 8 in Table 4-1) varied Entry whilst controlling Time Gap as 0 frames. Conditions 2, 3, 5, and 7 varied Time Gap whilst Entry changed to accommodate (see 4.2.2.4.3). The focal-object in these Time Gap conditions was always at 0% Entry once the Time Gap reached 0 frames.

Condition Number	Entry (% of object)	Time Gap (frames)
1	Random	Random
2	0	0
3	0	Random
4	50	0
5	-100	6
6	100	0
7	-50	3
8	Random	0
9	Random	Random

Table 4-1: Experimental Conditions

Two extra conditions were included: 1 and 9. Both Entry and Time Gap were randomized in these conditions. These conditions were used to test the *occlusion advantage hypothesis*. By comparing performance in condition 9 to condition 1 any adaptation to the Exit condition, which is consistent for each subject, should have been detected. Entry and Time Gap were randomized to ensure that any increase in performance was due to Exit not either of the other factors.

The full list of experimental conditions can be seen in Table 4-1. Conditions 1 and 9 were always presented first and last, respectively. The presentation order of the other seven conditions was rotated within Exit subject groups to protect against learning, task improvement, and fatigue. Seven different permutations were needed to ensure that each condition was presented at each position during the experiment. Balancing

this across the four subject groups (for the four Exit levels) required a minimum of 28 subjects.

4.2.2.6 Subject Task

The subject's task during this study was to follow the focal-object throughout the animation, over a series of cuts, and react as quickly and accurately as possible to a reaction time (RT) cue. The RT cue was a black digit from 2 to 9 superimposed on top of the focal-object. It was presented for 1 frame (41.67ms at 24fps); long enough for the digit to be identified if it is being fixated but not if it is only seen peripherally. The presentation time of 41.67ms is also too short for a saccade to be performed to fixate it. The subject's task was to identify if the digit was odd or even and respond by pressing the appropriate button: 'E' for even and 'O' for odd (on a QWERTY keyboard). The hands used to press each key were counter-balanced across subjects to protect against handedness bias. Subjects were told that accuracy and speed were important so they should respond as soon as they think they can correctly identify the digit. If they were unsure they were asked not to guess. This was done to limit the effect on guessing on the correct response rates.

The **correct response rate** and **average reaction time** for each cue position were recorded as dependent variables.

4.2.2.7 Reaction Time Cue Positions

The majority of the RT cues were placed immediately before and after the cut as these were the areas of interest. Any withdrawal of attention occurring in preparation for a saccade prior to the edit would be reflected in a decrease in response accuracy to the RT task. Similarly, predictive saccades would enable the subject to respond to RT cues immediately after the cut (within the first 200ms; 5 frames) with higher accuracy than if the saccade had only been initiated after the cut occurred.

Thirteen RT cue positions were used: 5 on the left screen edge, 3 in the middle of the screen, and 5 on the right screen edge: 1, 3, 4, 5, 7, 310, 420, 530, -6, -4, -3, -2, 0

(see Figure 4-9). The centre cue positions, 310, 420, and 530, were always positioned in the same horizontal position on the screen but they move vertically to follow the path of the focal-object (which traces a variety of curving paths across the screen centre)[64]. The negative cue positions refer to the number of frames it will take the focal-object to reach the Exit position. When Time Gap is 0 frames, the positive cue positions begin (cue position 1) with the frame immediately after the Cut. When Time Gap is 3, 6 or Random, cue position 1 always occurs when the focal-object has reached 0% Entry. Therefore, in the positive Time Gap conditions a RT cue is not displayed until after the Time Gap.

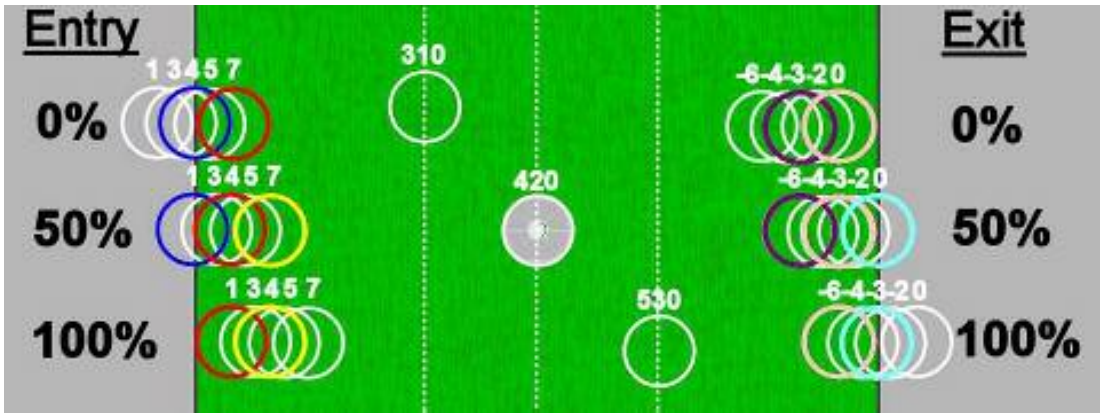


Figure 4-9: The cue positions (circles) associated with 0% (top left and right circles), 50% (middle), and 100% Entry (bottom). All positive and negative positions are always presented on the horizontal middle line of the screen. They are shifted vertically in this diagram for ease of viewing. Cue position 0 always matches the Exit value and position 1 matches Entry value when Time Gap is 0. Positions 310, 420, and 530 always stay in the same horizontal position but vary vertically. Coloured circles denote cue positions that are matched in other editing conditions.

In editing conditions where the focal-object was partially or fully occluded before or after the cut some of the first and last cue positions coincide with the screen edge (Figure 4-9 shows examples of this). In these positions the cue is superimposed on top of the screen edge at the position that matches the focal-object occluded below. This should not change the way that subjects allocate attention to the focal-object as it has previously been shown that attention spreads across the occluder, mirroring the location of the object as if the occluder did not exist (Haimson & Behrmann, 2001;

[64] The labels for these three cue positions refer to the number of pixels from the left edge of the display, incorporating the 120 pixel grey border, to these positions.

Moore et al., 1998; Yi et al., 2003). To ensure that there was no saliency differences between a cue superimposed onto the focal-object compared to the screen edge the same colour was used for both the edge and the focal-object (see Figure 4-9).

4.2.2.8 Experimental Procedure

The experiment consisted of trials testing one of the experimental conditions followed by a short questionnaire for each trial. A single trial contained many shots (a single passage of the focal-object from entry to exit), each presenting one RT cue. Once an RT cue was presented a timer began which registered when the subject hit either of the two buttons corresponding to 'even' or 'odd'. The accuracy and time of their response was recorded. If the subject had not responded after 1000ms (average response times were expected to be around 700ms) and the focal-object had reached the screen edge[65] a new RT cue would be planned. In these instances a null response would be recorded. Cue positions were chosen at random with each of the 13 positions being presented five times during each trial. The average duration of a trial was 3.75 minutes (although this varied depending on how quick the subject's reaction times were). Each subject was presented all nine trials, each followed by the same questionnaire and a voluntary rest period. The entire experiment took between 44 and 55 minutes depending on how long the subjects took to answer the questions and paused between each trial.

The questionnaire asked a series of multiple-choice questions about the subject's impression of the focal-object's motion across the cut. The questionnaire and description of its use can be found in Appendix 2. These questions were intended to indirectly probe the subject's perception of action continuity. The questionnaire provided introspective feedback about the editing conditions that could be used to supplement the reaction time measure when testing the experimental hypotheses.

[65] The quickest the focal-object could cross the screen was 2250ms.

4.3 Experiment 1: Results

Two main types of data were generated for each cue position within an experimental condition: **correct response rates** and **average reaction times.** Incorrect response rates, null response rates, and questionnaire answers were also recorded but these were viewed as supplementary to correct response rate and reaction time. During the design of this study it was not known whether correct response rates or average reaction times would provide the best measure of the attention withdrawal associated with the saccadic eye movement occurring across the cut. Before each of the experimental hypotheses can be investigated a method of analysis first needs to be developed. This requires the correct response rates and average reaction times over all experimental conditions to be examined to see which matches the expected pattern of attention withdrawal.

4.3.1 General statistics

4.3.1.1 Reaction times

Reaction times should be seen to increase as attention is withdrawn from the focal-object in preparation for a saccade. The attentional withdrawal associated with saccadic suppression begins ~75ms (i.e. 2 frames) before the actual eye movement. Attention should be completely withdrawn during the eye movement which lasts 30-60ms, i.e. 1-2 frames (depending on the length of the saccade). Attention then returns to maximum around 50ms (~1 frame) after the saccade ends (Diamond et al., 2000).

The mean reaction time for this study was 675.3ms (standard deviation = 114.4). This was the result of averaging reaction times across all subjects and all experimental conditions. When mean reaction times are generated for each cue position across all subjects a distribution of reaction times across an "average"[66] matched-exit/entrance cut can be seen (Figure 4-10).

[66] the cut is "average" as the focal-object leaves one side of the screen and re-appears at the opposite screen edge but the exact timing of the cut is variable.

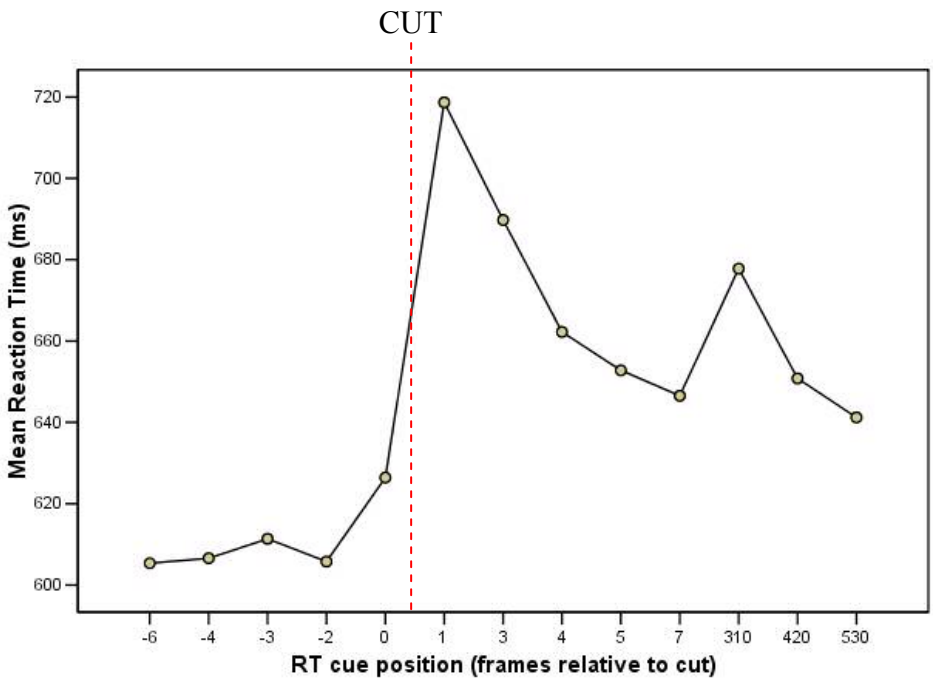


Figure 4-10: Mean reaction time (y-axis: ms) for each cue position (x-axis). Data is averaged across all subjects and experimental conditions.

Figure 4-10 depicts the mean reaction times at each cue position represented as a continuum from pre-cut (-6, -4, -3, -2, 0) to post-cut (1, 3, 4, 5, 7) and screen centre cue positions (310, 420, 530)[67]. The cut occurs between positions 0 and 1. A clear peak in reaction times immediately after the cut can be seen (cue position 1). When the difference between reaction times for each cue position is analysed using a repeated-measures ANOVA there is found to be a significant main effect of cue position ($F=22.81$, $df=4.855$, $p<.000$). This indicates that there is significant difference between the cue positions. Post-hoc comparisons between cue position show that this difference is mostly caused by two main differences: positions prior to the cut, -6 to -2 are significantly (at least $p<.01$) lower than all positions after the cut, and cue positions 1 and 3 are significantly (at least $p<.01$) higher than all other positions (except for position 310[68]).

---

[67] Screen distance between cue positions is consistent for positive and negative positions but 310, 420, and 530 are spaced at different intervals in the centre of the screen (see Figure 4-9).

[68] The slow reaction times at position 310 could be due to the fact that the target has just started moving in an unpredictable fashion as it begins traversing one of the many possible curving paths across the centre of the screen.

The quick reaction times before the cut could indicate that more attention is being allocated to the target immediately before the cut compared to during the rest of its screen time (which includes centre positions 310, 420, and 530 so can not be solely due to the saccade). The main assumption of this chapter is that occlusion by the screen edge *pushes* attention away from the screen edge. This assumption does not seem to be supported by this pattern of reaction times averaged across all experimental conditions. Analysis of the *occlusion expectation* and *advantage hypotheses* are expected to indicate that this lack of effect on attention pre-cut is due to the differences between Exit conditions.

The pattern of reaction times demonstrated in Figure 4-10 seems to be indicative of the expected pattern of attention withdrawal associated with a saccade. However, the mean reaction times are misrepresentative due to null responses. In Figure 4-11, the average correct/incorrect/null response rates across all subjects can be seen. What is clear from looking at Figure 4-11 is that the decrease in correct response rate around the time of the cut is solely due to an increase in the rate of null responses. Incorrect response rates are never very high (mean = 4.31%, s.d. 3.09) and proximity to the cut has very little effect. By comparison, null response rates are normally very low (~5%) but shoot up to ~50% at cue position 1. This indicates that a lot of subjects are either still fixating the exit screen edge or performing a saccade during condition 1. This increase in null responses causes a problem with statistical analysis of reaction times for condition 1. As null response rates increase the average reaction time becomes the product of less correct responses. This introduces more errors in the mean and makes it less representative of a larger population. Also, in some of the Exit groups (as will later be seen), null response rates reach 100% for cue position 1. This excludes cue position 1 from statistical analysis.

Therefore, whilst the reaction time data follows the pattern of attention withdrawal it should only be used for subsequent analysis if the null response rate can be shown to be low within a particular subject group.

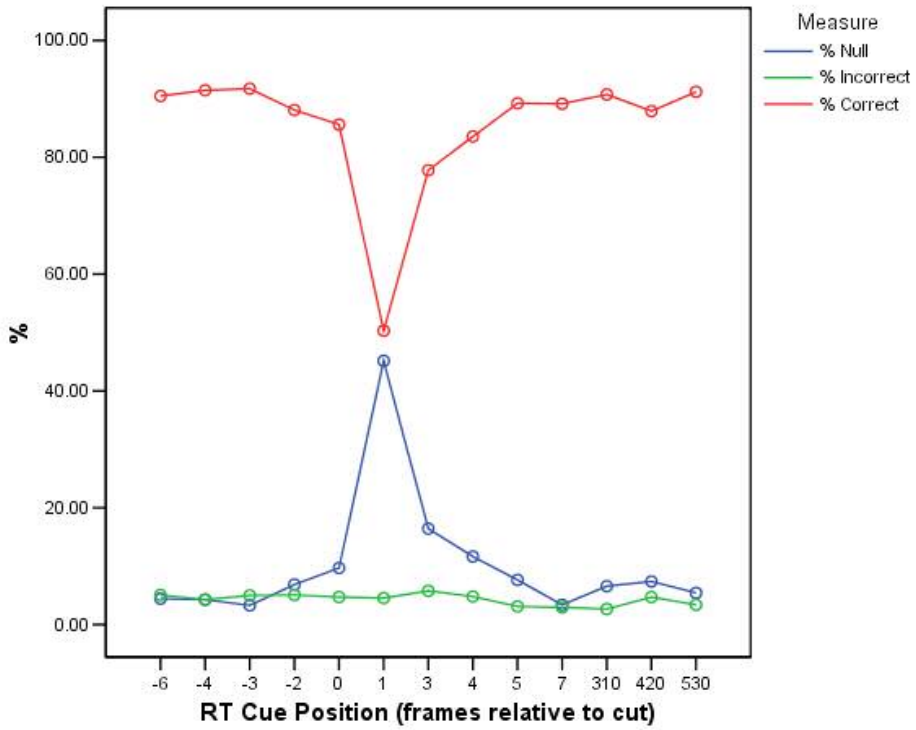


Figure 4-11: Mean percentage correct (red), incorrect (green), and null (blue) responses for each cue position (x-axis). Data is averaged over all subjects and conditions. Negative cue positions are pre-cut, cut occurs between 0 and 1, and positive cue positions are post-cut.

4.3.1.2 Correct Response Rate

The correct response rates should show the opposite pattern of distribution across cue positions compared with reaction times. This pattern can clearly be seen in Figure 4-11 (top red line). The correct response rate is high across most of the cue positions (~90%), decreases suddenly immediately after the cut (~50%), and then takes a few frames to return to the maximum level after the cut[69].

The mean correct response rate across all cue positions and subjects is 85.52% (s.d. 5.36). If the differences between correct response rates for each cue position is analysed using a repeated-measures ANOVA a significant main effect of cue position is found (F=55.276, df=2.985, p<.000). This can be attributed to a

[69] An interesting side point is the absence of an increase in incorrect response rate around the time of the cut. This seems to indicate that subjects were adhering to the instructions of not guessing and were either able to identify the cue as correct or not at all.

significant decrease in correct responses at cue positions -2, 0, 1, 3, and 4. The most significant decrease in correct responses is at position 1 ($p<.000$[70]), followed by positions 3 and 0 ($p<.01$), position 4 ($p<.05$) and 2 (only lower than position -4, $p<.05$, and -3, $p<.05$).

This decrease in correct response rates around the time of the cut lasts 7 frames (position -2 to 4) equivalent to a time period of 292ms. The fact that this decrease starts before the cut indicates that some preparation for the saccade may be happening before the cut as predicted in the *occlusion* hypotheses. As will be displayed during the rest of the analysis the timing of this period of attention withdrawal relative to the cut is highly dependent on the editing conditions.

Before analysis of the hypotheses begins in earnest a quick check for differences is required to ensure that all subjects can be treated as equal within the subject groups.

4.3.1.3 Gender Differences

There is no significant difference between genders for reaction time (female: 635ms, male: 651ms; $t=-.549$, $df=23$, $p=.589$), correct (female: 84.86%, male: 86.04%; $t=-.538$, $df=23$, $p=.596$), or null (female: 10.35%, male: 10.02%; $t=.194$, $df=23$, $p=.848$) response rates. There is also no significant difference when the genders are compared using an independent-samples t-test within each cue position for each of these four measures.

Now that the effect of gender has been dismissed the main interests of this study can be discussed. The four hypotheses will each be addressed in turn: occlusion expectation, occlusion advantage, pursuit initiation, and clearance hypothesis.

[70] The level at which correct response rates for this cue position are significantly lower than all other cue positions when compared in a pairwise comparison of a repeated-measures ANOVA.

4.3.2 Occlusion expectation hypothesis

Viewers expect a cut to occur when the focal-object is fully occluded by the screen edge.

If the subject expects the focal-object to only disappear from view when it is occluded by the screen edge they will exhibit preference for the 100% Exit editing condition. The experimental task demands that the subject track the focal-object at all times so that they can accurately respond to the reaction time cue. Therefore, the subject should time their saccades to coincide with the time at which they expect the cut to occur. If subjects exactly match their saccades to the cut the period of attention withdrawal associated with the saccade should start 2-3 frames before the cut (cue position -2), show lowest attention at the cut (cue positions 0 and 1), and recover by the third frame after the (cue position 4).

To eliminate any effect of learning of the Exit conditions only data from the Random Exit subject group will be used. These subjects were never able to predict exactly where, relative to the screen edge, the focal-object would be when a cut occurred. The focal-object could either be 0%, 50% or 100% occluded. This unpredictability should result in subjects using their default expectation of when the cut should happen. Comparing their correct response rates across the three Exit conditions should indicate what this default expectation is.

The data used in this analysis is averaged over experimental conditions 2, 4, 6, and 8. These conditions all kept Time Gap as 0 whilst varying Entry (0%, 50%, 100%, and Random, respectively). Time Gap was omitted as it was decided that only when there was no wait until the target appeared on the screen after the cut would the viewer be impelled to perform anticipatory saccadic eye movements for the next cut (as suggested by the *Clearance Hypothesis*).

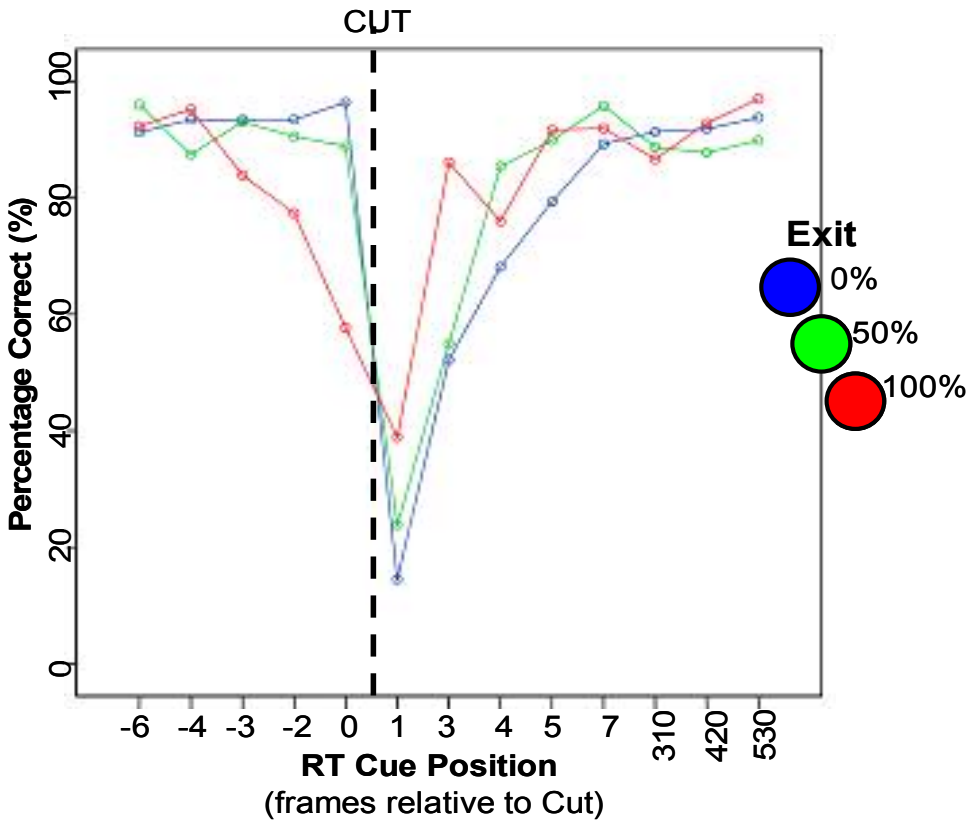


Figure 4-12: Average correct response rate (y-axis: %) for each cue position (x-axis: frames to cut) split across the three Exit conditions used within the Random Exit group: 0% (blue), 50% (green), 100% (red).

As can be seen from Figure 4-12, the data conforms to the previously observed pattern of high correct response rate across the majority of the cue positions with a significant decrease around the cut (see 4.3.1.2). Figure 4-12 also indicates that the time course and onset of this period of decreased accuracy differs between Exit groups. When a cut occurs after the focal-object is 100% occluded, the subjects show earlier attention withdrawal relative to the cut and a quicker recovery back to maximum attention after the cut in comparison to cuts with 0% or 50% Exit. Performing a Wilcoxon Signed Ranks test[71] between cue positions for 100% Exit shows that the decrease in correct response rates starts at cue position -3 (sig. lower than -4; $Z=-1.841$, $p<.05$, one-tailed) and lasts until position 1 ($Z=-2.371$, $p<.05$). By

[71] The decision was made to use non-parametric tests for the analysis of this data as the size of the subject group was quite small (7) and to split the Random Exit condition used in this analysis into the three constituent Exit conditions the data had to be further divided by three. Non-parametric tests are thought to produce more conservative measures of significance for small sample sizes.

position 3, accuracy is back to normal where it stays[72]. These results indicate that a cut occurring when the focal-object is 100% occluded is anticipated by 4 frames (167ms). The total period of decreased performance lasts 208-250ms[73].

By comparison cuts with 0% Exit show no significant decrease in correct response rates until after the cut. Cue positions 1, 3, 4, and 5 all have significantly lower correct response rates than prior to the cut[74] (position 1: $Z=-2.371$, $p<.05$; position 3: $Z=-2.371$, $p<.05$; position 4: $Z=-2.207$, $p<.05$: position 5: $Z=-2.023$, $p<.05$). This period of decreased performance lasts 208-250ms.

A similar pattern can be seen for 50% Exit although the recovery time after the cut is much shorter (see green line Figure 4-12). There is no significant decrease in response accuracy prior to the cut and only cue positions 1 (Wilcoxon Signed Ranks test: $Z=-2.371$, $p<.05$) and 3 ($Z=-2.197$, $p<.05$) have a significantly lower accuracy compared to the level before the cut (position 0, mean 88.9%). This indicates a period of decreased performance lasting only 125ms. Given that voluntary saccades usually take 150-200ms to execute (Rose et al., 2002), this shorter duration could indicate that the saccades performed in response to a cut with 50% Exit could be involuntary. This possibility will be discussed later (section 4.3.3).

These results seem to support the *Occlusion expectation hypothesis* by showing that attention withdrawal only occurs prior to a cut when the focal-object is fully occluded by the screen edge at the time of the cut. This hypothesis can be further supported by showing that, due to the unpredictable nature of the Exit condition for each cut subjects were displaying a single pattern of attention withdrawal across all three Exit conditions. The data used in this analysis was taken from the Random Exit subject group. Within this group, the Exit conditions of each cut were unpredictable.

---

[72] The sudden decrease in correct response rate at cue position 4 is not significant when compared with the surrounding positions.

[73] It cannot be known if performance returns to normal immediately after cue position 1 as there is no cue position 2.

[74] Cue position -2 was used as a comparison (mean = 93.5%).

Given that the subjects do not know when the cut will occur they should be seen to prepare for each cut in the same way. If subjects expect the cut to occur once the focal-object is 100% occluded irrespective of the predictability of Exit, this expectation should be visible as single pattern of attentional withdrawal across the data for all degrees of Exit. The reason such consistency of attentional withdrawal prior to the cut has not previously been seen is that the cue positions were not aligned on the screen. Figure 4-9 shows the cue positions for each Exit condition (right side). Cue position 0 is always the same as the degree of Exit. Therefore, the rest of the pre-cut cue positions are shifted across the screen edge based on the degree of Exit. When cue positions are compared across Exit conditions they represent a position in time relative to the cut, not a position in space. If they are to be taken as a measure of attention at a particular position in space the cue positions need to be re-aligned.

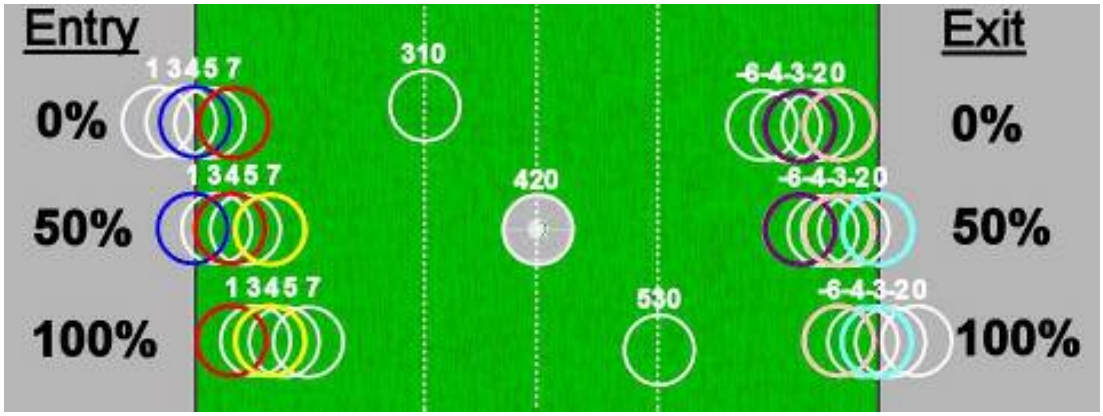


Figure 4-13: The cue positions (circles) associated with 0% (top left and right circles), 50% (middle), and 100% Entry (bottom). All positive and negative positions are always presented on the horizontal middle line of the screen. They are shifted vertically in this diagram for ease of viewing. Cue position 0 always matches the Exit value and position 1 matches Entry value when Time Gap is 0. Positions 310, 420, and 530 always stay in the same horizontal position but vary vertically. Coloured circles denote cue positions that are matched in other editing conditions.

The coloured circles on the right-hand side of Figure 4-9 illustrate cue positions that are spatially aligned across Exit conditions. For example, when the focal-object is touching the inside of the right screen edge (0% Exit) it aligns with cue positions 0 (0% Exit), -3 (50% Exit), and -6 (100% Exit). If all pre-cut cue positions across the three Exit conditions are spatially aligned in this way a single pattern of attentional withdrawal can be seen (Figure 4-14).

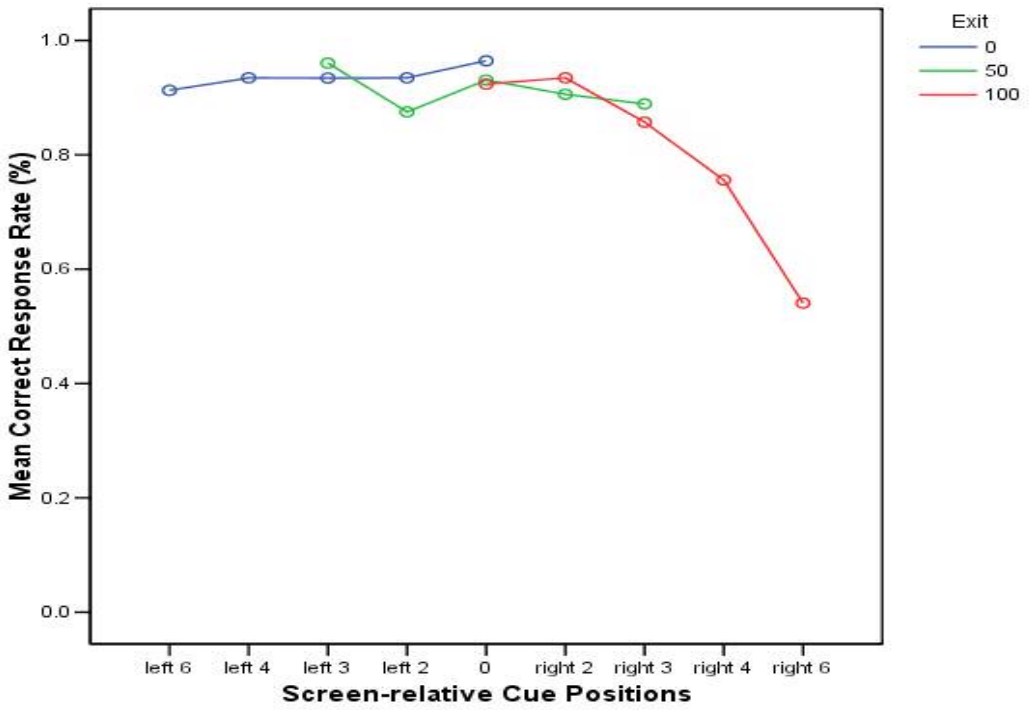


Figure 4-14: Correct response rate (y-axis: %) at cue positions as a function of their relative distance and direction to the screen edge (frames). The three lines represent the three degrees of Exit within the Random Exit subject group (0% blue, 50% green, 100% red) are aligned.

Figure 4-14 shows the accuracy for cue positions realigned in terms of their spatial position relative to the screen edge. Spatial location is represented as the number of frames to the screen edge and the direction e.g. when the focal-object is at 'left 6' it will take 6 frames to reach the screen edge (i.e. be at 0% Exit). This graph clearly shows that subjects are displaying a single pattern of attentional withdrawal. There are no significant differences between correct response rates for the spatially aligned cue positions:

- 0% and 50% Exit at left 3 (Wilcoxon Signed Ranks: Z=-0.736, p=.461, not sig.);
- 0% and 50% Exit at left 2 (Z=-1.095, p=.273, not sig.);
- 0%, 50%, and 100% at position 0 (Friedman: $\chi 2$ = 0.778, df=2, p=.678, not sig.);
- 50% and 100% at right 2 (Wilcoxon Signed Ranks: Z=-1.461, p=.144, not sig.);

- 50% and 100% at right 3 (Z=-0.137, p=.891, not sig.).

Given that there are no between-Exit differences all data can be averaged together within each spatially-aligned cue position. This creates a representation of the average trajectory of attention decrease approaching the cut. This new, spatially-aligned, data can now be used to see how subjects within the Random Exit group withdraw attention in preparation for the cut. Comparing correct response rates between spatially-aligned cue positions only 'right 4' and 'right 6' show a significant decrease compared to position '0' ('right 4': Z=-1.992, p<.05; 'right 6': Z=-2.201, p<.05). Right 4 is the same as cue position -2 under 100% Exit and 'right 6' is the same as cue position 0 under 100% Exit. The absence of a decrease in correct response rates before position 'right 4' suggests that subjects must be waiting until after the target is half occluded (50% Exit) before they start preparing their saccadic eye movement[75]. This seems to confirm the predictions of the *occlusion expectation hypothesis*: viewers expect a cut to occur when the focal-object is fully occluded, not before.

However, there is an alternate explanation for this lack of attention withdrawal before the focal-object reaches 50%: subjects could be choosing to wait until as late as possible before initiating their saccadic eye movement. Such a tactic would increase their chances at seeing all the cues presented before the cut, whilst sacrificing their ability to detect some of the cues presented immediately after the cut (when the cut occurs at 0% or 50% Exit). If this was true then it would suggest that the current pattern of attention withdrawal is not an indication of the subject's preference for 100% Exit. Instead the advantage is just a by product of 100% Exit being the last point at which a cut occurs in our Random Exit condition (i.e. a by-product of the methodology).

[75] The last cue position of 50% Exit spatially aligns with 'right 3' (right-most end of the green line in Figure 4-14). The decrease in correct response rates does not occur until after 50% Exit has occurred i.e. the right most points on the red line.

One way to distinguish between these explanations is to make Exit predictable. If Exit is predictable then subjects should be able to anticipate each cut to the same degree. By comparing performance across the three Exit subject groups (0%, 50%, and 100%) any advantage of 100% Exit will be seen. If the advantage for 100% Exit observed in this analysis was just due to it being the last point at which a cut would occur, performance when the other two Exit conditions (0% and 50%) are predictable should show the same signs of attention withdrawal before the cut. This analysis will be performed whilst testing the second hypothesis**,** *Occlusion advantage hypothesis*.

4.3.2.1 Summary of Occlusion Expectation Results

When the Exit condition is unpredictable (Random Exit subject group) subjects wait until after the object is 50% occluded by the screen edge before preparing a saccade. This is manifest as a significant decrease in correct response rates at cue position -3 (167ms before the cut) and a rapid recovery back to full accuracy after cue position 1 (42ms after the cut). If the cut occurs when the object is not occluded (0% Exit), or half occluded (50% Exit) by the screen edge there is no sign of predictive saccading and the subject takes correspondingly longer to recover full accuracy after the cut, not until after cue position 5 (208ms) and 3 (125ms) respectively. This absence of saccade preparation prior to the cut indicates that subjects would have the potential to be aware of the visual disruption caused by the cut.

When the correct response rates are aligned according to screen position rather than their position relative to the cut, a single pattern of saccade preparation incorporating all three Exit conditions can be seen. Subjects' expect a cut to occur only when the object is fully occluded, not before.

4.3.3 Occlusion advantage hypothesis

Viewers will find it harder to adapt to cuts that occur before the focal-object is fully occluded.

This hypothesis is intended to test whether subjects find it easier to withdraw attention from the focal-object in preparation for a saccade when the focal-object is occluded. The advantage of 100% Exit shown in the previous section should still be seen even when the Exit conditions are predictable. Over repeated presentations it is expected that 100% Exit group will show a greater increase in performance compared with 0%, 50%, or Random Exit group.

To test this hypothesis we first need to split the data into the four Exit subject groups: 0%, 50%, 100%, and Random. The Random Exit group will be kept in this analysis as a control group as the subjects were unable to predict when the cut would occur so they should show less withdrawal of attention pre-cut compared to the other three Exit groups. Subjects were allocated to one of these groups at the start of the experiment and every cut they observed was edited with this degree of Exit. Because the onset of the cuts were as predictable across all subject groups, if there is no advantage of 100% Exit, all Exit groups should show the same degree of predictive saccading. To ensure that subjects performed a saccade as quickly as possible any experimental conditions with a Time Gap larger than 0 was omitted (conditions 1, 3, 5, 7, and 9). The conditions left in this analysis were the same as in the *occlusion expectation* analysis: 2 (0% Entry), 4 (50%), 6 (100%), and 8 (Random).

The mean correct response rate for each Exit subject group (listed in decreasing size) was for 100% Exit 87.18% (s.d. 1.86), 50% Exit 82.31% (s.d. 10.03), Random Exit 81.35% (s.d. 7.25) and for 0% Exit 77.88% (s.d. 2.87). Performing a one-way ANOVA[76] between the groups shows there is no significant main effect of Exit for correct response rate ($F=2.005$, $df=3$, $p=.144$, not sig.). However, there is a significant difference between 0% and 100% Exit (Independent Samples t-test: $t= -6.661$, $df=10$, $p<.000$) and 0% and Random ($t=1.908$, $df=10$, $p<.05$, one tailed).

[76] Parametric tests are used as the distribution of correct response rates was previously shown to satisfy the tests of normality and the data being used in this stage of analysis has an N or 21.

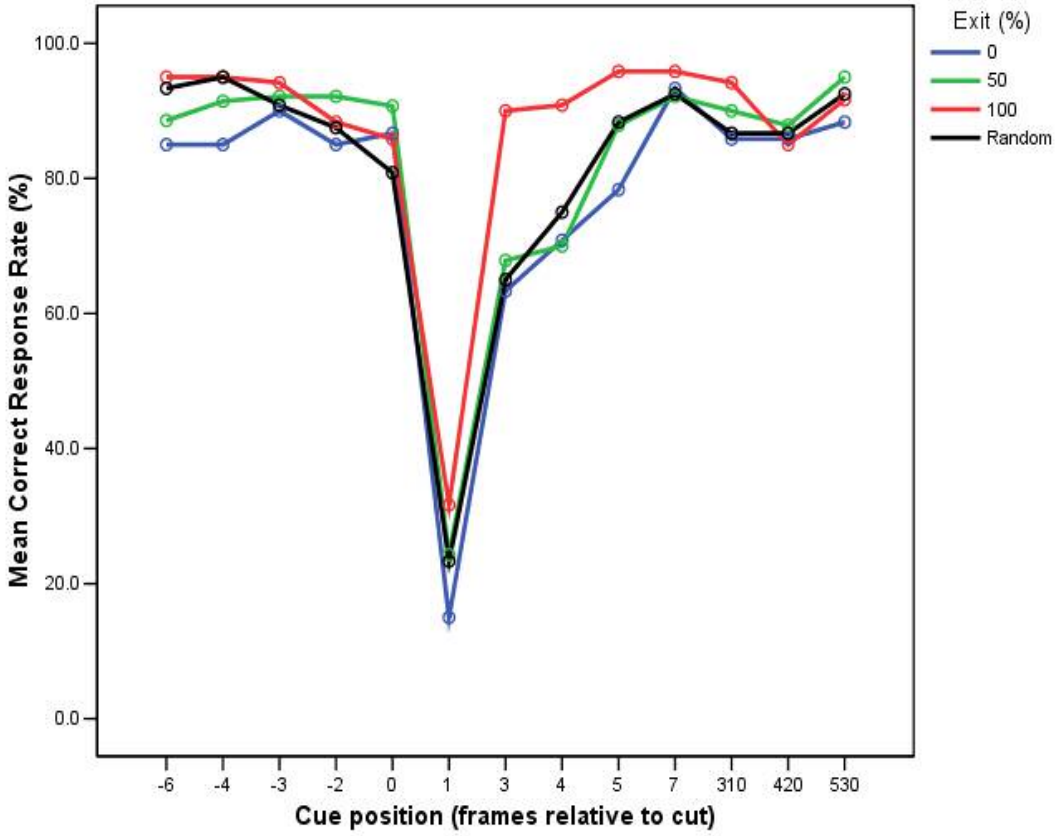


Figure 4-15: Correct response rates (y-axis: %) for each cue position (x-axis: frames relative to cut) across the four Exit groups (0% blue, 50% green, 100% red, Random black).

Splitting the data across cue positions (Figure 4-15) the familiar trend of a decrease in correct response rate immediately surrounding the cut can be seen. However, the pattern varies across Exit conditions. To investigate the pattern of attention withdrawal within each Exit group repeated-measures ANOVA comparing correct response rate across the 13 cue positions.

Within the 0% Exit subject group (blue line Figure 4-15) there is a main effect of cue position (Greenhouse-Geisser: F=27.656, df=3.879, p<.000) which can be attributed to a large decrease in correct response rate at cue positions 1, 3, 4, and 5. This decrease is significant when compared with cue position -3[77] (position 1: mean difference = 75.0, p<.000; position 3: mean diff. = 26.67 p<.05; position 4: mean

[77] Cue position 3 was used as a "baseline" comparison as it has a correct response rate representative of most pre-cut cue positions and is similar across all Exit groups (see Figure 4-15).

diff.=19.167, p<.05; and position 5: mean diff.= 11.667, p<.01). Correct response rate returns to normal between cue positions 5 and 7 (between 208ms and 250ms after the cut). There is no significant decrease in accuracy before the cut. This is a similar pattern as found for edits with 0% Exit within the Random Exit group (see section 4.3.2).

Within the 50% Exit subject group (green line Figure 4-15) there is also a main effect of cue position (GG: F=15.760, df=1.590, p<.01) which can be attributed to a significant decrease in correct response rate at cue positions 1 and 3 only (position 1 compared with -3: mean diff. = 67.86, p<.000; position 3 compared with -3: mean diff. = 24.29, p<.05). Cue position 4 shows no significant difference even though its mean appears lower than the pre-cut average. This lack of effect can be attributed to a high degree of variance. The period of decreased correct response rates for 50% Exit can be interpreted as 125ms.

Within the 100% Exit subject group (red line Figure 4-15) there is a main effect of cue position (GG: F=35.015, df=3.199, p<.000) and this can be attributed to a slight, but significant decrease in correct response rate before the cut at cue position 0 (compared to position -3, mean diff.=-8.33, p<.05) and a large significant decrease in correct response rate at position 1 (significantly lower than all other cue positions, p<.000). There are no other significant differences. This indicates that the decrease in correct response rate lasts 82-125ms (until between positions 1 and 3).

By comparison, Random Exit, which contains all the other three Exit conditions, also shows a significant main effect of cue position (GG: F=19.216, df=1.904, p<.000) but this is due to decreased accuracy at cue positions 0 (p<.05[78]), 1 (p<.000), and 3 (p<.05). This indicates that the Random Exit group shares the same signs of predictive saccading as 100% Exit but takes longer to recover after the cut. The decrease in correct response rates lasts 167ms.

[78] Compared to cue position -6. Cue position -3 could not be used in this Exit group as there was uncommonly large variance in the correct response rates at cue position -3.

This comparison of the duration and position of the decrease in correct response rates around the cut indicates that only 100% and Random Exit show any signs of preparation for the saccade before the cut and that 100% Exit produces the shortest period of attention withdrawal: 100% (decrease in correct response rates lasts 82-125ms), 50% (125ms), Random (167ms), 0% Exit (208-250ms).

4.3.3.1 Learning Effect

The apparent advantage of 100% Exit should also be visible as an increase in performance over the course of the experiment. The experiment was designed to provide two conditions, 1 and 9, that were identical (Random Entry and Random Time Gap) but presented at different times during the experimental session: first and last, respectively. Any improvement in subject's performance between these two conditions can be attributed to learning. It was assumed under the *occlusion advantage hypothesis* that 100% Exit would be the most intuitive editing condition so subjects should be able to adapt to it quicker than other Exit conditions. This should be visible as greater effect of learning.

Over all subjects there is a slight increase in mean correct response rates for condition 9 (mean =87.51%, s.d.= 5.30) compared with condition 1 (mean= 85.72%, s.d.= 6.47) although this was not significant (Greenhouse-Geisser: $F=1.717$ $df= 1$, $p=.204$). However, there is an increase in correct response rate at cue positions 1 and 3 for condition 9 (paired samples t-test: position 1, $t=-2.717$, $df=24$, $p<.05$; position 3, $t=-3.070$, $df=24$, $p<.01$). This increase in performance suggests that all subjects get faster at performing saccades in response to the cut. What is of interest for the *occlusion advantage hypothesis* is if this learning effect is the same across Exit groups.

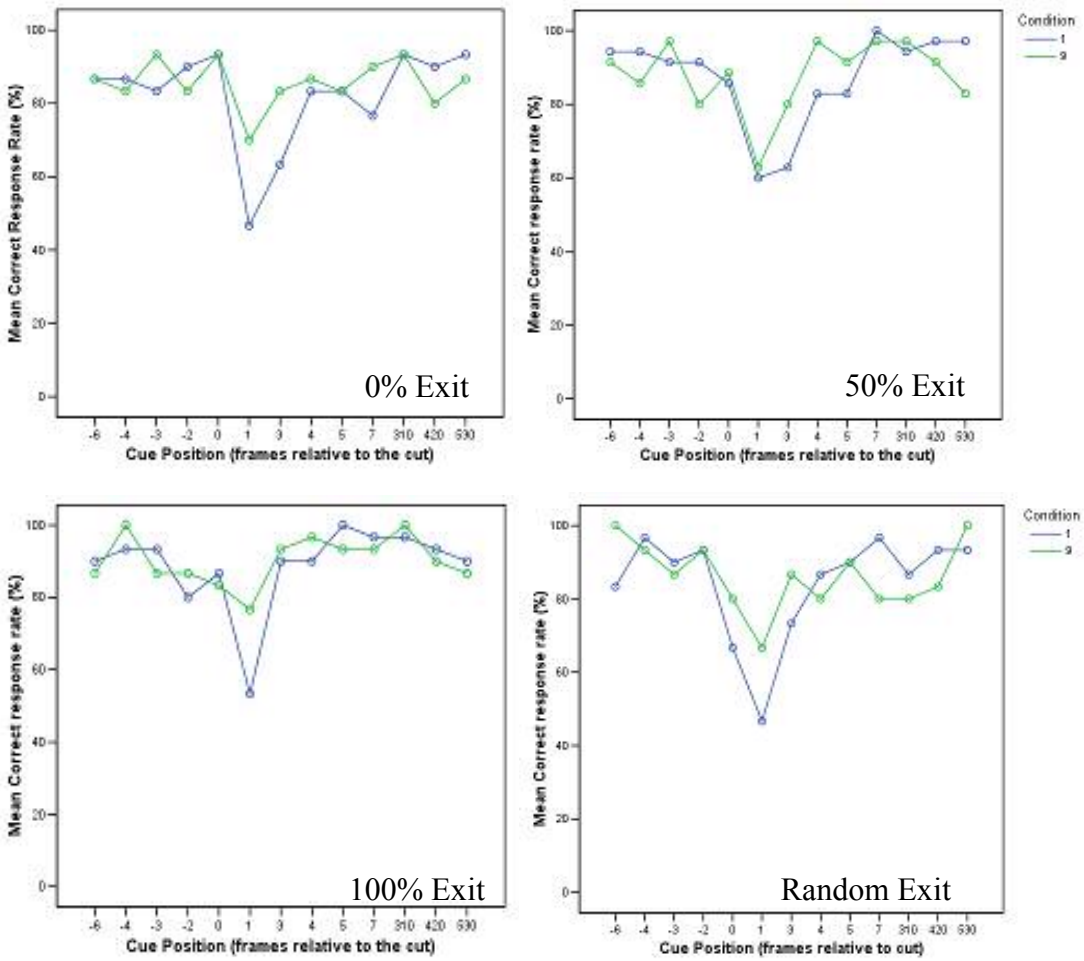


Figure 4-16: Difference in correct response rate (y-axis: %) between conditions 1 (blue) and 9 (green) for four Exit groups: 0% (top left), 50% (top right), 100% (bottom left), Random (bottom right).

If the data is split into the four Exit groups (see Figure 4-16) the different levels of learning can be seen. The 0% Exit group (top left, Figure 4-16) seems to show an increase in correct response rate at cue positions 1 and 3 (i.e. the green line is higher than the blue) but due to large variance this proves to not be significant (position 1: t=-1.472, df=5, p=.201; position 3: t=-1.732, df=5, p=.144). For 50% Exit (top right, Figure 4-16) there is a similar apparent increase in accuracy immediately after the cut but this is only significant for cue position 3 (t=2.521, df=6, p<.05). Unlike all other Exit groups, 50% Exit does not show a difference at cue position 1. Within the 100% Exit group (bottom left, Figure 4-16) the only significant increase between conditions 1 and 9 is at position 1 which sees an increase in mean correct responses from 53.33% to 76.67% (t=-3.796 df=5, p<.05). For Random Exit (bottom right, Figure 4-16) there are no significant differences between conditions.

These results indicate that only 100% and 50% Exit show a significant increase in correct response rates between conditions 1 and 9. For 100% Exit this effect is immediately after the cut (cue position 1) whilst for 50% Exit it occurs later (cue position 3). All groups (except 50% Exit) show an increasing trend in correct response rates at cue position 1 but, due to large variance across subjects, only 100% Exit is significant.

To attempt to extract the learning effect further reaction times at each cue position can be compared between conditions 1 and 9. If subjects were performing better in condition 9 this should be manifest as quicker reaction times.

Comparing the mean reaction times across conditions 1 and 9 there does seem to be signs of learning. Mean reaction times across all cue positions and subjects are significantly lower in condition 9 (mean= 640.42ms, s.d.= 83.41) compared to condition 1 (mean 663.56ms, s.d. 80.60: paired samples t-test: $t=2.646$, $df=24$, $p<.05$). However when the data is split across Exit groups, whilst all groups show a decrease in reaction times between condition 1 and 9, this is only significant in the 100% Exit group (mean diff.=23.42ms; $t=2.8$, $df=5$, $p<.05$).

When reaction times are split across cue positions within each Exit group (Figure 4-17) two distinct effects of learning can be seen. For 50% (top right, Figure 4-17) and 100% Exit (bottom left, Figure 4-17), the peak in reaction times around the time of the cut in condition 1 disappears in condition 9. By comparison the reaction times for 0% and Random Exit remain variable in condition 9 but there is an apparent increase in reaction times before the cut (cue position 0).

Performing paired-samples t-tests between conditions 1 and 9 for each cue position indicates that for 0% Exit reaction times for cue position -2 are significantly faster in condition 9 than 1 ($t=2.892$, $df=5$, $p<.05$) but the apparent differences at cue positions 0 and 1 do not prove to be significant. For 50% Exit reaction times in condition 9 appear considerably faster than in condition 1 but the difference is not

significant. For 100% Exit the only significant difference between conditions is a decrease between 1 and 9 at cue position 1 (t=2.126 df=5, p<.05, one-tailed). For the Random Exit group reaction times are slower than all other groups and reaction times appear to increase between conditions 1 and 9 not decrease as expected. However, the only significant difference between conditions 1 and 9 is at cue position 1 (9>1, t=-2.2m df=5, p<.05, one tailed).

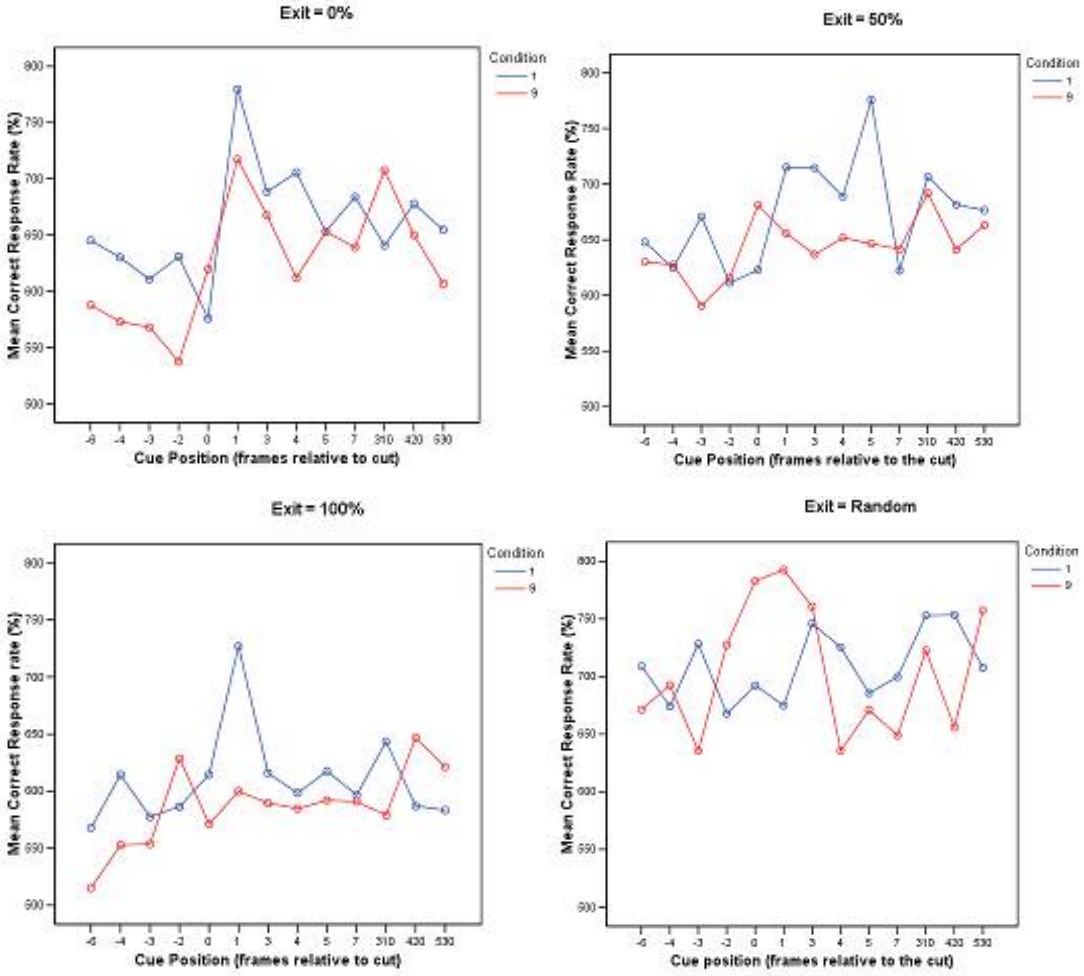


Figure 4-17: Mean correct response rate (%) for conditions 1 (blue) and 9 (red) across cue positions. The four graphs correspond to Exit conditions 0, 50, 100%, and Random (Clockwise from top left).

4.3.3.2 Summary of Occlusion Advantage Results

When subjects are repeatedly presented the same Exit condition they should be able to adapt to the Exit condition and prepare their saccades before the cut occurs. This analysis of the four Exit groups (0%, 50%, 100%, and Random) has shown that only the 100% and Random Exit group show any decrease in correct response rate

(associated with a saccade) prior to the cut. 100% Exit subjects also show the shortest period of decreased correct response rate (82-125ms) indicating that by preparing the saccade before the cut they ensure a quicker recovery after the cut. The 50% Exit group also show quick recovery (125ms) but saccades in this group appear to be made in response to the cut.

When performance is compared between the beginning (condition 1) and end (condition 9) of the experimental session the effect of learning can be seen. All Exit groups (except 50%) show an increase in correct response rates at cue position 1 between conditions 1 and 9 but this is only significant for 100% Exit. The 50% Exit group shows a later significant increase at cue position 3.

In terms of reaction times, only 100% shows significantly faster reaction times in condition 9 compared to 1 even though most Exit groups show a trend in that direction. For 50% Exit subjects appear to be successfully learning the Exit condition as the reaction times are very long and variable in condition 1 but by condition 9 they are quick and consistent. However, this difference is not significant. By comparison, 100% Exit already shows consistent reaction times in condition 1 and by condition 9 the only slow reaction time, cue position 1, has disappeared.

In combination, these results seem to suggest that, in general, 100% Exit produces the clearest signs of saccade preparation before the cut and quickest recovery after the cut. 100% Exit also shows an increase in correct response rates and decrease in reaction times over the course of the experiment. However, as performance for 100% Exit is already very good this improvement is minor. By comparison, 50% Exit shows a significant shortening of the period of decreased accuracy and a trend towards faster reaction times at the end of the experiment. This level of improvement is more pronounced than for 100% Exit as performance under 50% Exit was initially mush worse. Therefore, whilst learning does occur under 100% Exit, the effect of learning is more pronounced with 50% Exit. These results both support the o*cclusion advantage hypothesis* and suggest that the advantage of occlusion may also occur when occlusion is incomplete (e.g. 50%).

4.3.4 Pursuit initiation hypothesis

Viewers expect the focal-object to move back onto the screen from a fully-occluded position after the cut.

This hypothesis predicts that viewers will expect the focal-object to be fully occluded by the screen edge in the frame after the cut (i.e. have 0% Entry) and direct their attention to that position. This will result in more accurate responses to the cues presented at the cue positions immediately after the cut (1, 3, 4, 5, and 7). Performance under experimental conditions 2 (0% Entry), 4 (50%), 6 (100%), and 8 (Random) was compared.

The first evidence of any advantage for 0% Entry should be seen in the correct response rate across all subjects and cue positions. However, all Entry conditions produce very similar correct response rates (0% Entry= 81.72%, 50%= 81.97%, 100%= is 83.08%, and Random= 81.97%) with no significant difference between them (Greenhouse-Gesisser: F=.389, df=2.624, p=.735).

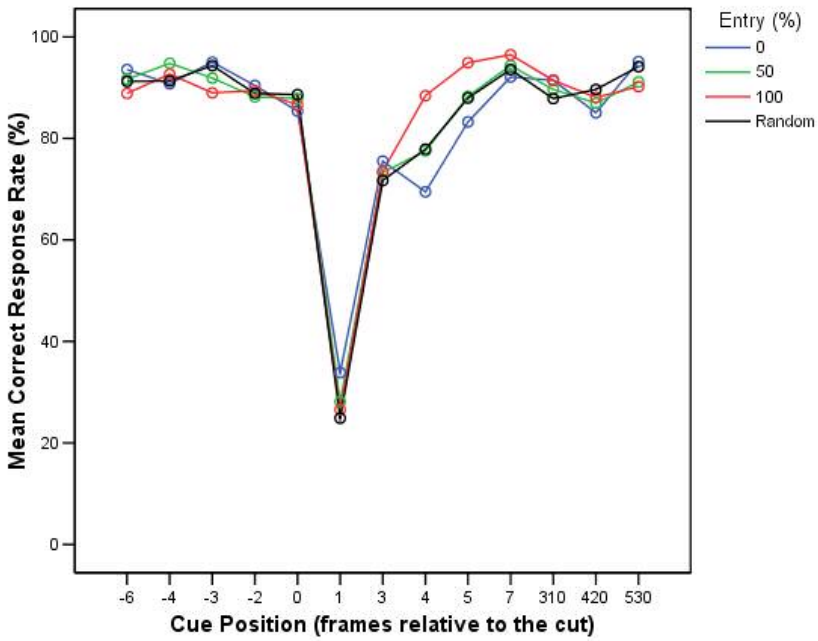
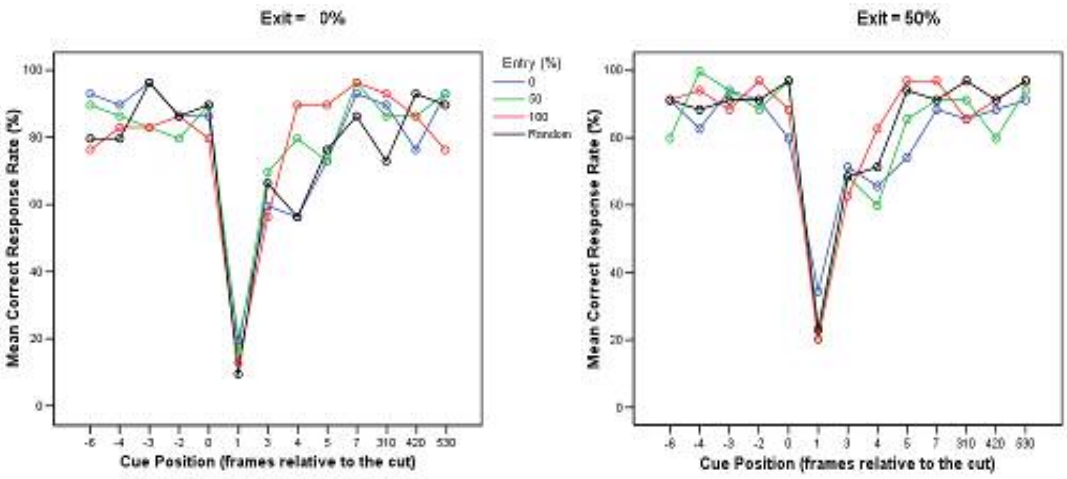


Figure 4-18: Mean Correct response rate (y-axis: %) at each cue position (x-axis: frames relative to cut) compared across 0% (blue), 50% (green), 100% (red), and Random (black) Entry

Figure 4-18 shows the correct response data at each cue position split across the four Entry conditions. Most cue positions show no difference between Entry conditions except for positions 4 and 5. At cue position 4 there is a main effect of Entry (significant repeated-measures ANOVA with Greenhouse-Geisser (GG) correction: F=4.403, df=2.768, p<.01) and this can be attributed to 100% Entry having a significantly higher correct response rate compared to all other Entry conditions (0% Entry: mean difference= 20.119, p<.01; 50% Entry: mean diff.= 11.548, p<.05, one-tailed; Random Entry: mean diff.= 11.190, p<.05).

At cue position 5 there is no main effect of Entry (GG: F=2.184, df=2.824, p=.103) but 100% Entry is significantly higher than 0% (mean diff. = -12.31, p<.05) and Random Entry (mean diff. = -5.0, p<.05, one-tailed). These results indicate that 100% Entry leads to the best performance after the cut. Performance under 0% Entry (blue line in Figure 4-18) was predicted by the *Pursuit Initiation hypothesis* to produce the best performance but it actually takes the longest to return to the pre-cut level of performance.

Analysis for the *occlusion advantage hypothesis* (see 4.3.3) has shown that there is a significant difference between correct response rates for the Exit groups. This difference may be confounding the effect of Entry. Therefore, the data will be split across the four Exit subject groups to see if there are any Exit-specific Entry effects.



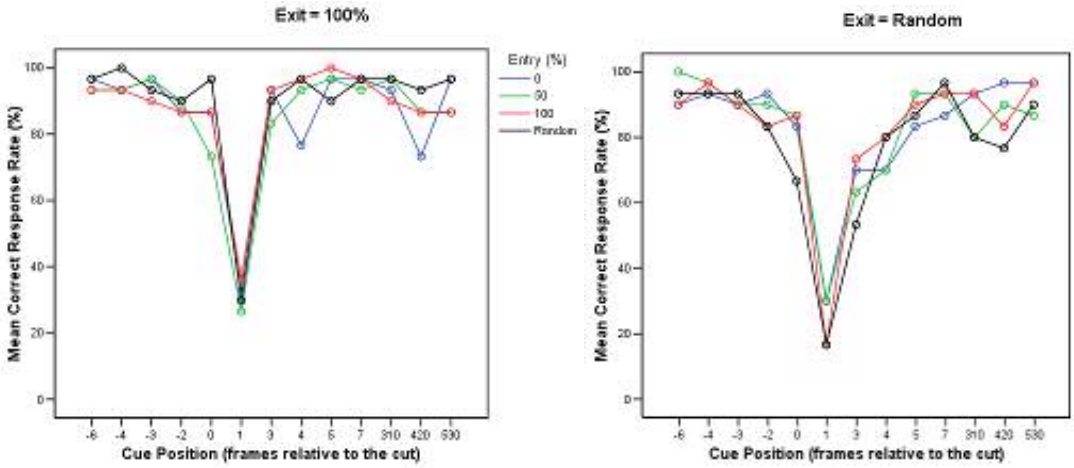


Figure 4-19: Mean correct response rates (y-axis: %) at each cue position (x-axis: frames relative to cut) split across the four Entry conditions, 0% (blue), 50% (green), 100% (red), Random (black). The four graphs indicate different Exit groups: 0% (top left), 50% (top right), 100% (bottom left), and Random Exit (bottom right).

When data is split across Exit groups (as in Figure 4-19) different effects emerge. If we consider the best presentation condition to be that which results in the shortest period of decreased correct response rates surrounding the cut we can identify the best Entry conditions for each Exit. For 0% Exit the best Entry condition appears to be 100% as this returns to the pre-cut level of correct responses quicker (by cue position 4) than any other Entry condition (0%, 50%, and Random do not recover until position 7). Performing a repeated-measures ANOVA between the Entry conditions at each cue position shows that 100% Entry produces a statistically higher correct response rate at cue positions 4 and 5 compared with 0% Entry (position 4: mean diff = -33.33, $p<.05$, one-tailed; position 5: mean diff.=-16.667 $p<.05$) and Random Entry (position 4: mean diff.=33.333, $p<.05$). This indicates that for the 0% Exit subject group, 100% Entry produces the best performance.

For the 50% Exit group (top right, Figure 4-19) it is hard to see which Entry condition is the best. It looks as if 100% Entry may return to pre-cut levels earliest (by cue position 5) but Random also seems to return at this point and 50% Entry's level of correct responses are not far off baseline. Comparing the correct response rate at each cue position to baseline (correct response rate at cue position -4) indicates that 100%, and Random Entry both return to baseline by cue position 4 (100%: $t=1.188$, $df=6$, $p=.280$.; Random: $t=1.441$, $df=6$, $p=.200$). 50% Entry returns

by cue position 5 (t=1.698, df=6, p=.140). By comparison, 0% Entry only shows a significant deviation from baseline at cue position 1 (t=4.599, df=6, p<.01) but this is probably due to the low mean and large variance at cue position -4 (mean=82.86%, s.d.=13.80). Comparing the correct response rates for each Entry condition across all cue positions shows that the only significant difference between Entry conditions is at cue position 0 where 0% Entry is significantly lower than all other Entry conditions (50%: mean diff.=-17.143, p<.05; 100%: mean diff.=-8.571, p<.05, one-tailed; and Random: mean diff.=-17.143, p<.05). This difference at cue position 0 indicates a pre-cut attention withdrawal for 0% Entry. However, this does not result in quicker recovery after the cut. The absence of any post-cut significant differences indicates that there is no single Entry condition that is clearly the best for 50% Exit, although 100% and Random do recover more quickly than the other conditions.

Given the analysis of the *occlusion expectation and advantage hypotheses* (see sections 4.3.2 and 4.3.3), there appears to be a consensus that 100% Exit is the best condition for saccade preparation prior to the cut and shows quick recovery afterwards. If we now look to see if one Entry condition produces the best performance within 100% Exit, we find that there is little difference between the Entry conditions as they all recover very quickly (Figure 4-19, bottom left graph). All Entry conditions show a significant decrease in correct response rates at cue position 1 but all also show a recovery back to baseline by position 3 (0%: t=.000, df=5, p=1.0; 50%: t=1.168, df=5, p=.296; 100%: t=.000, df=5, p=1.00; Random: t=1.464, df=5, p=.203). The only interesting differences between Entry conditions occur at cue positions 0 and 4. At cue position 0, 50% Entry shows a significant decrease in accuracy compared with the baseline[79] (t=2.739, df=5, p<.05) but between-Entry conditions this decrease only proves to be significant from Random Entry (mean diff.=-23.333, p<.05). However, the fact that 50% Entry is significantly different to its baseline does suggest that some degree of saccade preparation is occurring before the cut.

[79] Cue position -4

The other interesting effect is a sudden and isolated decrease in correct response rate at cue position 4 for 0% Entry (4<-4, t=2.712, df=5, p<.05). A similar decrease at position 4 can be seen for 0% and Random Entry under 0% Exit and 0%, 50%, and Random Entry under 50% Exit (see Figure 4-19). Cue position 4 with 0% Entry places the focal-object half on and half off the screen with the reaction time cue presented directly on top of the screen edge. This decrease in the subject's ability to correctly identify the RT cue may be due to the screen edge obscuring the cue in someway. An interpretation of this "edge-effect" will be left until the Discussion section.

Ignoring the "edge-effect", there does not seem to be any one Entry condition which produces better performance than the rest (Figure 4-19, bottom left graph). 50% Entry may show the only signs of predictive saccading but 0%, 100% and Random Entry all result in a shorter period of attentional withdrawal (83ms, cue position 1 to 3). These results are inconclusive.

Looking at the correct response rates for Random Exit it is hard to see one Entry condition that creates the best performance. Paired t-tests between each cue position and -4 ("baseline") shows that all Entry conditions recover back to baseline by cue position 4 with the only difference being that 50%, 100%, and Random Entry all also show a significant decrease in correct response rate at cue position 0 (50%: t=2.236, df=5, p<.05, one-tailed; 100%: t=2.236, df=5, p<.05, one-tailed; Random: t=2.390 df=5, p<.05, one-tailed). Comparing the correct response rates for Entry condition at these cue positions shows that the only significant difference is that Random Entry is significantly lower than 0% at cue position 0 (mean diff.=16.667 p<.05). This clear sign of pre-cut attention withdrawal seems to suggest that for a cut with Random Exit, Random Entry is expected.

4.3.4.1 Summary of Pursuit Initiation Results

This analysis of the Entry conditions indicates how important the Exit conditions are. Whilst differences between Entry conditions can be observed across all data, once

the data is split across Exit groups it becomes apparent that each Exit group has its own preferential Entry condition. Over all the data 100% Entry produces the consistently highest correct response rate. This is the opposite result from that expected under the *pursuit initiation hypothesis* where by full occlusion after a cut (0% Entry) was expected to lead to the best performance.

However, this analysis also indicates that preference for Entry appears to vary across Exit groups. 0% and 50% Exit groups show quickest recovery back to pre-cut correct response levels with 100% Entry but longest recovery with 0% Entry. There is no clear advantage of any one Entry condition for the 100% Exit group as all produce very rapid recovery after the cut. As for the Random Exit group it appears that if the onset of the cue is unpredictable the subject expects the Entry condition to be similarly unpredictable (i.e. Random Entry).

These results do not provide supporting evidence for the *pursuit initiation hypothesis.* 0% Entry does produce quick recovery after the cut under 100% Exit but then so do all the other Entry conditions. To attempt to detect if subjects perceived any difference between these Entry conditions the questionnaire answers can be examined. Question 4 (see Appendix 2) asked subjects to rate the difficulty they experienced performing the reaction time task under each experimental condition. The responses indicate that condition 2 (0% Entry and 0 Temporal Overlap) was rated as the "easiest" condition under 100% Exit. This seems to confirm the *pursuit initiation hypothesis.* However, over all Exit groups, condition 9 (Random Entry and Random Time Gap) was rated as the "easiest". Turning to the question concerning subject's impressions of the "smoothness" of the focal-object's motion across the cut (question 1, Appendix 2), condition 6 (100% Entry and 0 Time Gap) was rated as being the "smoothest" across all subjects and, specifically in the 100% Exit subject group. If subject's response to this question can be interpreted as a reliable measure of their impressions of action continuity[80] then this completely invalidates the *pursuit*

[80] Some of the feedback provided by subjects to the experimenter and the experimenter's own observation suggests that this "smoothness" rating may have been misinterpreted as a rating of the

initiation hypothesis. Instead of subjects expecting the focal-object to be fully occluded after the cut (0% Entry) they appear to find it easier to allocate attention to the focal-object when it is fully on-screen. Whether this is just an artefact of the experimental methodology cannot be deduced from this data.

4.3.5 Clearance hypothesis

Viewers expect the focal-object to begin re-entering the screen immediately after the cut. A time gap between the cut and the object's re-appearance will have adverse effects on attention.

Editors believe that if the frame is left empty at the beginning of a new shot the audience will abandon any expectations they had about the focal-object's motion (referred to as c*learing the frame*; Katz, 1991). This may result in less attention focussed on the object's expected entry location. The experimental conditions of this study were designed to provide four conditions in which the time between the cut and the focal-object's first appearance was varied: condition 2 (no Time Gap), 3 (Random Time Gap), 5 (6 frames Time Gap= 250ms), and 7 (3 frames Time Gap= 125ms). Any difference in attention at the Entry location across these conditions would be observable as decreased correct response rates and increased reaction times at cue positions 1-7.

Figure 4-20 (top graph, page 164) illustrates the correct response rate data for the four Time Gap conditions. Time Gaps greater than 0 appear to make the decrease in correct response rates after the cut smaller and a Time Gap of 6 frames actually increases correct response rates at cue position 1. Comparing the correct response rates at each cue position for the four Time Gap conditions indicates that at cue positions 1, 3, 4, and 5 there are significant differences between all Time Gap conditions (position 1: GG, $F=41.546$, $df=2.195$, $p<.000$; position 3: $F=7.316$, $df=2.79$, $p<.000$; position 4: $F=1.986$, $df=2.506$, $p<.000$; position 5: $F=5.363$, $df=2.144$, $p<.01$). This can mostly be attributed to the 0 Time Gap condition having a

---

focal-object's motion. When the focal-object starts a shot with 100% Entry it appears to accelerate on to the screen. The reason for this will be discussed later (see 4.4.2.1).

significantly lower correct response rate compared with Time Gaps of 3 and 6 frames.

To examine how Time Gap affects the pattern of correct responses across the cut a repeated-measures ANOVA was performed for each Time Gap condition across cue positions. When there is no Time Gap there is a main effect of cue position (GG: F=28.046, df=3.809, p<.000) which can be attributed to a decrease in correct responses at cue positions 0, 1, 3, 4, and 5 (all to at least p<.05). The decrease at cue position 0 indicates pre-cut saccade preparation. When the Time Gap is Random there is a main effect of cue position (GG: f=6.235, df=6.218, p<.000) and a significant decrease in correct responses at cue positions 1 (p<.01), 3 (p<.05, one-tailed), and 7 (p<.05). When 3 frame Time Gap is used there is still a main effect of cue position but it is weaker (GG: F=2.694, df=5.874, p<.05) and the decrease in correct responses only occurs at cue positions 1 and 3 (both p<.05). By comparison, a Time Gap of 6 frames shows no main effect of cue position (GG: F=1.336, df=7.052, p=.237) and there is no significant decrease in correct response rates after the cut.

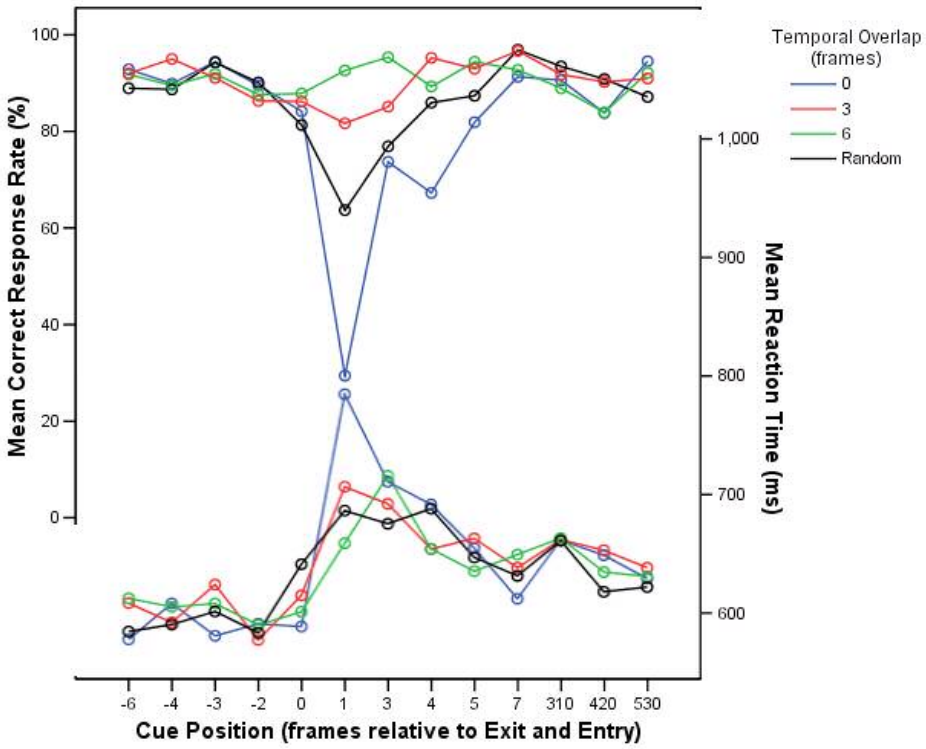


Figure 4-20: Top set of lines represent the mean correct response rate (left y-axis: %) at each cue position split across the four Time Gap conditions: 0 (blue), 3 (red), 6

(green), Random (black). The bottom set of lines represent the mean reaction times (right y-axis: ms) for the corresponding points.

The increase in correct response rates immediately after the cut seems to indicate that as the Time Gap between the cut and the focal-object's reappearance increases attention focussed to the focal-object also increases. This is opposite the predictions of the *clearance hypothesis.* The advantage of a Time Gap may be to give the subject time to complete their saccade before the first cue is presented. This also seems to remove any impetus the subjects have for preparing their saccade before the cut as only the 0 Time Gap condition shows a pre-cut decrease in correct response rates. However, when the Time Gap is 6 frames (250ms) the subject may find that they have to wait for the focal-object to reappear on the screen after their saccade. This lack of an object to which they can focus their attention may have detrimental effects on their performance. No such effect can be seen in correct response rates but reaction times might be more sensitive[81].

Figure 4-20 (bottom graph) shows that the advantage of Time Gap for reaction times is less pronounced than it was for correct response data. All four Time Gap conditions still show an increase in reaction times immediately after the cut. A repeated-measures ANOVA performed within the 3 frames Time Gap condition shows that this increase is significant at cue positions 1, 3, 4, 5, 7 (to at least $p<.05$, one-tailed). There is no increase in reaction times before the cut that might indicate saccadic preparation. The same analysis within the 6 frames Time Gap condition indicates that the increase in reaction times is significant at cue positions 1, 3, and 4 ($p<.05$). For the Random Time Gap condition the increase in reaction times begins at cue position 0 (mean diff.=-50.712, $p<.05$) and continues across all post-cut positions (at least $p<.05$).

[81] Statistical analysis of the reaction time data can be performed as, unlike 0 Time Gap, all positive Time Gap conditions have a null response rate of 0 at cue position 1.

4.3.5.1 Summary of Clearance Results

The *clearance hypothesis* predicted that a Time Gap between the cut and the first appearance of the focal-object might have a detrimental effect on subject's ability to focus attention on the focal-object once it appeared. Initial analysis of correct response rates across the four Time Gap conditions indicated that the reverse was true: as Time Gap increased, performance immediately after the cut improved. In the Random and 3 frames Time Gap conditions performance is still lower than pre-cut levels at cue positions 1 and 3 but when the Time Gap is 6 frames this decrease in performance disappears. However, the reaction time data indicates that attention is not fully focussed on the focal-object after the cut even after a Time Gap of 6 frames.

4.4 Experiment 1: Discussion

This study investigated whether the primitive nature of the matched-exit/entrance cut within the continuity style of filmmaking could be attributed to the screen edge functioning as an attentional *push* cue[82]. Editors believe that when an actor moves off-screen a viewer's attention is "repelled" back across the screen to the opposite screen edge (Dmytryk, 1986; Katz, 1991). This shift of attention provides a period of "blindness" (saccadic suppression) in which the visual disruption of the cut could be hidden.

To find evidence that such an attentional shift is caused by the screen edge a reaction time experiment was used to map attention across a matched-exit/entrance cut. The cut was decomposed into three factors: Exit, Entry, and Time Gap. Exit controlled the percentage of the focal-object occluded by the screen edge at the time of the cut. Entry controlled the percentage of the focal-object occluded by the screen edge immediately after the cut. Time Gap controlled the time between the cut and the focal-object's reappearance on the screen (measured in frames). These three factors were manipulated to test four hypotheses: *occlusion expectation, occlusion advantage, pursuit initiation,* and *clearance hypothesis.* The first two hypotheses relate to the Exit factor, the second to Entry, and the final hypothesis to Time Gap. The results of this study will be discussed in relation to these three factors.

4.4.1 Exit

The results of this study indicate that subjects expect the principal object to fully leave the screen before a cut (100% Exit). Both "occlusion" hypotheses are supported by these results. 100% Exit is expected even if the cut is unpredictable

[82] An attentional cue that shifts a person's attention away from the cue itself. The other main category of cues. *pull* cues attract attention to themselves.

(*occlusion expectation hypothesis*) and subjects adapt most successfully to 100% Exit (*occlusion advantage hypothesis*).

The *occlusion expectation hypothesis* predicted that when the cut was unpredictable subjects would default to expecting a cut to occur once the principal object was fully occluded. This was supported by the results. The pattern of correct response rates prior to the cut indicated that subjects were employing one viewing pattern across all three Exit conditions. Subjects waited until the object was half occluded then began preparing the saccade. This was evident in the significant decrease in correct response rates at cue position -3 in the 100% Exit condition. The total duration of the period of attention withdrawal was 208ms. When the cut occurred before the focal-object was fully occluded the saccade was performed in response to the cut. This resulted in periods of attention withdrawal 208-250ms in length when the cut occurred at 0% Exit and 125ms at 50% Exit. 125ms is insufficient time to prepare and carry out a voluntary saccade (Rose et al., 2002) indicating that the saccades performed in response to a cut at 50% Exit may have been reflexive, i.e. not due to deliberate control by the subject.

The preference for 100% Exit was also observed during analysis of the *occlusion advantage hypothesis*. As predicted, even when 0% and 50% Exit were consistent across all cuts (within Exit subject groups) signs of attention withdrawal prior to the cut were only seen in the 100% Exit group. The 100% Exit group recovered full attention after only 82-125ms. 50% Exit also showed a very rapid recovery back to pre-cut accuracy (125ms after the cut) although this was in response to the cut. This evidence of short duration saccades further suggests that under some conditions saccades were reflexive.

When the effect of learning was examined all Exit groups showed some improvement in correct response rates over the course of the experiment but the differences were only significant for 100% Exit (cue position 1) and 50% (cue position 3). Reaction times were also found to be quicker for 100% and 50% Exit at the end of the experiment although the difference was only significant for 100% Exit.

Both the *occlusion expectation* and *advantage hypothesis* are supported by the results of this study but there is not the clear advantage for 100% Exit that was predicted. Subjects only appear to be able to anticipate a cut when it occurs after the focal-object is fully occluded. This anticipation is seen in the withdrawal of attention associated with a saccade to the opposite screen edge. As the period of attention withdrawal starts before the cut it suggest that subjects are "blind" to a 100% Exit cut. This supports the matched-exit/entrance convention of using occlusion of the focal-object by the screen edge to hide a cut. However, the precise convention used in matched-exit/entrance editing is for the focal-object to be 50% occluded before the cut (4.1.2.2). The results of this study suggest that such a cut would not be anticipated by the viewers. As 50% Exit shows no sign of anticipatory attention withdrawal the viewers should remain aware of the cut. However, the rapid recovery of attention after the cut observed for 50% Exit suggests that the saccadic eye movement performed in response to the cut may be reflexive.

Reflexive saccades do not require preparation as they are usually the result of attentional capture. This leads to saccades with significantly shorter durations than if the saccade was performed voluntarily (Sheliga, Riggio, & Rizzolatti, 1995; Walker, Walker, Husain, & Kennard, 2000). Reflexive saccades under the 50% and 100% Exit conditions could either be a result of the edit screen edge *pushing* attention across the screen in the same way an observed gaze-shift does, or being *pulled* by the sudden appearance of the focal-object on the other side of the screen. Attention could only be *pulled* across the screen if the focal-object was visible immediately after the cut. Therefore, if attention is being *pulled* there should be a clear difference in the duration of attention withdrawal across Entry conditions. This possibility will be discussed in reference to the *pursuit initiation hypothesis*.

Even though 50% Exit shows rapid recovery of attention after the cut there could still be an advantage for 100% Exit. The *occlusion* hypotheses were based on the assumption that full occlusion before disappearance would lead to continued perception of the focal-object across the cut (Gibson et al., 1969). Such continuity of

perception should not occur when attention is captured by the object's relocation (e.g. under 50% Exit) as such sudden object relocation will violate spatiotemporal continuity. Therefore, 100% and 50% Exit may provide two useful but perceptually different types of cut. If an editor's intention is to create a cut with perceived continuity of action they may be able to create this effect with 100% Exit. However, if they just want to ensure that their viewer's attention immediately shifts to the focal-object of the new shot they can use 50% Exit to capture attention. When presenting editing conventions such as those associated with matched-exit/entrance cuts editors always attach a disclaimer stating that the conventions will vary depending on the action depicted and the editor's intention. To be able to identify exactly how editor's use occlusion by the screen edge to cue cuts a survey of matched-exit/entrance cuts would need to be performed and their perceptual consequences judged. Such a survey is beyond the scope of this thesis. Instead the perceptual consequences of such cuts will be discussed in chapter 5.

In terms of the main question of this chapter: *Does a viewer expect a cut to occur as the focal-object is fully occluded by the screen edge?*, the results related to Exit indicate that viewers do expect a cut only after full occlusion. There also seems to be some advantage of 50% Exit but as this appears to be due to involuntary attentional orienting it can not be due to viewers expecting a cut at 50% Exit. The clearest result indicated by analysis of Exit conditions is that 0% Exit is by far the worst place to cut. A cut when the focal-object is still on the screen, irrespective of whether the cut is expected or not, results in a extended period of decreased performance after the cut, less improvement over repeated presentations, and slower reaction times. Whatever the degree, occlusion by the screen edge creates a less disruptive cut.

4.4.2 Entry

The effect of Entry on attention across matched-exit/entrance cuts was investigated under the *pursuit initiation hypothesis.* Based on existing evidence that an object needs to reappear from behind an occluder for it to be perceived as the same object (Gibson et al., 1969; Michotte, 1991; Spelke et al., 1995; Xu & Carey, 1996). The

hypothesis stated that subjects would expect the object to re-enter the screen from a position of full occlusion after the cut. If the object appeared partially or fully on the screen after a cut it was predicted that viewers would have difficulty allocating attention to the object (reflected in low correct response rates immediately after the cut) and perceive the motion of the object as more erratic (rated less "smooth" on the questionnaire).

In general, the results of this study indicated that the opposite was true: locating the object fully on the screen immediately after the cut (100% Entry) led to highest correct response rates and highest "smoothness" ratings. The results indicate that 100% Entry produces the shortest period of attention withdrawal under 0%, 50%, and Random Exit conditions. The condition predicted by the *pursuit initiation hypothesis* to produce the best performance, 0% Entry, actually leads to the longest periods of attention withdrawal. In the 100% Exit group there does not appear to be a benefit of any Entry condition as all produce a very short period of attention withdrawal. In fact, the only decrease in correct response rates for the 100% Exit group occurs in the frame immediately after the cut (cue position 1). This indicates that when subjects are repeatedly presented cuts with 100% Exit they perform saccades across the cut with an average duration less than 83ms. Saccades with such short durations must be reflexive (Sheliga et al., 1995; Walker et al., 2000).

Previously, the benefit of 100% Exit was associated with the voluntary withdrawal of attention prior to the cut (see 4.4.1) and 50% Exit was seen as the main benefactor of reflexive saccades. This evidence indicates that 100% Exit also seems to benefit from reflexive saccades. A voluntary saccade performed in anticipation of a cut would be expected to show a decrease in performance lasting 150-200ms (Rose et al., 2002). The 100% Exit subject group show a decrease lasting an average of 83ms. This occurs even if the focal-object is fully occluded after the cut (0% Entry) and if there are signs of saccade preparation prior to the cut (seen in the 50% Entry condition). Both these conditions indicate that the reflexive saccade begins before the focal-object has relocated across the screen. If the reflexive saccade was only seen in response to a cut with 100% Entry the speed of the saccade could be attributed to

attention capture by the suddenly appearing object (see 2.3.2). The focal-object would be functioning as an attentional *pull* cue. However, when the focal object relocates to a position under the opposite screen edge (0% Entry) it cannot attract attention. This suggests that it is the focal-object's occlusion by the screen edge that is redirecting attention. Based on this evidence the screen edge can be viewed as an attentional *push* cue similar to a gaze shift.

What cannot be known, given the design of this study, is where the screen edge redirects attention to. The convention of relocating the focal-object to the opposite screen edge after a matched-exit/entrance cut seems to suggest that attention is pushed directly across the screen. This could be seen as the opposite to Hirsch's spatiotemporal continuity during occlusion. Spatiotemporal continuity during normal occlusion is believed to function by viewers projecting the focal-object's path prior to occlusion forwards across the space of the occluder to predict where the object will re-appear (Hirsch, 1982). If the same projection occurs in reverse when a focal-object is occluded by the screen edge it would predict that the object would reappear at the point at which this projection intersected with the opposite screen edge. Unfortunately, in this study the focal-object always appeared at the same position along the screen edge so no effect of location (other than Entry) can be seen. To examine if viewers had a preference for the focal-object's re-entry position, attention would need to be probed at various positions around the screen edge immediately after a matched-exit/entrance cut. Alternatively, eye tracking could be combined with an unexpected Time Gap to detect if viewers shifted there gaze to a specific screen location without the focal-object being present.

An alternate explanation of the rapid saccades observed after 50% and 100% Exit could be that a matched-exit/entrance cut functions both as an attentional *push* and *pull* cue. Occlusion by the screen edge may initially disengage attention from the focal-object, *pushing* attention back on to the screen in a distributed form. When either the focal-object or a reaction time cue suddenly appears on the screen the viewer's attention is then captured ("*pulled"*) by the new object. Such a combination of *pushing* and *pulling* of attention would enable the editor to redirect the viewer's

attention to any part of the visual scene, not just the screen edge directly opposite the occluding edge. An empirical study such as that described above would be required to investigate this explanation.

The big question raised by this evidence of reflexive saccades after 100% Exit is: *Do these reflexive saccades maintain the perception of existence constancy created by occlusion?* The sudden, unnatural relocation of an object through space should violate the expectation of spatiotemporal continuity created by occlusion (Hirsch, 1982; Michotte, 1991; Spelke et al., 1995; Xu & Carey, 1996). However, the primitive status of matched-exit/entrance cuts within the continuity editing styles seems to suggest that they do result in the perception of "continuity". Whether the editor's concept of "continuity" can be interpreted as the same concept referred to by perceptual and developmental psychologists will be discussed in Chapter 5:.

4.4.2.1 Flash-lag, Fröhlich, and Edge Effect

One final result emerging from the Entry manipulations should be discussed before moving on: the "edge-effect". Across a lot of experimental conditions, when the reaction time cues were presented 125ms after the cut (cue position 4) a decrease in correct response rates was observed which deviated from the smooth increase in correct response rates either side i.e. cue positions 3 and 5 (see section 4.3.4). This difficulty in reacting to the cues presented so soon after a cut can be attributed to an interaction between two established visual phenomenons: the *flash-lag* and *fröhlich effect.*

The *Flash-lag effect* refers to the illusion that when a flash and a moving object are presented at the same location the flash is perceived as located behind the moving object (Eagleman & Sejnowski, 2000). In this study the effect occurred whenever a reaction time cue was flashed on top of the focal-object. The impression was of the cue lagging behind the focal-object[83]. Given that the velocity of the focal-object was

[83] This illusion was the source of many hours of head scratching for the experiment's designer who checked and rechecked the positioning of the cues before being informed of the *flash-lag effect*.

consistent throughout its screen time, any negative effects the *flash-lag effect* might have on performance should have been consistent across all cue positions.

However, the *flash-lag's* effect on performance may have increased immediately after saccades due to its interaction with another visual illusion: the *fröhlich effect* (Fröhlich, 1923). When subjects are asked to determine the position at which a fast moving stimulus first enters a screen, they will typically mislocate the entry position further along the stimulus' path (Fröhlich, 1923). This phenomenon occurs when the moving target is viewed peripherally (Müsseler & Aschersleben, 1998), pursued foveally after an unexpected onset, or saccaded to after the onset (Yarrow et al., 2005). In this experiment, the *Fröhlich Effect* would have occurred whenever the focal-object entered the screen edge. The impression would have been of the focal-object accelerating on to the screen. This distortion of the focal-object's location combined with the perceived displacement of cue and focal-object position caused by the *flash-lag effect* may have combined to make response to the cue very difficult. In addition to this, cue position 4 for 0% Entry also lay directly on top of the screen edge. The disadvantage this appears to have caused was an unfortunate side-effect of the methodology used in this study. However, the ability to measure attention at so many specific screen locations offsets this disadvantage.

4.4.3 Time Gap

The *clearance hypothesis* was based on a technique editors use for breaking "continuity". A cut is made to a new shot but the shot is left devoid of a focal-object for a few frames. This is believed to make it difficult for viewers to shift attention across the cut and erases any expectation they have about the focal-object's motion (Katz, 1991). To test this hypothesis degrees of Time Gap were introduced between the cut and the focal-object's first appearance on the screen: 0, 3, 6 frames, and Random.

Comparison of the correct response rates between these four Time Gap conditions indicates the opposite effect to that predicted by the *clearance hypothesis*:

performance increases as Time Gap increases. This is clearly caused by the Time Gap providing the subjects with more time to perform and recover from the saccade required to follow the focal-object across the cut. However, when the effect of Time Gap on reaction times is analyzed, the benefit of Time Gaps greater than zero becomes less clear. All Time Gap conditions produce the same pattern of fast reaction times before the cut and slower reaction times after the cut. This suggests that a longer Time Gap does not result in more attention being focused to the focal-object immediately after the cut; just enough to identify the cue. This could be due to difficulty shifting attention to a location without a clear target, maintaining focus on an empty location (with longer Time Gaps the subject has to fixate the screen edge and wait for the object to appear), or just maintaining a high level of attention when the task is undemanding.

The results associated with Time Gaps indicate that, whilst there does not appear to be a disadvantage to waiting for a focal-object after a cut as predicted by the *clearance hypothesis*, the advantage also appears to be less pronounced than the correct response rates first indicate. The experimental stimuli used in this study are probably not ideal to test this hypothesis. The main intention of *clearing the frame* is to eradicate perceptual expectations. By removing the focal-object from the screen when then viewer expects it, *clearing the frame* seems to violate the spatiotemporal expectations associated with existence constancy (Hirsch, 1982; Michotte, 1991; Spelke et al., 1995; Xu & Carey, 1996). The exact mechanism of this can not be explored further in this study due to design limitations, but it will be investigated in the next chapter and empirical study (see sections 5 and 6).

4.5 Experiment 1: Conclusion

The intention of this chapter was to identify how editors use attentional cues to redirect viewers' attention and create periods of perceptual "blindness" in which the visual disruption of the cut can be hidden. A primitive type of continuity cut was identified, a matched-exit/entrance that appears to use occlusion of the focal-object by the screen edge to *push* attention across the screen (Dmytryk, 1986). If this *push* results in a saccadic eye movement initiated before the cut, then the cut will be hidden by saccadic suppression. Also, by occluding the focal-object before the cut the editor may be ensuring that the viewer perceives the object as being continuous across the cut (Gibson et al., 1969; Michotte, 1955). Evidence for such anticipatory attentional shifts and perceived continuity were sought in a reaction time experiment.

Matched-exit/entrance cuts were generated that varied in the percentage of the focal-object occluded by the screen edge at the time of the cut (Exit), the percentage of the focal-object occluded by the opposite screen edge immediately after the cut (Entry), and the time between the cut and the focal-object's reappearance (Time Gap). Attention was probed across the focal-object's entire path using a reaction time task.

The results of this study indicate that subjects are only able to time a saccadic eye movement to coincide with a cut when the cut occurs as the focal object is fully occluded by the screen edge (100% Exit). This anticipatory saccade results in a period of perceptual "blindness" (due to saccadic suppression) which may limit the viewer's awareness of the cut.

Repeated presentation of cuts with the same degree of Exit does not change viewers' tendency to only prepare a saccade in advance of 100% Exit. However, when the focal-object is half occluded by the screen edge at the time of the cut (50% Exit) viewers rapidly shift their attention to the opposite screen edge in response to the cut. These attention shifts can be interpreted as reflexive as they have a shorter duration than is normally required to perform a voluntary saccades (Sheliga et al., 1995; Walker et al., 2000). There are also signs that reflexive shifts of attention occur when the focal-object is fully occluded at the time of a cut (100% Exit).

For most Exit conditions these reflexive attentional shifts seem to be caused by the focal-object's sudden appearance on the screen immediately after the cut (100% Entry). This can be attributed to the focal-object capturing attention. However, cuts with 100% Exit appear to result in reflexive saccades even without the sudden appearance of a focal-object. This seems to suggest that the exit screen edge is functioning as an attentional *push* cue: as the focal-object moves behind the occluding edge attention is detached from the object and redirected back onto the screen.

This evidence both supports and expands the editing conventions associated with matched-exit/entrance cuts. Editors believe that cutting when the focal-object is half occluded before the cut (50% Exit) and half-occluded after the cut (50% Entry) creates the best impression of "continuity of action" across the cut (Katz, 1991). The results of this study indicate that this can be attributed to the sudden appearance of the focal-object *pulling* attention across the cut. Such attentional capture is not usually seen as satisfying the spatiotemporal expectations required for the continued perception of an object (Gibson et al., 1969; Michotte, 1955) but its combination with partial occlusion may result in some form of perceived "continuity". Viewer's rating of cuts with 100% Entry as creating the "smoothest" impression of motion would seem to indicate that some form of "continuity" is perceived across these cuts.

The issue of continuity perception across the cut is also raised when the Time Gap between the cut and the focal-object's first appearance is manipulated. Longer Time Gaps show no detrimental effects on performance in this study but there are some signs that viewers find it hard to focus their attention on the screen edge without a focal-object. Such a Time Gap is believed to *clear the frame* of any perceptual expectations the viewer may have about the focal-object's motion. This study failed to show any effect of Time Gap on perceived continuity as the questionnaire intended to probe this issue failed to be effective. This highlights viewers' inability to reflect on the editing concept of "continuity". More direct measurement of "continuity" is required. However, to achieve such a measurement an understanding

of what actually constitutes perceived “continuity” needs first to be developed. This will be the goal of the next chapter.

Chapter 5:
What is Continuity?

The last two chapters presented evidence indicating that visual attention may be withdrawn from the scene prior to certain continuity cuts. This is an important prerequisite of continuity editing as without an absence of attention the visual transients caused by the cut would capture attention, potentially leading to awareness of the editing. However, whilst the cut itself might not capture attention there is no evidence that the cut's effect on the objects within the scene does not capture attention once the period of withdrawal has ended. Why does the unnatural transformation of objects within the scene not lead to attention capture and how can this discontinuous visual information lead to the perception of continuous action? These questions will be investigated in this chapter. The result of this chapter will be insight into the second main question of this thesis and suggestions on how to answer the third question.

How does continuity editing

1. minimise awareness of a cut,
2. **create the perception of "continuity" across a cut, and**
3. ensure that "continuity" is not violated as a consequence of the cut?

5.1 Preserving Continuity

As previously discussed in the Background chapter (2), the goal of continuity editing is to create the perception of continuous action from discontinuous sensory

information. This is a widely accepted view[84] yet what is actually meant by "the perception of continuous action" has never formally been identified. For most film theorists this definition is in itself sufficient for understanding continuity. Yet such a definition does not permit the constituents of continuity to be identified and manipulated in the deliberate fashion employed by editors. An editor's skill is in their ability to control many different factors of a viewer's experience of a film that may all be seen as contributing to the phenomenon of "continuity". For a detailed understanding of what these factors are and how continuity editing rules utilise them a more comprehensive formalisation of continuity is required.

5.1.1 Continuity Errors

The best place to start looking for the constituents of continuity is where continuity is absent. These instances are variously referred to as continuity errors/mistakes/flaws/goofs/gaffes/slip-ups or production errors. Most filmmakers have the goal of presenting a coherent and believable storyworld that does not appear in anyway artificial[85]. However, as indicated by the number of websites[86] and books (e.g. Sandys, 2005) devoted to the task, audiences love nothing more than spotting the mistakes which undermine this coherence and believability. These errors range from the common and easy to spot, such as a cigarette suddenly changing length or a piece of clothing changing colour across a match-action shot, to the obscure and infrequent, such as the identification of an anachronism[87] by an expert.

---

[84] Similar definitions can be seen in chapter 2.

[85] Of course, not all filmmakers have this goal. Some Avant Garde and experimental filmmakers de-emphasise narrative in favour of other factors such as theme, commentary, or artistic composition. Discontinuities have also been used intentionally by some stylistics groups (such as Nouvelle Vague and Dogme 95) as a way of exposing the act of film creation and its conventionalised mode of communication.

[86] The three main websites are www.moviemistakes.com , www.continuitycorner.com (both dedicated to the task), and www.imdb.com which includes a section of 'goofs' for each film.

[87] A violation of the historical ordering of events or facts.

Continuity errors mostly occur when the person in charge of monitoring continuity, the script supervisor, fails to record a detail or notice its change between periods of filming. In general, continuity errors are very uncommon (relative to the number of cuts in a film) and those that do occur usually involve peripheral details that the production crew were not paying attention during filming. As most errors do not involve objects or characters that were at the centre of attention, detection of errors requires a very inquisitive viewer who may be repeatedly or abnormally watching a film (not just following the plot). Jon Sandys, the author of *Movie Mistakes* (Sandys, 2005) and expert on identifying and cataloguing continuity errors, has stated (in private communication with the author of this thesis) that most errors are initially detected as a "feeling that something isn't quite right" and only by replaying the scene can the error be properly identified. In the process of collating these errors, Sandys has identified several categories of continuity errors that may provide some insight into the constituents of continuity perception. Sandys' categories have been adapted and combined with other categories (www.imdb.com) as well as personal insight to create the following taxonomy of continuity errors:

1) Pre-production: Scripting errors

a) **Factual errors** e.g. errors of real-world geography, facts, and anachronisms that require specialist knowledge to be detected.

b) **Plot holes** e.g. a causal, logical, or dramatic error in the film's plot that "reveal a failure in the consistency of the created fictional world." (Wikipedia, 2005).These might also be introduced when scenes are removed during editing.

2) Production errors

a) **Clothing/appearance** e.g. an actor's top shirt button is closed in the first shot and open in the second.

b) **Stains and marks** e.g. stains miraculously disappear between shots.

c) **Food and Drink** e.g. a drink that never goes down even though it is repeatedly drunk from across shots.

d) **Lights and shadows** e.g. an entire cave impossibly lit by a match or a shadow that moves across shots.
e) **Visible crew or equipment** e.g. the classic "Boom shot": the microphone appears in the shot revealing the produced nature of the scene.
f) **Revealing mistakes** e.g. absences that reveal the artificiality of the film such as rocks being too light, walls obviously made of hardboard, etc.
g) **Audio problems** e.g. audiovisual synchrony problems, absence or incorrect sound effect, or inaccurate stereo. These could be errors of sound recording during production but are mostly caused during post-production.

3) Post- Production: Editing errors

a) **Time/clocks** e.g. a background clock that jumps forwards in time across shots even though the foreground action is continuous or a swimmer who manages to hold their breath for a superhuman length of time.
b) **Repetitions** e.g. an action being shown multiple times (often used to lengthen the duration and increase the impact of action scenes).
c) **Impossible Relocations** e.g. an object changes position within the scene, disappears completely or appears whilst the editing implies no omission of time[88].
d) **Swapping sides** e.g. actors switch which side of the screen they appear on across cuts, the whole scene reverses as the camera crosses the line of action, or the film is flipped over during editing so an actor's left hand becomes their right.

All of these errors would generally be referred to as "continuity errors" but, as can be seen from the groupings, most are not actually caused by a misapplication of the continuity editing rules. Only the errors listed under Post-Production, 3.a (Time/clocks), b (Repetition), c (Impossible Relocations), and d (Swapping sides), seem to be caused by a mismatch between viewer expectations and the action as presented by the editing. Most of the remaining errors are caused by errors during

[88] Sandys referred to these errors as "Continuity" indicating how important they are in comparison to all other errors. Ensuring that these errors do not occur is the primary job of the script supervisor.

script writing (Factual errors, and Plot Holes), or filming (Visible crew or equipment, Revealing mistakes, Clothing/appearance, Stains and marks, Food and Drink, and Lights and shadows). However, as will become clear during the course of this discussion editing plays a vital role in making viewers aware of most continuity errors. To begin understanding why this is the case the continuity errors directly caused by editing, the post-production errors, will first be discussed.

The four post-production errors can be divided into two groups: Time/clocks and Repetitions are caused by a violation of **temporal** expectations, and Impossible Relocations and Swapping Sides are caused by violations of **spatial** expectations. Time/clock and Repetition errors occur when there is a mismatch between the duration of an action depicted in the film and the duration expected by a viewer or implied by the editing. For example, a match-action cut joining two shots of the same action implies temporal continuity (i.e. that time is continuous across the two shots) but if these two shots are filmed in two different locations this spatial discontinuity indicates that time must have elapsed between the two shots. The viewers expect temporal continuity due to the match-on-action and the violation of this expectation creates a temporal continuity error. Every time an editor cuts from one point in time to another, whether it is a millisecond ellipsis during an action or a hundred year jump, they are creating a temporal discontinuity. However, this will not always lead to a temporal continuity error as the viewer may have had no expectation that time would be continuous across the cut. For example, when a scene ends, all action within the scene is brought to a close, removing any expectations the viewer might have about temporal continuity.

Spatial continuity errors, Swapping Sides and Impossible Relocations, occur when viewers expect the 3D space of the scene to be constant or for an object/actor to be located at a certain point on the screen but the film presents it elsewhere. A cut could transport the viewer to a new location, creating a spatial discontinuity, but this wouldn't be detected as a spatial continuity error unless the viewer had expected the second shot to share the same space as the first (i.e. have spatial continuity). The detection of spatial continuity errors is dependent on temporal continuity. If an object

is located on a table in the first shot but on the floor of the same scene after a cut, a spatial continuity error will only be perceived if no time has elapsed between the two shots (i.e. if there is temporal continuity).

By default, viewers expect temporal and spatial continuity. They expect continuity as the real-world is continuous and the actions occurring within it do not contain discontinuities. When these actions are filmed and presented on a cinema or TV screen there is nothing to indicate that this expectation of continuity disappears (evidence for this will be discussed in the rest of this chapter). Therefore, within each shot, the depicted actions will be expected to conform to spatial and temporal continuity. These expectations will naturally extend across cuts. For discontinuity to occur across a cut without it being perceived as a continuity error the editor must remove or distort the expectation of continuity. This can be achieved by using various techniques. The most obvious and explicit way is to use symbolic transitions such as fades and dissolves to indicate an omission of time. The alternative is to communicate to the viewer that the current scene has ended, erasing all expectations of continuity and freeing the editor to cut to a completely new scene. The common way this is achieved is to make all action/dialogue come to a clear end and possibly also indicate a decrease in the significance of the current events by pulling the camera back to a longer shot (Bordwell & Thompson, 2001; Katz, 1991; Reisz & Millar, 1953).

As well as the default expectation of spatial and temporal continuity the large number of production errors (Clothing/appearance, Stains and marks, Food and Drink, and Lights and shadows) also seems to indicate that properties of objects within the scene are also expected to be continuous. This category of expectations will be referred to as **object** and includes properties such as an object or actor's identity and appearance. The existence of these production errors and their frequent identification (they are by far the most common type of error listed on Sandys' website) indicates that viewers often expect object continuity and are sensitive to discontinuities. Within a shot, object continuity is expected by default (e.g. an actor's tie cannot spontaneously change colour) but across a cut the expectation of object

continuity is dependent on temporal continuity. If a fade is used to join two shots of the same actor wearing different clothes viewers cannot perceive an object continuity error as the fade implies temporal discontinuity[89]. In the time between the two shots the actor could have changed their clothing[90]. Similarly, if the same two shots are used but they are joined by a cut and the second shot depicts the actor in a different location the viewer cannot perceive an object continuity error as the presence of the same actor in a different location implies that time has elapsed since the first shot (i.e. object continuity + spatial discontinuity = temporal discontinuity).

Expectations of spatial, temporal, and object continuity mostly occur simultaneously, such as when a cut occurs during the middle of an action the viewer expects the next shot to pick up the action where the previous shot left off, in the same location, and with the same actor. However, unlike in reality, these expectations can occur independently. The best example of this is viewers' acceptance of 'crosscutting' (see Bordwell & Thompson, 2001; Reisz & Millar, 1953 for discussion of the technique). This technique allows multiple simultaneous events taking place in different locations to be shown by serially cutting between them. Each cut creates a spatial discontinuity as the locations are changed but temporal continuity can be preserved by repeatedly cutting back to each location and showing that time has only advanced as much as was presented in the previous shot. The clear separation between the locations (spatial discontinuity) and actors (object discontinuity) ensures that there is no indication that time has elapsed between the events. If the same actor were to appear in two different locations the cut joining these two locations could not be seen

---

[89] This is assuming that the viewer understands the symbolism of a fade. An uninformed viewer or a viewer who was momentarily distracted might not be able to 'read' the transition and will fail to interpret a passage of time. Without the fade the viewer would have to deduce a passage of time from the impossible change of clothing.

[90] Another technique often used (see Reisz & Millar, 1953) is to insert another shot between the two shots with the object continuity error. Editors have learnt that by focussing the viewers' attention on some other object (e.g. another character in the scene) when the discontinuous object is later presented viewers do not notice the continuity error. This technique indicates that attention seems to be important. This idea will be expanded and explained during the rest of this chapter.

as temporally continuous (i.e. object continuity + spatial discontinuity ≠ temporal continuity).

However, there does appear to be one exception to this rule: matched-exit/entrances. A matched-exit/entrance, as used in the first experiment, depicts an object (typically an actor) leaving the screen in the first shot and then re-entering in the next shot. The two locations are typically different indicating a spatial discontinuity[91]. However, the matched-action of the same object departing the screen and then re-entering implies temporal continuity which is incompatible with spatial discontinuity[92]. This incompatibility will only be noticed if the viewer has existing knowledge of the spatial relationship between the two depicted locations and can use this to detect the spatial discontinuity. This difference between absolute continuity and perceived continuity will be expanded during the rest of this chapter.

Before dissecting the three dimensions of continuity (temporal, spatial, and object) here is a complete listing of the categories of cut created by mixing continuity and discontinuity across the three dimensions (Table 2)[93]. This is intended as a point of reference during the remainder of this chapter. Each category is defined after the table.

---

[91] Different in the sense that the space depicted in the first shot is not also depicted in the second shot. However, the level of spatial discontinuity can be any size, for example a matched-exit/entrance cut can occur across locations physically separated by thousands of miles.

[92] A spatial discontinuity refers to two locations that are not the same and not adjacent. In short, two locations that cannot be moved between without a passage of time (i.e.a temporal discontinuity).

[93] The names chosen for the eight categories of edit (right-most column of Table 2) are just intended as representative examples of a type of cut using that particular mixture of the three continuity dimensions. Other types of cut could also share the category.

No.	Temporal	Spatial	Object	Cut Categories
1	continuity	continuity	continuity	Match-action
2	continuity	continuity	discontinuity	POV or Reverse
3	continuity	discontinuity	continuity	Matched-Exit/Entrance
4	continuity	discontinuity	discontinuity	Crosscutting
5	discontinuity	continuity	continuity	Ellipsis or Repetition
6	discontinuity	continuity	discontinuity	Montage
7	discontinuity	discontinuity	continuity	Following or Flashback
8	discontinuity	discontinuity	discontinuity	Establishing shot

Table 2: The three dimensions of continuity and the resulting categories of cut.

Definitions:

1. **Match-action:** The action depicted in the shot before the cut is continued in the shot after the cut. The viewer's default expectation of continuity across all three dimensions is maintained.
2. **POV or Reverse:** Any cut that moves the camera within a scene so that the object of the previous shot is no longer in view. The new shot can either represent the point-of-view (POV) of the previous object/actor or the camera can be positioned somewhere close to the object's position and pointed in the opposite direction (Reverse). The presence of the previous object is implied by temporal and spatial continuity and may even be indicated by other actors interacting with the off-screen object.
3. **Matched-exit/entrance:** The object (typically the main character) departs the screen in the first shot and is depicted entering a new scene in the next shot. Whilst the change in scene indicates a spatial discontinuity the match-on-action implies continuity. This technique can be used to construct impossible spaces from spatially non-adjacent shots of the same action. If the spatial discontinuity is perceived then temporal discontinuity will be assumed resulting in a Following shot.

4. **Crosscutting:** Multiple events are shown simultaneously by serially cutting between them. Each event has to involve different objects otherwise temporal discontinuity will be perceived resulting in a Following shot or Flashback.
5. **Ellipsis or Repetition:** Time is omitted or repeated across a cut but the absence of a change in location or object makes it hard for the viewer to detect the temporal discontinuity. This technique is used to shorten or extend the duration of actions. If too much time is repeated or omitted artefacts may start to appear in the object or spatial information (e.g. an actor's change in clothing or position in the scene) which would lead to a Time/clock or Repetition continuity error.
6. **Montage:** The object is intentionally changed across a cut to imply a passage of time. A classic example of this is the pages of a calendar quickly ripped off or the sudden setting or rise of the sun.
7. **Following shot or Flashback:** The object of the previous shot is followed into a new location (Following) or backwards in time to a previous event (Flashback). A match-on-action or matched-exit/entrance is typically not used so that the temporal and spatial discontinuity is clear. If the temporal discontinuity is large the object may start to change in appearance (e.g. old-fashioned clothing) and the temporal relationship between the shots would have to be re-established (category 8).
8. **Establishing shot:** This is the default category of cut that all other categories become when a viewer's assumption of continuity is violated. If the viewer becomes unclear whether the cut has spatial, temporal or object[94] continuity the cut will be perceived as a category 8 and these relationships will need to be re-established.

[94] Viewers can become unsure of object continuity if the object from the previous shot seems to have changed significantly. A good example of this is a flashback to a character's youth. The change from adult to child might make the character unrecognisable so their identity would need to be re-established. A classic way of doing this is to have some other character say the main character's name as soon as the new shot starts. Once the identity is established then the viewer perceives the cut as a category 7, a Flashback.

What is apparent from this discussion of common continuity errors is that all editing errors and most production errors (excluding Visible crew or equipment, Revealing mistakes, and Audio problems) are caused by a mismatch between what the viewers expect to see and what the film depicts. These expectations occur along three dimensions: **object, spatial,** and **temporal**. Object seems to refer to all properties that allow one object to be distinguished from another object. These include visual properties such as surface colour, texture, pattern, as well as the object's shape and semantic information such as its name (e.g. 'Rover'), and type (e.g. 'Dog'). Spatial refers to the position of this object within the scene and Temporal, to its position in time relative to the previous shot. As can be seen from Table 3, these expectations seem to vary over time enabling editors to create eight different categories of cut. The main hypothesis of this chapter is that continuity editing functions by accommodating perceptual expectations about the visual event so that discontinuity isn't perceived after the cut. To find evidence to support this hypothesis the factors controlling the variation in these expectations needs to be understood. If evidence of these factors can then be found in the suggestions of continuity editing rules this hypothesis can be said to be supported.

Evidence for these three dimensions of continuity can be found in existing research areas. The distinction between spatial, temporal, and object information has previously been used when investigating infants ability to perceive objects during occlusion (referred to as existence constancy; see 4.1.3.2). It can also be seen in theories of object perception, change blindness, transsaccadic memory, and object-based attention. The rest of this chapter will entail a quick survey of each of these areas. All evidence relevant to the understanding of how these expectations are constructed, maintained, and influenced by such factors as viewing conditions and attention will be recorded. This evidence will then be used to formalise an empirical test of continuity perception across various categories of cut so that the assumptions of continuity editing can be assessed

5.2 Object Continuity

The majority of continuity errors previously listed (see 5.1.1) seem to result from a mismatch between how the viewers expect an object to appear and the way that it is depicted after a cut. Some continuity errors, such as Impossible Relocations may involve changes to the entire scene but, as indicated by the category of cuts in Table 2, even these spatial continuity errors are mostly detected when the viewer first notices an object discontinuity (e.g. the previous object is no longer present). As such, to begin identifying what constitutes viewer's expectations it seems best to start with the most noticeable category of expectations: object continuity.

Perceiving object continuity requires the retention of information about the object's properties across a cut. Understanding which object information is retained across a cut first requires an understanding of what it is that defines *objecthood*. Adults experience an object as an entity that persists over time even though the sensory referent for that object may be intermittent as the object and its viewer move about the 3D space (Spelke et al., 1995). The viewer must retain some form of conceptual representation which allows them to identify the object once it has come back into view. To identify what constitutes this representation a paradigm is required that manipulates object features and detects whether viewer's notice the manipulations. Such manipulations have been performed by change blindness studies.

5.2.1 Change Blindness

We are not as aware of our surroundings as we think we are. This is the conclusion of the field of research known as Change Blindness. Adults, whilst believing they are highly aware of the world around them, often fail to notice large changes to objects when they occur during brief distractions (e.g. occlusions, saccades, blinks, mud-splashes, cuts, etc; see Simons and Levin, 1997 for a review). For example, 50% of observers failed to notice when two cowboys sitting on a bench exchanged heads (Grimes, 1996). If the same change occurred without the observer's view of the stimuli being obscured they would instantly be aware of the change (Rensink et al.,

1997). This is because the sudden change is accompanied by visual transients such as changes in colour, brightness, contrast, or motion signals that under normal viewing conditions capture attention (Simons & Rensink, 2005). The removal of these transients and the resulting inability to detect changes to the scene and objects within it indicates how much emphasis our perceptual system places on the transients to indicate when a change has occurred. However, most change blindness experiments actually set subjects the task of detecting a change and they still take an inordinate length of time to do so (e.g. Rensink et al., 1997). This seems to indicate that viewers only retain a very sparse representation of the scene and its objects under normal viewing conditions (Rensink, 2000). If this is the case, achieving object continuity across a cut should be relatively easy as observers only *"encode the gist of the scene (in this case, the specific action and a few characteristics of the actor) and ignore the visual details. As long as the gist remains the same, change detection is unlikely as observers have not expended the effort to encode more detail"* (Simons & Levin, 1997; page 266).

The startling findings of these initial change blindness studies led some theorists to claim that little or no visual representation was necessary for scene perception (O'Regan, 1992; O'Regan & Noë, 2001; Rensink, 2000; Rensink, O'Regan, & Clark, 2000). This conclusion was supported by evidence showing that observers were slow to detect changes to an object's identity (Levin & Simons, 1997; Angelone et al., 2003), colour (Levin & Simons, 1997; Rensink et al., 1997; Angelone et al., 2003), position within the scene (Levin & Simons, 1997; Rensink et al., 1997), orientation (Henderson & Hollingworth, 1999; Hollingworth & Henderson, 2002b), as well as the sudden disappearance or appearance of objects (Rensink et al., 1997). However, recent studies have shown that this insensitivity functions differently for different types of changes (Brockmole & Henderson, in press) and for the location of the change relative to the saccade target (Henderson & Hollingworth, 1999; Henderson & Hollingworth, 2003) or point of fixation (Rensink et al., 1997).

It has also been shown that even in the absence of explicit change detection, incidental memory for the changed item is above chance (Angelone et al., 2003).

This indicates that subjects are retaining some degree of information about the objects but, for some reason, they are unable to use this information to detect a change. Similar evidence of relatively detailed representations generated incidentally has been shown in visual search experiments. When subjects were instructed to count the number of white telephones in an object array their incidental memory (not told to memorize the arrays) for all non-telephones that shared the target property "white" was significantly better than other non-white distractors (Williams et al., in press). Recognition accuracy also increases as the number of fixations received by each object increases (Williams et al., in press). These results indicate that detailed visual information of fixated and task related objects is retained long after the object has disappeared from view.

5.2.2 Change Blindness during film viewing.

Whilst change blindness is a useful paradigm for indicating if viewers are aware of changes to a particular object feature if viewers fail to notice the change it does not indicate that the feature isn't represented. Recent evidence of implicit change detection suggests that change blindness occurs not because of a failure to represent the object, but because the new information is not compared to this representation (Hollingworth & Henderson, 2002a; Mitroff, Simons, & Levin, 2004). This comparison is believed to require attention and, given that the visual disruptions used in the change blindness experiments (e.g. blank frames, saccades, blinks, mudsplashes) remove the visual transients which would usually capture attention the comparison does not occur (Rensink et al., 1997).

The visual disruptions usually occur during the presentation of an approximation of a real visual scene e.g. a static photograph. As such, the viewer may assume that object information will be continuous as object discontinuities do not occur in reality. However, one type of visual disruption used in change blindness experiments does not occur whilst viewing a continuous visual scene: a cut (Levin & Simons, 1997). In these experiments a deliberate change to a single object e.g. change of identity or clothing was made across a cut (see Figure 5-1). Even when this change was

significant, such as the change of actor across a matched-exit/entrance cut (Levin & Simons, 1997), only 33% of subjects noticed the change. The difference between these studies and other change blindness studies is that the cut does not provide a period of occlusion during which the change can occur. Instead the cut produces a whole-field change with a mass of visual transients: apparent rotations, relocations, changes in size, etc. The visual transients of the target object are obscured by the scene transients. This means that the probability that attention will be captured by the transients of the target object should be significantly less than with static stimuli.

Figure 5-1: an example of identity change across a matched-exit/entrance cut from Levin and Simons' (1997). The sequence is constructed from two shots (AB and CD) joined by a cut.

However, in the 'Hiding a Cut' chapter eye tracking evidence was presented indicating that during or immediately after a continuity cut most viewers saccade to the same position on the screen. This position usually relates to an object that was either present in the previous shot or has been previously introduced. If, as indicated in the last two chapters ('Hiding a Cut' and 'Experiment 1'), the cut occurs during the saccade, when attention returns to the scene the target of the saccade will have changed significantly from how it appeared before the saccade. This suggests that in Levin & Simons' change blindness film the changed object (e.g. the face of the changed actor) should have received full attention after the change and given that incidental memory has been accumulated for the object during previous fixations

(Angelone et al., 2003) all the prerequisites for change detection should be in place (Rensink et al., 1997).

The fact that the rate of change detection is still very low when the change occurs across a cut suggests that, whilst attending to an object is necessary for change detection, it is not sufficient (Levin & Simons, 1997). Some other factor must exist for the viewer to perceive a discontinuity. Levin & Simons' evidence seems to suggest that the spontaneous deformation of object properties during a cut is insufficient to cause the perception of object discontinuity. An object's properties must, therefore, be represented in a form resilient to these deformations. To understand how object properties can be represented in such an amorphous form and to identify what factors cause the perception of discontinuities a theory of object perception is required.

5.2.3 Object Files

One theory of object perception that has received a lot of support over the past 15 years (Enns et al., 2001; Henderson, 1994; Irwin, 1992; Leslie, Xu, Tremoulet, & Scholl, 1998) is Kahneman and Treisman's (Kahneman & Treisman, 1984) theory of **object files**. Object files are temporary "episodic"[95], mid-level[96] representations of objects and events. They are regarded as "files" as they are repositories of information that can be added to or updated at any point much in the same way a police case file will be updated as new information about a crime becomes available (Palmer, 1999). Files are initially created when a feature of the object captures attention (such as its abrupt onset of motion; Enns et al., 2001) or the object is first fixated. This focussing of attention on the object "binds" together all the visual features occurring in that location (Kahneman et al., 1992). This binding is important as before attention is focussed on an object there is evidence to suggest that the object's visual features are stored as a "shapeless bundle" (Wolfe & Bennett, 1997).

[95] Specific to the point in time that they were established.

[96] More than the low-level independent representations of sensory factors such as colour but less than the high-level of concepts and semantic associations.

The precise mechanism that allows this amorphous collection of sensory signals processed by different areas of the visual system to be joined together into a recognisable object is not known but there is increasing evidence that focal attention is important (Treisman & Gelade, 1980; Wolfe & Bennett, 1997).

Object files are addressed according to their location at a particular time not by any feature or identity label (Kahneman et al., 1992). The information is stored as abstract structural descriptions which can be matched to long-term object representations to identify or classify the object. However, association of a semantic label with a file is not essential. When an object file is first opened it may just consist of sparse information about vague form and location. As more information becomes available this will be added to the file, changing the objects identity whilst maintaining the same object file. This allows an object file to evolve over repeated fixations[97]. This feature of Kahneman and Treisman's object file theory is key to active dynamic object perception.

The number of total object files open at any one time is limited to around 4 (Kahneman et al., 1992). Whenever a new visual input is detected it is first checked against existing object files to see if it refers to the same object and if this fails a new object file is allocated to it. If the maximum number of object files already exists one of the existing files must be abandoned before the new object can receive a file. This checking of existing files is known as *correspondence* and has to incorporate expectations of acceptable transformations of stored information (Ullman, 1979). If the new information is seen to represent an acceptable change (e.g. through object or viewer motion) a *reviewing* process retrieves the matching object file. The existing

[97] A famous film example of this is Omar Sharif's entrance into Lawrence of Arabia (1962). Sharif's character is first seen as a black blob on the horizon surrounded by a dust cloud and blurred by the heat haze of the desert. As Peter O'Toole's Lawrence gazes expectantly at him more detail is revealed and we begin to see that he is a black-robed Bedouin riding a camel. As he nears he shoots Lawrence's companion and the dialogue of the scene begins. The long static shot emphasises the importance of Sharif's character as an indistinct object gradually accumulates information until at the end of the scene he is a richly detailed character.

content of the object file is then combined with the new information through a process known as *impletion* to create a percept for the change or motion connecting the two (Shepard, 1984).

If the object moves behind an occluder or its visual features are temporarily unavailable (e.g. during a saccade away from the object) object files can stay open for at least 600-700ms (Kahneman et al., 1992) and possibly even beyond 8 seconds (Noles, Scholl, & Mitroff, 2005). However, in the absence of visual reference the object file is susceptible to erasure if a new object captures attention (Enns et al., 2001). To ensure that an object file is maintained after a visual disruption, such as a period of occlusion, saccadic eye movement, or even a cut, the object would need to be immediately present in the scene after the disruption and similar enough to the features stored in the object file for correspondence to succeed. If the *correspondence* process fails (Ullman, 1979) or the object cannot be found an error signal is generated that could result in the viewer becoming aware of the change (overt change detection) or just prioritising the source of the change for closer examination (e.g. increased fixation durations and tendency to be fixated; Brockmole & Henderson, in press; Henderson & Hollingworth, 2003).

Under normal viewing conditions, change blindness shows us that viewers are not predisposed to overt change detection. This could either indicate that correspondence is not performed between the post-disruption information and that of the *object file* or the lack of correspondence is insufficient for the change to trigger an error signal. Distinguishing between these two interpretations is not possible given the evidence from change blindness. Even if viewers are encouraged to perform correspondence in a change detection task (e.g. Levin & Simons, 1997) or they are alerted to the presence of change by the accompanying visual transients (Brockmole & Henderson, in press; Franconeri, Hollingworth, & Simons, 2005; Rensink et al., 1997) the increase in detection rate can be explained either as correspondence being performed or an increase in sensitivity due to attention being focussed on the object. For the purpose of this thesis the distinction is not important. What is important is that, under normal viewing conditions object discontinuities are not perceived.

Instead, the information about object features stored in the object file is updated to accommodate the new information (Hollingworth & Henderson, 2002b). In other words, object continuity is perceived even though the contents of the object file do not correspond with the current object features.

If the evidence of change blindness studies is considered in isolation there seems to be a consensus that viewers are very insensitive to changes in a whole range of object features such as colour, identity, orientation, even presence. However, the majority of change blindness studies use static stimuli (e.g. object arrays, 3D scenes, photographs) in which the only thing changing over the time is the feature the subject is meant to detect. The only studies that use dynamic stimuli are Levin & Simons' film (Levin & Simons, 1997) and real-world experiments (Simons & Levin, 1998) and these, as will soon become apparent, very cleverly control other factors that might encourage change detection. By using only static images, most change blindness studies are limiting the range of information that could be used to check for correspondence across a disruption. Principally, they are excluding spatial and temporal information about the object as it moves through the visual scene. One phenomenon that incorporates precisely these factors is existence constancy.

5.2.4 Existence Constancy

Existence constancy is the name given to the belief that an object continues to exist even when out of sight (Piaget, 1929). As previously discussed (4.1.3.2), viewers will continue to perceive an object in the absence of a visual referent as long as the object disappears from view via one of a limited number of acceptable transformations (Michotte, 1955; Gibson et al., 1969; Kaplan, 1969; Gibson, 1979). One of these transformations occurs when an object moves behind an occluding edge: local, gradual, perspectival transformation (Michotte, 1955; Gibson et al., 1969). Once occluded a representation of the object and its passage through space is maintained. Evidence for this comes from studies showing that the object continues to be tracked by the eyes (Churchland et al., 2003; Johnson, Amso, & Slemmer, 2003) and neuronal activation corresponding to the object is maintained during

occlusion (Assad & Maunsell, 1995; Baker et al., 2001; Olson et al., 2004). This representation has been described as an "abstract conceptual representation of objects" (Baker et al., 2001) much like Kahneman & Treisman's object files (Kahneman & Treisman, 1984).

Existence constancy has been principally investigated by developmental psychologists who have used it to explore the development of object perception over the first few years of life. Unlike adults, infants less than 10 months old do not exhibit the ability to distinguish between objects based on visual features and identity (Bower, 1967; Fantz, 1964; Spelke et al., 1995; Xu & Carey, 1996). Up to 10 months infants show no shock when shown a toy duck that moves behind an occluder only to reappear on the other side as a ball (see right image in Figure 5-2; Bower, 1974; Xu & Carey, 1996). However, they will display shock if the occluder has a gap in the middle and the object does not appear in the gap as it moves from the first occluder to the opposite side of the second (see left image in Figure 5-2; Spelke et al., 1995; Xu & Carey, 1996). In this instance they will indicate they expect two objects to exist behind the occluders rather than one that has jumped across the gap (Spelke et al., 1995; Xu & Carey, 1996).

These findings have lead developmental psychologists to infer that adults store three types of information about an object during occlusion: spatiotemporal, object property, and object kind information (Xu, 1999). Spatiotemporal information refers to the objects motion over time and its adherence to a series of constraints: an object can trace only one continuous path over space and time (**continuity constraint;** Hirsch, 1982), two distinct objects cannot share the same space at the same point in time (**solidity constraint;** Hirsch, 1982), and the motion of the object must continue along the same path at the same speed unless acted upon by external forces (the **smoothness constraint;** Spelke et al., 1995). This spatiotemporal information can be seen as an amalgamation of the information previously identified in change blindness as referenced when perceiving spatial and temporal continuity.

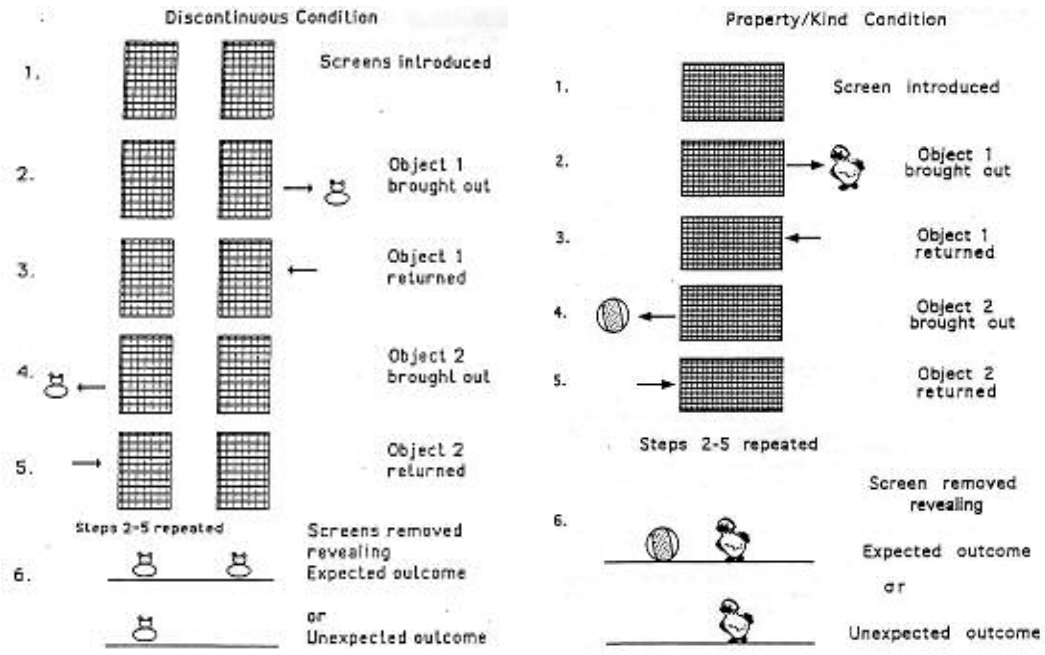


Figure 5-2: Two illustrations of stimuli used in existence constancy experiments (taken from; Xu, 1999). Left image illustrates infants' ability to detect violations of spatiotemporal continuity. Right image illustrates infants (< 10 mths) inability to detect violations of object/property kind information.

Object property information refers to the object's features (e.g. colour, shape, size, or texture) and their acceptable deformations over time (Xu, 1999). **Object kind information** refers to our general knowledge about acceptable behaviour and transformations of an object within a certain object category (Xu, 1999). For example, change in the spatial relationship between sub-parts may be acceptable in an animal (e.g. its limbs) but not for a piece of furniture. Both object property and kind information can be seen as referring to the information used to perceive object continuity.

During the first year of life, infants can use information about object properties to group objects into categories but they show no ability to use it to identify an object after occlusion (Xu, 1999). This ability does not develop until around 12 months at which point the infant infers that even though an object has spatiotemporal continuity, the change in object property information indicates that it is a different object (Spelke et al., 1995). This change has been explained as coinciding with the first stages of language comprehension, the naming of objects believed to be associated with an integration of the spatiotemporal properties of an object with its

identity (Xu, 1999). When adults are asked to perform the object individuation task illustrated in Figure 5-2 they will immediately identify that there must be two objects behind the occluder (Spelke et al., 1995). However, Levin and Simons created similar object-level changes to people during occlusion both in real-world interactions (Simons & Levin, 1998) and film (Levin & Simons, 1997) and adults failed to notice the change. Levin and Simons concluded that this occurred due to a failure to attach a label to the object (Levin & Simons, 2000). Without a label to indicate the change in identity or attention specifically focussed on the task of change detection, adults assumed that spatiotemporal continuity implied object continuity (Levin & Simons, 2000)[98]. Using Kahneman and Treisman's terminology (Kahneman et al., 1992), successful *correspondence* between the object's current spatial and temporal information and that predicted by the object file is sufficient for the object to be perceived as continuing to exist.

5.2.5 Fingers of Instantiation (FINST)

The evidence from both change blindness (Levin & Simons, 1997; Simons & Levin, 1998) and existence constancy studies (Kellman and Spelke, 1983; Spelke et al, 1993; Xu & Carey, 1996) indicates that spatial and temporal information is prioritised over all other object information when identifying an object. Kahneman and Treisman realised this priority when proposing object files: "Perception appears to define objects more by spatiotemporal constraints than by their sensory properties or by their labelled identity" (Kahneman, Treisman, and Gibbs, 1992, pg 177). Object files are first constructed by indexing (i.e. pointing to) an object within the visual scene (Kahneman, Treisman, and Gibbs, 1992). Once this spatial index is established all other featural information can be added (Kahneman, Treisman, and Gibbs, 1992).

[98] Whether the Levin & Simon's film experiments actually satisfied spatial expectations will be discussed in section 5.3.3.

This initial spatial index was initially proposed by Pylyshyn to explain object-based attention (Pylyshyn, 1989). Traditionally attention was seen as a *spotlight* (e.g. Eriksen & Hoffman, 1972; Posner, Snyder, & Davidson, 1980) or zoom-lens (e.g. Eriksen & St. James, 1986) which could be focussed on individual areas of the visual scene. Recently this idea was expanded by the discovery that attention could be allocated to a small number of moving objects distributed over the visual field (see Egeth & Yantis, 1997; for review). Pylyshyn proposed visual indices that could be pre-attentively allocated to objects and tracked over time. He called these indices Fingers of Instantiation (FINSTs) due to the fact that they work in a similar way as fingers: "Even if you do not know anything at all about *what* is located at the places that your fingers are touching, you are still in a position to determine such things as whether the object that finger number 1 is touching is to the left of or above the object that finger number 2 is touching….. [T]he access that the finger contact gives makes it inherently possible to track a *particular* token, that is, to keep referring to what is, in virtue of its historical trace, the *same* object." (Pylyshyn, 1989; page 68). These indices are limited to four[99] at any one time and they are *sticky*: they move with their indexed objects (Scholl & Pylyshyn, 1999). As visual indices are pre-attentive they segment the visual scene into areas of interest (i.e. objects) before attention is focussed on these areas. These indices can then be searched by focal attention without the need for scanning of the entire scene.

Visual indices seem to be the foundation of Kahneman's object files (Kahneman, Treisman, and Gibbs, 1992; (Leslie et al., 1998; Scholl & Leslie, 1999). A visual index points to "shapeless bundles" of visual features that all share spatiotemporal continuity (i.e. they all move together; Wolfe & Bennett, 1997). When focal attention is first allocated to this index the visual features are "bound" together into an object file (Treisman & Gelade, 1980). This object file "sticks" via the visual index to the object as it moves about the visual scene (Leslie et al., 1998).

[99] The same number as object files. This is a sign of the dependency between object files and indices.

Visual indices can also be maintained during occlusion as long as the object disappears from view in the expected way (i.e. local, gradual, perspectival transformation; (Gibson et al., 1969; Flombaum & Scholl, in press; Scholl & Pylyshyn, 1999). If the visual index is lost, such as in the spatiotemporally discontinuous motion during occlusion (see left image Figure 5-2; Spelke et al., 1995; Xu & Carey, 1996) the object file will be abandoned and a new object file allocated to the object when it reappears (Flombaum & Scholl, in press; Scholl, Pylyshyn, & Franconeri, 1999). This means that any featural or identity change to the object occurring during occlusion will not be detected as the object file containing the previous form of the object's information no longer exists (Flombaum & Scholl, in press; Scholl et al., 1999). However, if the object appears from behind the occluder when and where it was expected (i.e. with spatiotemporal continuity) and the viewer is aware of the possibility of a change, viewers will detect a change of feature or identity by performing *correspondence* on the existing object file (Flombaum & Scholl, in press).

Applying this theory of visual indices to ecologically valid (i.e. non-change detection) viewing conditions, object continuity will be assumed after periods of visual disruption (e.g. occlusions, saccades, blinks, etc) as long as spatiotemporal expectations are fulfilled. This was exactly the finding of Levin & Simon's real-world study (Simons & Levin, 1998). In this study the identity and some visual features of an actor changed during a naturally occurring period of occlusion (a door carried between the actor and the viewer; Simons & Levin, 1998). As the actor reappeared in the same position once the door had moved off, fulfilling spatiotemporal expectations, only 50% of viewers noticed the change (Simons & Levin, 1998). This finding seems incredible: a person with whom you are conversing suddenly changes identity and you fail to notice. However, the effect has since been show to be dependent on the size of the change (Williams & Simons, 2000) and its significance to the viewer (Simons & Levin, 1997). If the viewer is familiar with the actor, from a similar age or social group, or inclined to name the actor, their chances of detecting the change will increase (Levin & Simons, 2000; Simons & Levin, 1997). Therefore, whilst visual indices appear to be primarily allocated based on

spatiotemporal information there does seem to be some evidence that expectations about object features can also influence how indices are allocated (Leslie et al., 1998). Whether there is a difference in the perception of existence constancy resulting from these two forms of allocation is still a question waiting to be addressed.

5.2.6 The 'what' and 'where' systems

The combination of object files (Kahneman et al., 1992) with visual indices (FINSTs; Pylyshyn, 1989) is currently just one proposal for how object perception may operate. It is a mid-level theory as it does not deal with how the objects are segmented from the sensory data (see Marr, 1982) or how they are recognised and categorised based on long-term memory (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). It also fails to specify in detail how the "binding" of conceptual object file information to the sensory information of the index is achieved (Treisman & Gelade, 1980) or how this object file is represented in the brain (e.g. McClelland & Rumelhart, 1985). However, there is growing evidence to support the existence of object files and visual indices. Multiple object tracking indicates that attention can be object-based (Egeth & Yantis, 1997; Pylyshyn, 2001) and inhibition-to-return studies have shown that once attention is withdrawn from an object there is a cost to re-indexing the same object (Yi et al., 2003). Transsaccadic memory studies have shown the need for abstract conceptual representations of object properties that can be updated over multiple eye fixations (Henderson, 1994; Henderson & Anes, 1994; Henderson & Hollingworth, 2003; Irwin, 1992). The spatial basis of these representations (i.e. the index) has also been supported by evidence that short-term memory is preferentially accessed by spatial information (Jiang, Olson, & Chun, 2000).

The distinction between spatiotemporal visual information (i.e indices) and featural information (i.e. object files) can also be found in the neuroanatomy of the visual system (Van Essen & Maunsell, 1983). Featural information is processed in circuits joining the visual cortex, through the extrastriate cortex, to the inferior temporal

cortex. This functional pathway is known as the parvocellular pathway or ventral stream. Location and motion information is processed in circuits joining the striate to the parietal cortex. This is known as the magnocellular pathway or dorsal stream. Magno cells respond quicker than parvo cells and in parallel which makes them perfectly suited to maintain multiple spatial indices (Van Essen, Anderson, & Felleman, 1992). This functional and anatomical distinction between the processing of visual spatiotemporal information and featural information has been seen to support the distinction between visual indices and object files (Leslie et al., 1998).

5.2.7 Summary of Object Continuity

Most continuity errors involve a mismatch between viewer's expectations about how, where, and when an object should appear after a cut and its actual appearance. Change blindness studies have confirmed that, even though there are those viewers who will detect continuity errors, in general changes to an object's appearance (identified here as **object** information) go unnoticed (Levin & Simons, 1997; Levin & Simons, 2000). For a viewer to become aware of a change they either need to have their attention captured by the visual transients caused by the change (Rensink et al., 1997) or be able to compare the object features as presented after the cut to a conceptual representation of the expected features. Such a comparison is known as *correspondence* (Ullman, 1979) and the conceptual representation is an *object file* (Kahneman et al., 1992). Object files are thought of as collections of abstract conceptual information about an object's visual features, properties, and identity (Kahneman et al., 1992). These files are "bound" together by focal attention (Treisman & Gelade, 1980) and follow the object through the visual scene by "sticking" to it via a spatiotemporal index (Leslie et al., 1998; Pylyshyn, 1989).

Developmental studies investigating how infants perceive objects as persisting over time (i.e. existence constancy) even in the absence of visual information (e.g. during occlusion) have shown that during the first 10 months infants are unable to use information about an object's identity or visual features to distinguish between objects (Spelke et al., 1995). Instead, they base their expectations on spatiotemporal

information (Xu, 1999). Only once the infant has developed language are they able to retain object information during occlusion (Xu & Carey, 1996). Similar studies performed on adults have shown that in the absence of focal attention (Scholl, Pylyshyn, & Franconeri, 1999) or an explicit change detection task (Simons & Levin, 1997) adults will perceive existence constancy based solely on spatiotemporal continuity even if the object information has changed. These studies indicate that existence constancy is principally dependent on the preservation of a visual index (Pylyshyn, 1989). If spatiotemporal expectations are fulfilled by the object after the visual disruption, the visual index will be allocated to the object and the contents of the associated object file will be updated to match the new object information (Hollingworth & Henderson, 2002b).

The goal of this chapter is to understand how the discontinuous visual information presented across a cut can lead to the perception of "cinematic continuity". During this discussion of continuity errors and object continuity, it has become clear that the "cinematic continuity" referred to in the film literature is actually existence constancy: the continued perception of objects over time. If continuity editing rules are to maintain existence constancy across cuts they need to accommodate viewers' spatial and temporal expectations. These expectations are of critical importance to the maintenance of the visual indices upon which the object's representation is constructed. If spatial and temporal expectations are violated the index will be lost and the object will cease to be perceived. Therefore, to understand how continuity editing rules accommodate these expectations a greater understanding of how space and time are perceived is first required.

5.3 Spatial Continuity

Spatial continuity refers to the continued perception of an object at a particular location when static, or across a series of adjacent locations without any gaps when dynamic. If spatial continuity is expected a violation will result in a spatial continuity error[100]. The most obvious examples of such errors occur during film viewing when a character is suddenly relocated to a new position in the scene or a new location across a cut (an Impossible Relocation error; see section 5.1.1). Similar continuity errors (although they would not typically be identified as such) occur during real-world viewing when an object moving behind an occluder is not seen through a gap in the occluder (Spelke et al., 1995). Spelke et al believed that such a visual event was perceived as an anomaly as it violated the **spatiotemporal continuity constraint:** observers expect an object to traverse a single unbroken path through space (Hirsch, 1982).

This constraint can be clarified in light of recent evidence about visual indices. Viewers maintain an object's visual index during occlusion and this index is projected forwards along the same path (Assad & Maunsell, 1995; Baker et al., 2001; Churchland et al., 2003; Olson et al., 2004; Scholl & Pylyshyn, 1999; Yi et al., 2003). The path projected across the occluder conforms to the Gestalt law of 'good continuation': the object should follow the simplest, most regular, and continuous path under the occluder without any abrupt changes (Koffka, 1935). Eye tracking evidence indicates that the speed at which the object was first occluded is used to update the hypothetical position of the object as it passes behind the occluder even if the object was accelerating before occlusion (Churchland et al., 2003). This direction and speed information is used to update the object's index so that predictions can be made about where and when the object will reappear (Leslie et al., 1998). This ability

[100]The viewer could also expect a spatial discontinuity such as at the end of a scene when all action has ended and the viewer expects the film to move on to a new location. If the film then continues to show the same scene (spatial continuity) a spatial continuity error will be perceived. Editors seem to be aware of the negative effects of violating spatial expectations, referring to the technique as "clearing the frame" (Katz, 1991). See section 4.1.2.2.

to anticipate when an object will appear from behind an occluder is manifest as predictive saccading to the expected location (Johnson et al., 2003).

This evidence of viewers' tendency to anticipate where and when an object will appear from behind an occluder leads to a reinterpretation of Spelke et al's classic occlusion experiment (Spelke et al., 1995). When the object initially moves behind the occluder its index is maintained and updated according the projected motion. Given that there is a gap between the two occluders the first point at which the object is expected to reappear is the gap. When the object doesn't reappear when expected a violation of the **spatiotemporal smoothness constraint** is perceived (*"an object continues along the same path at the same speed unless acted upon by external forces"* Michotte, 1963; but see Spelke et al., 1995 for discussion). The viewer assumes that the object has either slowed down or stopped, either way the object is perceived as continuing to exist behind the first occluder. When an object with the same identity appears on the far side of the other occluder a new visual index has to be allocated to it as the previous index is still pointing to behind the first occluder. When the occluders are then raised to reveal that there is no object behind the first occluder, the viewer exhibits shock, not because they are unable to resolve the motion of the object present with a path that doesn't violate the spatiotemporal continuity constraint (Hirsch, 1982) but because the visual index that was pointing behind the first occluder is now without an object (Leslie et al., 1998).

Three spatiotemporal constraints are believed to be used when predicting object motion: continuity (Hirsch, 1982), smoothness (Michotte, 1963), and solidity (Hirsch, 1982). In light of visual index (Pylyshyn, 1989) and object file theories (Kahneman et al., 1992) these constraints can be reinterpreted as being equivalent to spatial, temporal, and causal expectations, respectively. The spatiotemporal continuity constraint is violated when an object fails to trace a continuous path through space (Hirsch, 1982). Given that visual indices must be constantly maintained during occlusion in order for them to be reallocated to an object when it reappears (Scholl & Pylyshyn, 1999), this constraint expresses the viewer's ability to predict **where** the object should reappear, i.e. formalise spatial expectations. The

spatiotemporal continuity constraint is violated when an object fails to reappear **when** expected (Hirsch, 1982). Therefore, the constraint expresses the viewer's ability to formalise temporal expectations. Finally, Hirsch's (1982) spatiotemporal solidity constraint is violated when two (or more) objects share the same position in space at the same time. If two moving objects are presented as merging into one, only one of the objects' original visual indices will remain (Mitroff, Scholl, & Wynn, under review). If the objects both move in opposite diagonal lines from the bottom of the screen to the top, when they meet in middle their indices will be switched as if the object "bounced" off each other and continued along the opposite diagonal (Mitroff, Scholl, & Wynn, 2005). These effects of merger and bouncing indicate that visual indices adhere to Michotte's perceptual laws of **causality** (Michotte, 1963).

Causality will not be discussed in depth in this thesis as, whilst it is critical to our ability to predict how objects will continue to exist during periods of visual disruption (e.g. occlusions, saccades, cuts), it requires an understanding of stereotypical object behaviours, schemas, as well as the ability to interpret other peoples intentions. All of these factors need to be considered when deciding how to present an object across a cut, especially if time and space are to be omitted. One of the most common styles of editing is question-and-answering. This relies on viewers' ability to predict the outcome of an action (Katz, 1991). Representing cause and effect is critical to this style of editing and is a very powerful tool for driving forward the film's narrative (Bordwell & Thompson, 2001). However, causality is a layer of continuity existing on top of temporal and spatial continuity. Causal expectations can be removed from a cut and temporal and spatial expectations will still remain. If, however, causal expectations exist they appear to be able to modify spatial and temporal expectations (as in the "bouncing" example Mitroff et al., 2005). The scope of this thesis is only large enough to consider temporal and spatial continuity without the influence of causality. This will be a topic worthy of in-depth future investigation.

5.3.1 The basis of spatial expectations

In the existence constancy experiments discussed above (Spelke et al., 1995), spatial information was only modified in a binary fashion: either the object appeared in the gap or it didn't. This level of manipulation does not provide any insight into the precision of the spatial expectation. Does the viewer expect the object to appear in a precise location and experience a spatial continuity error if it deviates even slightly from this location or are they tolerant to a degree of deviation? Temporal expectations, by comparison, have been shown to be very flexible (Nevarez & Scholl, in preparation; Spelke et al., 1995). When a moving object reappears from behind an occluder in the expected location but at a much earlier time than expected, viewers still allocate the same visual index to the object (Nevarez & Scholl, in preparation). Only if the object takes considerably longer than expected to reappear do viewers have trouble maintaining the visual index (Nevarez & Scholl, in preparation). A similar degree of manipulation of spatial discontinuity has not been performed. Whilst it is assumed that dynamic spatial expectations adhere to the Gestalt law of 'good continuation' (Koffka, 1935, see Michotte, 1963) there is no empirical evidence to indicate how deviant an object's path must be from the expected path for a discontinuity to be perceived.

Evidence from change blindness and studies investigating eye movements and recognition memory during film viewing indicate that spatial expectations are not formalised in terms of an object's position within 3D space, all that seems to be important is that the object is located in the right position relative to the viewer. Levin & Simon's study testing change blindness across matched-exit/entrance cuts found that viewers were able to construct a coherent representation of the object's motion even though it underwent a sudden relocation from one side of the screen to the other (Levin & Simons, 1997). If the object's direction of motion is reversed during the cut, causing the object to re-enter the screen from the same edge it departed, viewers' ability to construct a coherent long-term representation of the action deteriorates (Frith & Robson, 1975). These studies suggest that the sudden and impossible restructuring of the 3D space caused by the cut does not disrupt viewers'

ability to maintain the object's index. Eye tracking results (d'Ydewalle et al., 1998) also indicate that viewers saccade directly to the object after such cuts as well as dialogue scenes cut according to the 180° Rule (see Figure 5-3). If the cut violates continuity editing rules (e.g. a cut from 1 to 4 in Figure 5-3), viewers have to perform ocular search of the scene to locate the object (d'Ydewalle et al., 1998).

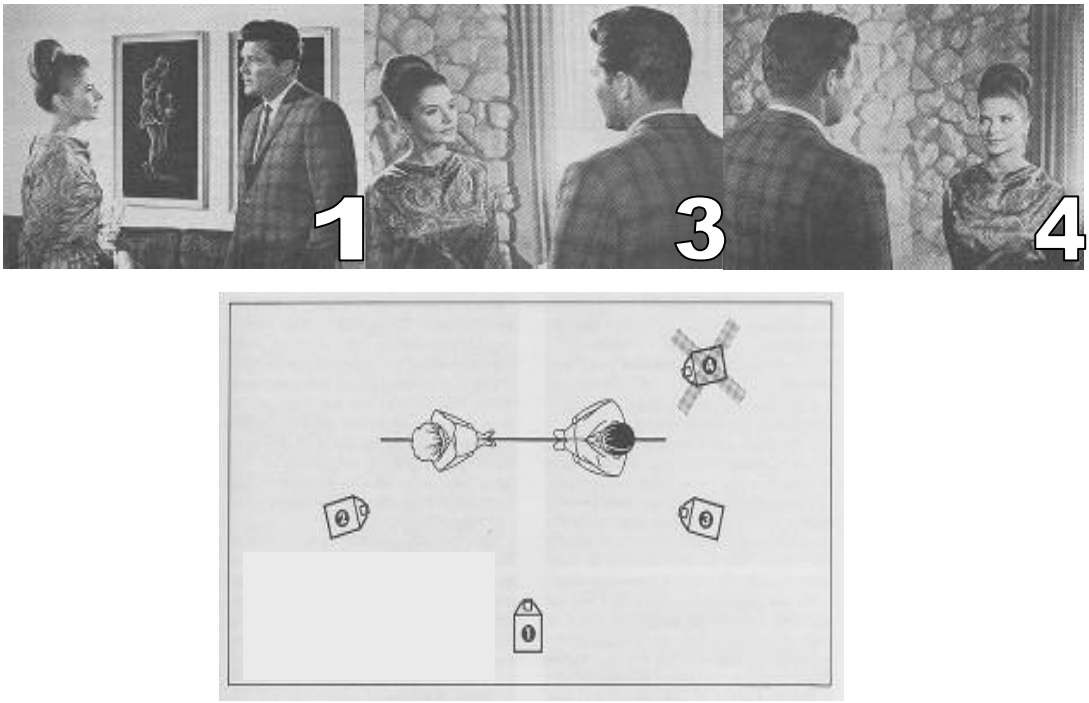


Figure 5-3: The 180° Rule. Once the screen position of each character has been established (camera position and shot 1) these positions have to be preserved. A cut from shot 1 to shot 3 is acceptable as the woman remains on the left of the screen but a cut to shot 4 isn't acceptable as the woman shifts to the right.

These film studies seem to suggest that viewers use egocentric spatial information to formalise expectations about where an object will be located after a visual disruption (a cut or saccade). An object is expected to be located at a certain position within the viewer's visual field not within the 3D space of the scene. If a cut occurs whilst the viewer is performing a saccade to an object, viewers will only perceive a spatial continuity error if the object is no longer present where the eyes land (d'Ydewalle et al., 1998). Viewers will still perceive spatial continuity even if the change in camera position has changed the saccade target's spatial relationship to the rest of the 3D scene. Evidence for this was provided by a follow on study mentioned by d'Ydewalle et al (1998). In this study they switched both the position of two actors

within a scene and the side from which they were filmed. This maintains the actors' positions within the viewer's visual field but changes their location within the 3D space of the scene (the same technique was used to construct shot 3 in Figure 5-3; notice the change in background from shot 1). When viewers are presented this new shot they behave as if nothing has changed, saccading directly to the main actor and when later probed they show no sign of noticing the changed background (d'Ydewalle et al., 1998). This technique of preserving a character's position on the screen whilst changing their position within the 3D scene or even their geographical location is a long-standing and integral part of continuity editing (Bordwell et al., 1985)[101].

Intuitively it might seem hard to believe that the visual scene could suddenly rotate around the saccade target during a saccade without your becoming aware of the rotation but evidence for such a blindness does exist. When viewers perform a saccade to an object in the visual scene they frequently fail to notice when the object is rotated during the saccade (Henderson & Hollingworth, 1999). The probability that they detect the rotation increases as the amplitude of the saccade decreases but even when the saccade only covers 1° of a visual angle they still only detect the rotation 40% of the time (Henderson & Hollingworth, 1999). By comparison, deletions of the saccade target are detected 80% of the time with similar saccade amplitudes (Henderson & Hollingworth, 1999).

In a related study, viewers were shown to be highly insensitive to incremental rotations of the entire visual scene (Hollingworth & Henderson, 2002b). When a visual scene was incrementally rotated 1° away from the original viewpoint during

[101] The most extreme example of sacrificing spatial continuity within the 3D scene for screen-relative spatial continuity can be seen in Cheat Cuts. These refer to cuts where an actor or object has been moved within the 3D scene between shots to preserve their position on the screen. This sometimes occurs because an actor/actress is shorter in real-life than they wish to appear on the screen or when the director wishes to place the camera in a position where a false wall used to be. Cheat cuts usually go unnoticed unless the viewer is very astute and inclined to dwell on the logical construction of a scene.

periods of masking, viewers required an average total rotation of 31° before they would explicitly report the rotation (Hollingworth & Henderson, 2002b). However, when the scene rotates more than 20° during one period of masking the majority of viewers noticed the rotation (Hollingworth & Henderson, 2002b). This indicates that viewers automatically update their representation of the scenes spatial relationships (i.e. their visual indices) to accommodate small deviations without becoming aware of the deviation (Hollingworth & Henderson, 2002b). Only when the size of the deviation becomes too large does explicit change detection occur. This can be attributed to a failure of correspondence between the new visual information and that stored in the visual indices and object files.

5.3.2 Visual Stability

This tolerance for small discrepancies between received spatial information and that associated with the object file is necessary as the extraretinal signals generated by eye movements are inaccurate (Grüsser, Krizic, & Weis, 1987). These signals represent the size and direction of eye movements so they could be used to cancel out the relocation of visual objects on the retina caused by the eye movements (Helmholtz, 1867/1925). Unfortunately, the extraretinal signals are not identical to the size of the eye movements (Grüsser et al., 1987). The inaccuracy of extraretinal signals combined with shift of covert attention towards the saccade target immediately prior to a saccade (Duhamel, Colby, & Goldberg, 1992) causes large distortions of perceived space (see Ross et al., 2001 for a review). To compensate for these distortions, spatial information present in the visual scene after the saccade are believed to be used to reconstruct perceived space (Deubel & Schneider, 1994; Henderson & Hollingworth, 2003; McConkie & Currie, 1996).

If an object is perceived as remaining in the same position across an eye movement it is said to be “visually stable”. In the absence of reliable extraretinal information about the eye movement other information is required to check stability. Three theories have been proposed that attempt to explain visual stability: *reference object theory* (Deubel & Schneider, 1994), *saccade target theory* (McConkie & Currie,

1996), and the *visual memory theory* of dynamic scene representation (Henderson & Hollingworth, 2003; Hollingworth & Henderson, 2002a). All three identify the saccade target as being of primary importance when perceiving visual stability of the entire visual scene across a saccade. The processing of visual stability across a saccade proceeds by first focussing covert attention on the saccade target (Duhamel et al., 1992). The peripheral sensory information available about the saccade target is used to form an abstract visual representation (equivalent to an object file and index; Henderson & Hollingworth, 2003). As the eyes move, the sensory information quickly decays (Sperling, 1960) leaving only the abstract representation. When the eyes land correspondence is performed between the stored information and that encoded from the current fixation. If there is a mismatch an error signal is generated which triggers either covert or overt change detection (Henderson & Hollingworth, 2003). If correspondence succeeds the new information is integrated into the existing representation, irrespective of whether there is an absolute match.

Recent evidence indicates that this level of spatial information might be insufficient as "landmark" objects are sometimes needed to locate the saccade target (Deubel, 2004). The spatial information stored across the saccade extends to include a small number of adjacent objects. If the saccade target is not immediately located after the saccade these "landmarks" will be used to first locate the target and then, by comparing its relative position to the landmarks, detect if the saccade target moved during the saccade (Deubel, 2004). These "landmarks" must be indexible objects close to the saccade target, general background information cannot be used (Deubel, 2004). A similar allocentric[102] representation of the saccade target's position has been proposed under the *visual memory theory* of dynamic scene representation (Henderson & Hollingworth, 2003).

The tolerance for small discrepancies between the saccade target's location and that predicted prior to the eye movement extends a few degrees from the landing position of the eyes after the saccade (Deubel, 2004). This range of possible locations has

[102] Relative to other objects.

been called a 'constancy window' (Currie et al., 2000). The viewer has 30ms to locate the saccade target within this window before visual instability is perceived and the extra-retinal eye movement signal is used to reconstruct the spatial representation of the scene (Deubel, 2004).

5.3.3 Perceiving spatial continuity across film cuts

This 'constancy window' is evidence that small discrepancies between where we expect an object to be located and where it appears after a saccade are attributed to errors of planning and performing eye movements, not an instability of the saccade target (Deubel, 2004) or the visual scene (Hollingworth & Henderson, 2002a). In the absence of any clues to the contrary, spatial continuity is assumed (Dennett, 1991). To ensure spatial continuity the visual index targeted in the periphery prior to the saccade must be located within 30ms and a few degrees of the fixation position. As a visual index is a limited representation of an object's position in space (Pylyshyn, 1989) and does not contain information about the object's orientation, size or shape it is insensitive to rotations (Henderson & Hollingworth, 1999; Hollingworth & Henderson, 2002b), enlargements (McConkie & Currie, 1996), changes in visual features (Henderson & Hollingworth, 2003) or identity (Levin & Simons, 2000). Therefore, the spatial transformation of a visual scene caused by cut is compatible with the perception of spatial continuity as long as:

a) the transformation occurs during a visual disruption such as a saccade,
b) the saccade target can be targeted prior to the cut, and
c) the saccade target is located in roughly the same position after the cut.

If these conditions are fulfilled and the viewer is not actively performing a change detection task, there is no reason, given the existing theories of visual stability, why the visual scene should not be perceived as spatially continuous across the cut.

The first condition outlined above, “the transformation [i.e. cut] occurs during a visual disruption such as a saccade”, is essential as otherwise the visual disruption caused by the cut itself should capture attention and make the viewer aware of the scenes visual instability. However, the coinciding of a saccade with a cut has already been shown to occur across a range of different types of edits (see chapters 3 and 4). In chapter 3 eye tracking evidence was used to show how viewers had a tendency to coincide their saccades with cuts during conversations, cuts to Point of View (POV) shots, Over-the-shoulder (OTS) shots, and other shots containing objects predictable from the first shot (May et al., 2003).This evidence was supported by the results of the first experiment of this thesis which showed that viewers would time a saccade to coincide with the point at which a moving object was fully occluded by the screen edge, saccading to the opposite screen edge to continue tracking the object (i.e. during a matched-exit/entrance cut; 2.1.4). These results indicate that, whilst not occurring during every cut, there does seem to be an indication that viewers are able to predict when some cuts will occur, choose a saccade target prior to the cut, and not notice a significant change to this target after the cut (see Schröder, 1990 for evidence of a lack of awareness of continuity cuts) .

The second condition required to perceive spatial continuity across a cut is that “the saccade target can be targeted prior to the cut”. In real-world vision this isn’t a problem as the visual scene is relatively static on the viewer’s retina except for a few moving objects and the motion caused by body and head movements which are automatically compensated for. Saccade targets can be identified by identifying a visual object in the periphery and it is highly unlikely, given the short duration of a saccade (~150-200ms) that the saccade target will have moved after the saccade. In film viewing, however, there is no guarantee that an object chosen as a saccade target prior to a cut will remain in the same position after the cut. A camera can physically move to any vantage point within a scene creating an almost infinite number of variations of shot.

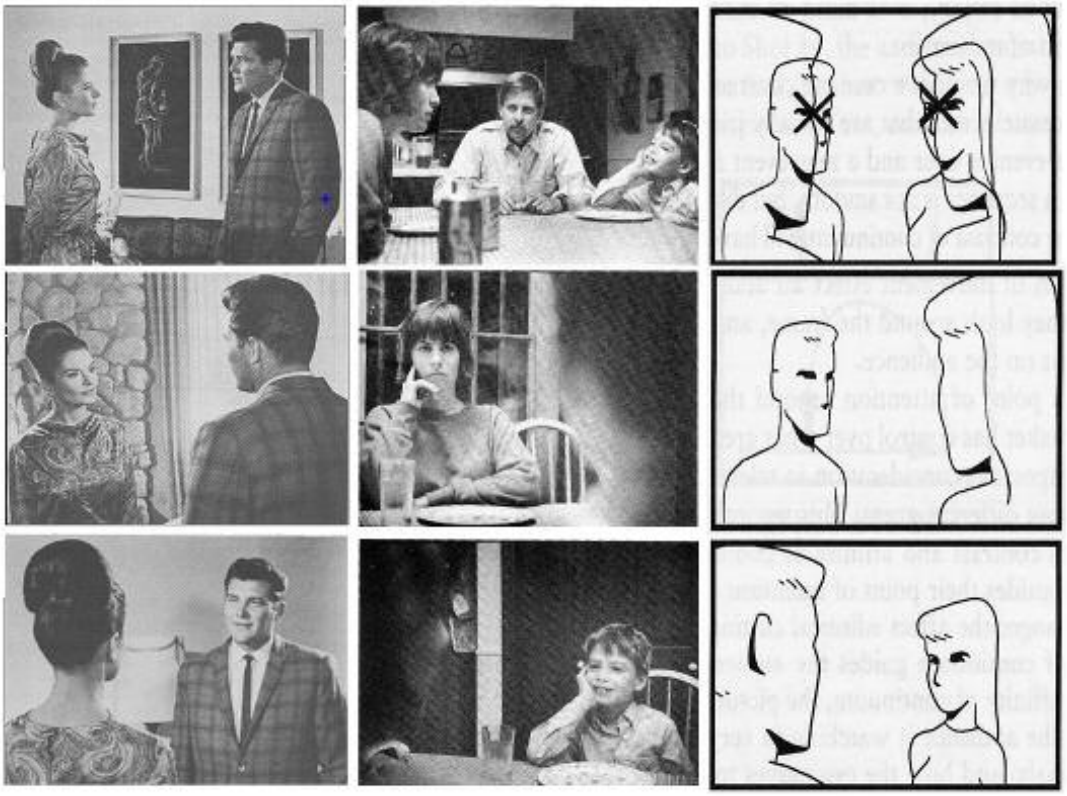


Figure 5-4: Ensuring perceived spatial continuity by adhering to 180° Rule. Each establishing shot (top line) establishes a character's position on the screen. This is then preserved across all subsequent shots (see how each character stays within their own spatial region). Right column taken from Block (2001). Top image in right column shows a rough superimposition of the two bottom images.

However, full use of this range of camera positions is rarely seen, principally because it is not permitted according to continuity editing rules. Figure 5-4 illustrates three different shot sequences (running vertically) created according to the 180° Rule. Each sequence begins with an establishing shot[103] which introduces the viewer to the characters involved in the scene, their location within the 3D space of the scene, and, of principle importance for the perception of spatial continuity, their location on the screen. In the left sequence of Figure 5-4, the woman is established on the left of the screen and the man on the right. Across all subsequent shots they always remain within their own small region of the screen: woman on the left, man on the right. A

[103] Except for the right column. The top image is a composite of the bottom two images (taken from Block, 2001). It is intended to show the close relationship between the focal points of interest in the two OTS shots.

continuity error would be seen as occurring (according to editing theory) if the wrong character was ever to enter into another character's screen space (e.g. Block, 2001).

This emphasis on preserving screen location can be explained in terms of saccade targeting. If, during the establishing shot the viewer is fixating the man, when the woman begins talking they will want to saccade to her (see Chapter 3 for evidence of this consensual movement of gaze). The woman is targeted by separating the part of peripheral scene relating to her head from the background information (the saccade will probably go to her eyes or mouth; Yarbus, 1967). If a cut to the second shot then occurs during the saccade to the woman, when the viewer regains perceptual acuity after the saccade they will find the woman's head roughly in the same position it was when they targeted it. This satisfies the third condition for perceiving spatial continuity across a cut: "the saccade target [must be] located in roughly the same position after the cut". Her head is now rotated and may have changed size slightly due to the change in camera position but, as long as the visual index created prior to the saccade can be attributed to her head the viewer should assume visual stability. The cut has completely transformed the visual scene (see the change to the man) yet the viewer is unaware of any spatial discontinuity.

The 180° Rule has always been assumed to guarantee the viewer's clear understanding of the scene's 3D space (e.g. (Bordwell et al., 1985; Katz, 1991; Reisz & Millar, 1953) but the precise mechanism by which the rule achieved this has never been explained. Bruce Block (2001) is the only film theorist to indicate that the viewer's point of fixation in combination with their eye movements is important for creating "smooth" and "invisible" editing. Block understood that the location of an object during one shot should be matched after a cut so that a viewer saccading to or fixating that object would not have to perform another saccade to relocate the object after the cut (Block, 2001). He illustrated this by superimposing two Over-The-Shoulder shots filmed according to the 180° rule (top right, Figure 5-4). As in the other shot sequences in Figure 5-4, each character stays in their own region of the screen across a cut and an object targeted for a saccade prior to a cut (such as the man's eyes) would still be located in exactly the same position after the cut. This

ensures that the viewer is not performing saccades to compensate for the editor's manipulations, just to follow their own perceptual enquiry e.g. "Who is speaking now?"

Block believed that the need for compensatory eye movements after a cut increases the chances that the viewer will become aware of the editing and produces a more abrupt, visually jarring experience (Block, 2001). This can now be explained as being due to the viewer's inability to locate the saccade target within the "constancy window" after the saccade. The visual index associated with the object prior to the saccade is lost as the object has moved too far from the original position. Without a visual index the object file cannot continue to exist. This causes the old file to be erased and a new visual index and object file to be allocated to the object. As there is no representational connection between the old object file and the new file the object cannot be perceived as continuing to exist. Conceptually, the viewer will be able to associate the same identity to the object as before the cut but without the spatiotemporal connection (i.e. the visual index) no continuity will be perceived.

This theory of how spatial continuity is perceived across cuts is not universal; the sequence of shots illustrated in Figure 5-4 is only a very small subset of all the types of shot sequences possible in film. Not all shots present all characters/objects from a scene on the screen at the same. If an object is not present in a shot it can not be chosen as a saccade target[104]. However, dialogue sequences (such as in Figure 5-4) filmed and edited according to the 180° rule have been shown to constitute 31% of all cuts making it the most frequent type of cut (May et al., 2003). If our sensitivity to visual stability across saccades is as flexible as implied by the existing evidence (Currie et al., 2000; Deubel, 2004; Hollingworth & Henderson, 2002b) and maintenance of the visual index is all that is required for the perception of existence constancy under non-change detection viewing conditions (i.e. correspondence is not

[104] For example, how would the man present in the establishing shot of the centre sequence in Figure 5-4 be targeted if he were to speak during either the middle or bottom shot? Could he be targeted based on the viewer's memory for his position within the scene or would the saccade have to wait until the cut had occurred and he had appeared on the screen?

performed on the object file; Hollingworth & Henderson, 2002b) then viewers should perceive existence constancy when viewing film sequences edited according to the 180° Rule. Further empirical evidence is required to clarify this theory. Specifically, evidence is needed to check if "true" visual stability[105] is perceived across a cut or whether some degree of implicit change detection compensates for the instability at some conceptual level.

5.3.4 The screen edge as saccade target

One type of cut not compatible with this theory of visual stability is the matched-exit/entrance cut used in the first experiment (see Figure 5-5). This cut depicts an object moving off one side of the screen before a cut and then moving back on to the opposite side of the screen after the cut (see Figure 5-5, next page). The saccade that must be timed to coincide with the cut (see chapter 4) needs to originate from the same object it will end on. This object cannot be the saccade target as it is not in the same position before and after the cut. Therefore, given the theory of visual stability outlined in the previous section, the object cannot be perceived as having spatial continuity. Some other explanation is required.

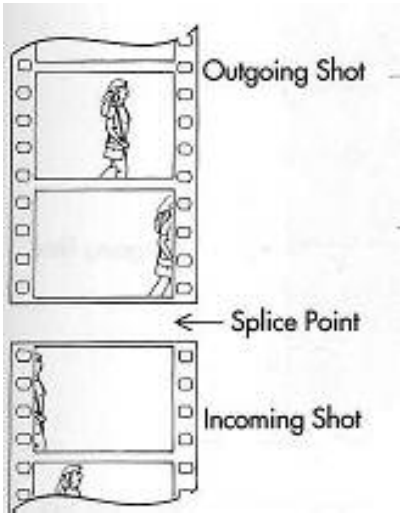


Figure 5-5: Matched-exit/entrance cut (taken from Katz, 1991).

[105] What constitutes "true" visual stability is unclear but it is assumed that the continued association of a visual index to an object across a visual disruption forms the foundation of perceived visual stability.

The action depicted in a matched-exit/entrance cut is similar to the occlusions used in the classic existence constancy experiments (Spelke et al., 1995). An object is tracked as it moves behind an occluder and then reappears at the opposite edge of the occluder moving in the same direction and at the same speed. The key difference is that the occluding edge under which it disappears is not connected to the edge from which it will appear by "a single unbroken path through space" (Hirsch, 1982). The path of the object is spatiotemporally discontinuous. However, Hirsch's constraint has previously been reinterpreted based on evidence that viewers tend to saccade to the occluder's edge in anticipation of the object's reappearance (Johnson et al., 2003). All that is essential for the perception of spatial continuity is that the object reappears from behind the occluding edge in the expected position (see section 5.3).

Under real-world occlusion conditions, the viewer can predict where the object will reappear by extrapolating from the object's direction of motion prior to occlusion. When watching film, no such extrapolation can occur as the motion has to be relocated to the opposite screen edge. However, viewers must be able to predict where the object will re-appear after a cut as there is evidence that spatial continuity is perceived across a matched-exit/entrance cut (Williams & Simons, 2000). Williams & Simons showed that viewers' ability to detect featural changes to an object across a matched-exit/entrance cut is not significantly worse than during spatiotemporally continuous occlusion (Williams & Simons, 2000). This indicates that viewers must be maintaining the visual index across such a cut as otherwise the object file required for them to perform the comparison task would not continue to exist (Flombaum & Scholl, in press).

If all that is required for the perception of spatial continuity is that the object appears from behind the occluder at the expected location this constraint can be fulfilled by a matched-exit/entrance cut. All that is required is that the viewer identifies the opposite screen edge as the occluding edge from which the object will reappear. Evidence for such an expectation can be seen in memory experiments where matched-exit-entrance cuts result in intact memory for the object's motion where as cuts that depict the object as re-entering the screen from the same edge it departed

are recalled poorly (Frith & Robson, 1975). This suggest that viewers do identify the opposite screen edge as the point at which the object will appear. The fact that the saccade required to reach this edge is in the opposite direction to the object's motion should not hinder the perception of spatial continuity as the extraretinal signals indicating the direction of eye movement will only be referenced if the saccade target cannot be located (Deubel & Schneider, 1994; McConkie & Currie, 1996; Hollingworth & Henderson, 2002a). Locating the saccade target shouldn't be the problem as the screen edge is a large, reliably static, and highly salient part of the visual field. Once the screen edge has been located spatial stability should be assumed. All that is then required for the object's motion to be perceived as spatiotemporally continuous is that the occluded object reappears at the expected time. This temporal expectation will be discussed in depth in the Temporal Continuity section (see section 5.4).

5.3.5 Summary of Spatial Continuity

In summary, an object's spatial information is retained across saccades and projected across periods of occlusion (Leslie et al., 1998). This information is used to detect deviations from the object's expected passage through space, however, it is imprecise and based only on the spatial relationships between the target object and a small number of landmark objects (Deubel, 2004). Only if the saccade target and adjacent landmarks cannot be found immediately after the saccade within a few degrees of the eye's landing position will information about the observer's eye movement be used to reconstruct their impression of the 3D space (Deubel & Schneider, 1994; McConkie & Currie, 1996; Hollingworth & Henderson, 2002a).

This dependency on the saccade target for the perception of visual stability across saccades provides a mechanism by which the instantaneous change in viewpoint caused by a cut can be perceived as spatially continuous. The 180° Rule ensures that objects remain in roughly the same position across all cuts (Block, 2001). This ensures that when a saccade coincides with a cut (as found by the first experiment of this thesis and (May et al., 2003) the saccade target will be located immediately after

the saccade. This ensures that the visual index remains allocated to the object and that it is perceived as spatially stable (Leslie et al., 1998). Visual stability of the saccade target (and possibly a few spatially adjacent landmarks such as the screen edge) is sufficient for the scene to be perceived as spatially continuous across the saccade (Deubel & Schneider, 1994; McConkie & Currie, 1996; Hollingworth & Henderson, 2002a).

When the film depicts an object moving out of shot (i.e. a matched-exit/entrance cut) viewers' attention is *pushed* back across the screen to the opposite screen edge (see chapter 4). Saccadic eye movements are initiated before the object re-appears suggesting that the opposite screen edge is chosen as the saccade target. As the screen edge is constant across the cut the saccade target will not violate spatial expectations. The reflexive attentional shift may also result in spatiotemporal expectations about the object's motion also shifting to the opposite screen edge (see chapter 4 and Frith & Robson, 1975). The reappearance of the object at the opposite screen edge could then satisfy expectations and lead to the perception of existence constancy.

Matching temporal expectations across a cut is not as straight forward as in the real-world as it is not known if the cut causes any distortions in time perception. Spelke et al (1995) found that infants' expectations about when an object should reappear from behind an occluder were inaccurate. Nevarez & Scholl (Nevarez & Scholl, in preparation) supported this view with evidence that adults were very insensitive to objects that suddenly sped up during occlusion (causing them to reappear much earlier than expected) but were unable to continue tracking an object when it took longer to appear than expected. These studies investigated temporal expectations during occlusion, the film equivalent of which has already been shown to lead to unclear expectations about when an object should re-enter the screen (see chapter 4), but what is not known is how temporal expectations are formalised across more common types of cut such as those adhering to the 180° Rule. Investigating how temporal expectations are formalised and temporal continuity perceived across a cut will be investigated in the final empirical study of this thesis. However, before

formalising hypotheses about the perception of time during film viewing an insight must be gained about time perception in the real-world.

5.4 Temporal Continuity

"Time is a mental construction" (Pöppel, 1997). Our experience of time is independent from any external veridical referent and not the result of a dedicated sensory system (Zakay & Block, 1996). This seems very odd given the critical role accurate time perception plays in everyday life (Michon, 1985). If, for example, we perceived the time taken to process a sensory event as less than the actual time, any response we made to that event would occur too late. Our perceptual experience would be "out-of-synch" with the world. This example seems fantastical but it is not far from the truth. Our subjective experience of temporal duration is subject to large variability due to factors such as body temperature (Campbell & Birnbaum, 1994), pleasure (Carpenter & Wojtaszcyk, 2002), arousal (Penton-Voak, Edwards, Percival, & Wearden, 1996), memory (Block, 1978), and attention (Zakay & Block, 1996). The distortions caused by these factors are most obvious for long durations e.g. we experience a holiday as passing much quicker than the same duration spent at our tedious job, but they occur every time we process sensory information. The reason why we are not always aware of the distortions is that we are not always confronted by a conflicting temporal referent immediately after the processing period. Without an indication that our perception of time is erroneous we continue interacting with the world as normal.

These distortions arise due to the way the brain measures duration. Many theories of time perception have been proposed (e.g. (Block, 1990; Pöppel, 1997; Treisman, 1963) but all share similar concepts. Time perception is seen as being comprised of many separate yet hierarchically dependent phenomena such as simultaneity, successiveness, temporal order, subjective present, temporal continuity and duration (Pöppel, 1997). The common belief is that sensory data is processed in a continuous fashion, creating a continuous internal representation of the outside world (see Watson, 1986 for review). However, due to differing rates of processing across

sensory modalities (the auditory system has the best whilst the visual system has the worst temporal acuity) there exists a minimum period of time necessary for two sensory stimuli to be represented as successive. Subjects will perceive successively presented stimuli as simultaneous unless there is 30ms between them (Hirsch & Sherrick, 1961; von Steinbüchel, Wittmann, & Pöppel, 1996). This 30ms limit functions as a discrete unit of time from which all other phenomena of time perception can be constructed.

The origin of this 30ms time unit is the source of most debate in the area of time perception. Models of time perception can be categorised into two major groups: with-a-timer and without-a-timer models (Ivry & Hazeltine, 1992). With-a-timer models propose some sort of "pacemaker" (Zakay & Block, 1996), internal "clock" (Treisman, 1963), or neuronal metronome that meter out these 30ms "beats" (Coull, Vidal, Nazarian, & Macar, 2004). Without-a-timer models propose that these units are defined by changes between periods of sensory processing and storage (Block, 1990). Recent neuroanatomical evidence seems to support the with-a-timer models by identifying a network of brain regions involved in time processing, the supplementary motor area (SMA), premotor cortex, frontal operculum, basal ganglia and cerebellum (Macar et al., 2002), with the anterior portion of the SMA being responsible for the production of these time "beats" (Coull et al., 2004). Whilst there is not believed to be a single "neuronal clock", there does seem to be evidence for a primitive unit of subjective time based on a rhythmical pattern of neuronal activation (referred to as "neuronal oscillations"; Pöppel, 1970).

Treisman (1963) proposed that the timing mechanism consisted of a pacemaker, a switch, and an accumulator. The pacemaker generates semi-regular "pulses". Based on recent evidence these can be equated with the 30ms neuronal oscillations. The pacemaker works continuously, not only when time is being perceived (as proposed by Pöppel, 1997), but the "pulses" are only used once an external timing signal is perceived. Once the beginning of a time period is perceived the switch allows "pulses" to pass into the accumulator. This counts the total number of "pulses" passing through the switch during an event and associates the count with the event

when stored in long-term memory. Similar representations of pulse counts can be compared in working memory to formalise duration judgements.

As well as producing a measure of time, this accumulation of time units also allows the perception of duration and the "subjective moment". Whilst 30ms is the minimal period of time needed to resolve sensory information there also seems to exist a maximum duration within which sensory stimuli can be perceptually grouped into a single event. Short-term memory has been shown to hold information for only ~3s without rehearsal (Peterson & Peterson, 1959). The temporal segmentation of speech or a string of auditory stimuli occurs in 3s intervals (Szelag, von Steinbüchel, Reiser, de Langen, & Pöppel, 1996; Vollrath, Kazenwadel, & Krüger, 1992). The spontaneous change in the perception of ambiguous figures (e.g. Necker cubes) occurs every 3s as attentional mechanisms are elicited that seek out new information (von Steinbüchel et al., 1996). All these results highlight the significance of 3s as a primitive unit of time and its functional approximation with the concept of "subjective presence" i.e. a sense of "now-ness" distinct from the past or the future (Pöppel, 1997).

As this "subjective presence" is the largest period of time within which sensory events can be perceived as a unit, events lasting longer than 3s require another mechanism to enable us to perceive them as continuous: temporal continuity. An object or event is perceived as occurring continuously over time unless conflicting evidence is perceived e.g. the end of the event or its transformation into another event. As the real-world is believed to be temporally continuous it seems a fair assumption to that our perceptual systems would be adapted to perceiving temporal continuity. However, if our experience of time is constructed from discrete 3s units then the perception of events longer than 3s require the active integration of multiple time units. At the end of each 3s unit lies a potential discontinuity. Pöppel suggests that in order for us to perceive an event as having temporal continuity, these discontinuities need to be masked by the successful semantic integration of information from the adjacent time units (Pöppel, 1997). If this semantic integration fails attention will be directed towards new information and a temporal discontinuity

will be perceived (Pöppel, 1997). This semantic integration can be interpreted as analogous to Kahneman & Treisman's (1992) *correspondence* performed between the new information and that of the object file associated with the current event.

5.4.1 Subjective Duration

The culmination of this temporal processing system is the ability to perceive duration. 30ms time units accumulate during the course of sensory stimulation and are integrated into 3s "subjective moments". Sequences of these are then bound together by the similarity of their semantic content into a representation of the sensory event over time. If attention is then focussed on the temporal information associated with this event the number of primitive time units accumulated during the event can be used to identify its duration. Given that this is constructed from 30ms neuronal oscillations it could be assumed that our perception of subjective durations was quite accurate. However, a wealth of evidence[106] along with personal experience tells us that this isn't the case. Numerous factors have been shown to effect our perception of durations. If a subject is asked to judge a duration after having experienced it (retrospective duration estimation) their estimates will typically be shorter than when they are told of the duration estimation task before the event (prospective duration estimation; see Block and Zakay, 1994 for review). If the subject is instructed to perform some non-temporal task during an event, retrospective estimates of the events duration will be longer (Coren, Ward, & Enns, 1993; Ornstein, 1969) whilst prospective estimates will be shorter than the same event without a secondary task (Zakay, Nitzan, & Glicksohn, 1983b). However, the effect depends on the complexity of the task. If the task is complex and difficult to perform it will occupy cognitive resources leading to shorter prospective durations (Zakay, 1993). However, if the task is simple it will not require so many resources and prospective durations will get longer (Zakay, 1993). By comparison, increasing task or information complexity leads to increasing retrospective durations (Ornstein,

[106] Brown (1997) reviewed over 80 individual experiments performed between 1924 and 1995 all examining the effects of task demands on time judgement.

1969; Poynter, 1989; Boltz, 1995). This "interference" (Brown, 1997) or "filled-duration" illusion (Fraisse, 1963; Zakay & Block, 1996) is one of the most consistent findings in the time perception literature and indicates the importance of attention for duration perception.

5.4.2 Prospective vs. Retrospective

A similar effect, and one which emphasises the difference between prospective and retrospective duration estimation, is commonly referred to as the "watched pot" illusion. When attention is focussed on an event, the prospective experience of the duration is longer than when attention is distracted (Block, George, & Reed, 1980; Zakay, 1992) or divided between stimuli (Zakay, 1989; Zakay & Block, 1996). This effect gets its name from the old adage "a watched pot never boils" and was initially investigated by testing precisely that (Block et al., 1980). This effect of concentrated attention also occurs (to a lesser degree) when the duration estimation task is performed retrospectively (Block et al., 1980). However, attention, in itself does not seem to be as important to retrospective time perception as prospective. This can be explained in terms of their primitive temporal units.

If the subject isn't aware of the time estimation task until afterwards (i.e. retrospective estimation) they have no accumulated record of time to reference. Instead they must base their temporal judgements on what they remember about the event. However, not all memories are useful for duration estimation as they are devoid of any temporal information. What seems to be used is the number, magnitude, and salience of significant changes within an event (Poynter, 1989). The amount of "memory space" required by an event is a good place to start constructing a perceived duration (Ornstein, 1969) as this will increase as the complexity of the event increases. However, size alone does not seem sufficient to explain subjective retrospective duration. Instead what seems to be most important is the number and saliency of changes in stimuli during an event (Block, 1982; Block, 1990; Fraisse,

1963; Poynter, 1989). Poynter believes that change is important for time perception as "it is the psychological index of time passage" (Poynter, 1989; page 309).

An event is constructed from sub-events all of which have their own smaller durations. Our expectations about these events, their regularity, order, and stereotypical duration allows us to base our duration judgements of the overall event on their summation (Brown, 1995). This means that our ability to retrospectively perceive duration is based upon our ability to accurately segment events which, as we have already seen (chapter 3) relies on prior knowledge of the hierarchical structure of events (Jones & Boltz, 1989) as well as low-level visual features such as changes in motion (Brown, 1995; Zacks, 2004). The more event boundaries we perceive in a period the longer its duration will appear to be when judged retrospectively (Block, 1978; Block, 1982; Block, 1990; Block & Reed, 1978; Zakay, Tsal, Moses, & Shahar, 1994). There also seems to be an interaction between event segmentation and difficulty in extracting event boundaries. Boltz (1995) showed that subjects were quite accurate at retrospectively identifying the duration of an event when its presentation accentuated the event boundaries[107]. However, when the presentation accentuated points between event boundaries, reproduced durations were longer with more disagreement across subjects.

Unlike retrospective duration estimation, prospective estimations can be based on an explicit measure of time: the 30ms neuronal oscillations[108]. When subjects are informed before the event of the time estimation task they can flick the "switch" (Treisman, 1963) that allows "pulses" from the pacemaker to be accumulated. These pulses are then used to label the event with temporal information. The main problem with this model is that its components are seen as rigid and systematic which leaves

---

[107] Boltz (1995) used musical compositions and filmed narratives as her events. Accentuation was provided by prolonged notes or commercial breaks (as if the filmed narrative was presented on TV).

[108] Subjects may also use other techniques for judging time prospectively e.g. sub-vocal counting, physical tapping, blinking, etc. These could be seen as confounds to "pure" time perception but as they are all just rhythmical acts generated by the same neuronal timer they can be just interpreted as external manifestations of the timer.

no room for the resulting distortions of duration perception so often observed (e.g. filled-duration illusion and watched pot illusion).

To accommodate this apparent flexibility in the timing model attention must be incorporated. Zakay and Block (1996) proposed an attentional-gate model based on Treisman's model. They extended Treisman's model by proposing an attentional gate between the pacemaker and the switch. This gate increases in size as more attention is allocated to it (or in practice, allocated to the task of time estimation). The pacemaker constantly produces "pulses" at a rate influenced by arousal (both due to circadian rhythm and induced by the stimulus). When the subject perceives the beginning of an event which they wish to time the switch opens. The rate of flow of "pluses" from the pacemaker through the switch is then metered by the attentional gate. The greater the degree of attention focussed on the time estimation task, the more "pulses" are allowed to pass through the gate. These pulses are then counted by a "cognitive counter" (Zakay and Block's equivalent to Treisman's "accumulator") from where the count can be referenced as the events duration or compared to similar counts stored in long-term memory.

Whilst the attentional gate may seem to have taken the role of Treisman's original "switch" the "switch" in Zakay and Block's model serves a different function. The switch identifies the beginning and end of events so that the "pulses" accumulated between these two points can be associated with the event as its duration. This requires the processing of event structures similar to those seen to effect retrospective durations. The effect of increasing the degree of event segmentation (i.e. "switching" on and off) has been shown to be less pronounced for prospective compared to retrospective estimates when using static stimuli (Zakay et al., 1994) but just as strong when the events boundaries are created by changes in stimulus motion (Brown, 1995). This effect of segmentation cannot just be attributed to memory as it can in retrospective conditions as event boundaries are also associated with changes in cognitive load and attention which may modify the attentional gate (see 3.3.1).

In summary, prospective time perception is based on an "internal clock" whilst retrospective perception is based on memory. Prospective estimates get shorter as attention is diverted away from the time estimation task and retrospective estimates get longer ("filled-duration" illusion; Fraisse, 1963). When attention is concentrated on the time estimation task, prospective estimates get longer as do retrospective estimates, although significantly less so (the "watched pot" illusion; Block, George, and Reed, 1980). The internal structure of events also appears to be critical to duration perception. For retrospective perception this appears to be due to the increase in event boundaries stored in memory whilst prospectively the increasing complexity of the event may increase demands on cognitive resources such as attention which is responsible for changes in perceived duration.

5.4.3 Chronostasis

So far distortions of time perception have been discussed as occurring either due to the structure of the sensory stimuli or its cognition. Attention has been seen to effect time perception but only by increasing the cognitive resources allocated to the processing of sensory information once it has been internalised. Nothing has been said about the act of extracting information from the outside world. Under active time perception conditions when a subject is overtly redirecting their attention to judge the duration of an external event, further distortions are seen.

The best known example of this is the "stopped clock" illusion or chronostasis effect (Yarrow, Haggard, Heal, Brown, & Rothwell, 2001b). When making a saccadic eye movement to a clock the second hand seems to stick longer than expected on the first second before continuing on at its normal pace. Time momentarily seems to stand still. This phenomenon is believed to be caused by the extension of the visual information falling on the retina after the saccade backwards over the course of the saccade to cover the period of saccadic suppression. Unfortunately, this estimate of how long perception has been suppressed during the saccade is systematically 50ms larger than the actual time (Yarrow et al., 2001b). This leads to the perceived

duration of the stimuli immediately after the saccade being 50ms longer than its actual presentation time.

The chronostasis effect has been shown to be remarkably consistent occurring across a range of saccade lengths (Yarrow et al., 2001b), types of saccade (e.g. self-timed, reflexive, antisaccades, and, express saccades; (Yarrow, Johnson, Haggard, & Rothwell, 2004), and under different levels of arousal (Yarrow, Haggard, & Rothwell, in press). The consistency with which the effect has been reproduced in the laboratory suggests that every time we perform a saccade we over perceive the duration of the eye movement by ~50ms. However, it is only when we are saccading to a temporal referent (such as a clock or a rhythmical or predictably paced event) that we become aware of the effect.

This 50ms over-extension matches the time course of other known pre-saccadic mechanisms. Saccadic suppression begins around 50ms before the actual eye movement (Diamond et al., 2000) and during this period space is perceived as being compressed towards the saccade target (Lappe, Awater, & Krekelberg, 2000). There is also evidence that the receptive fields[109] of neurons in the lateral interparietal areas of monkeys are remapped to the saccade target ~80ms prior to a saccade (Duhamel et al., 1992). This is thought to equate to the shift of covert attention to the saccade target prior to the overt eye movement (Posner et al., 1980). These phenomena are all signs of perception preparing for the saccade and chronostasis may be a way of compensating for these inaccuracies by extending the post-saccadic percept[110] back over this period.

[109] Receptive fields refer to the portion of the retinal information processed by neurons in the visual cortex. Prior to a saccade, this relationship is remapped so neurons are momentarily associated with the receptive fields currently corresponding to the saccade target.

[110] The product of perception.

5.4.4 Summary of Temporal Continuity

This discussion indicates that time perception is subject to large variability due to attention, memory, arousal, the distribution of cognitive resources, as well as the form of the sensory stimuli being timed. In short, there is no absolute perception of time. However, this variability does not mean that temporal expectations are not formed and their violations not detected. The reason we are able to experience the chronostasis effect is because we expect a saccade to take longer than it actually does and when this expectation is used to "fill in the blank" of saccadic suppression we perceive the event containing the saccade as lasting longer than it would usually last.

If this conception of time perception is now applied to the spatiotemporal expectations associated with visual indices the problem of knowing what form these expectations take becomes very apparent. Temporal continuity of an object will be perceived differently depending on whether a saccade has occurred, the object is moving, perception of the object is cognitively effortful or the viewer is performing a secondary task, whether the viewer was aware of the time estimation task beforehand and if the viewer was distracted at any point during the object's duration. Under real-world viewing conditions, this variability will only be detected if the viewer's attention is directed towards their temporal expectations such as when looking at a clock of waiting for an occluded object to reappear. When watching film, such situations are much more frequent as every cut requires the viewer to reassess temporal continuity in order to assign the visual index. As such, a match-action cut would have to accommodate these distortions of time otherwise temporal discontinuity would be perceived and the object depicted across the cut would lose its visual index and object file. This would create the perception of two distinct objects not connected by spatiotemporal continuity. If temporal expectations are accurately accommodated and the object is presented on screen where expected (spatial continuity) the visual index should track the object across the cut and there will be no need for the contents of the object file to be checked. This suggests that in Levin and Simons' actor-change films (Levin & Simons, 1997), and all match-action

films where the action is perceived as temporally continuous across the cut, editors must be accommodating distortions of time perception.

5.5 Summary of Continuity Perception

To detect continuity errors the shot after a cut must be compared to what is remembered of the previous shot. It is a task that, for most film viewers is unnecessary and distracts from the main point of watching a film: following the action. However, if the continuity error involves a violation of spatial or temporal expectations about the main object's motion the error will be perceived automatically. This is because spatiotemporal information is critical for the successful indexing of objects across visual disruptions (such as visual occlusions, saccades, or blinks). Viewers can also detect continuity errors related to an object's features (such as costume changes) but only if they actively look for such changes. Such object features are stored in an object file which is added to/updated every time the object is fixated. This object file 'sticks' to the moving object via a visual index. This index contains no information about the object other than its location in space and time. The object will continue to be perceived across any visual disruption as long as this visual index is immediately assigned to the object after the disruption. If the object moves out of a spatial 'constancy window' a few degrees of a visual angle wide or cannot be found within 30ms[111] of the visual disruption the visual index will be lost and the object will cease to be perceived (Deubel, 2004).

[111] Notice how this time constaint for perceiving visual stability of an object matches Pöppel's primitive unit of time. Pöppel believes that if two sensory events occur within 30ms of each other then they will be perceived as a single, unified event. If there is more than 30ms between them they will be perceived as two successive events. This idea explains why an object must be found within 30ms after a visual disruption. If the visual index cannot be assigned to the object within 30ms the viewer will perceive a unit of time between the old index and any new index that is later assigned. This temporal discontinuity deletes the existing visual index making the object distinct from whatever the index is next assigned to, even if that happens to be the same object.

The evidence and theories presented over the course of this chapter indicate that the "continuity" referred to in the film literature is actually existence constancy of the principle object across a cut. This is reliant on the object depicted in the shot after the cut satisfying spatial and temporal expectations. These expectations are based on the visual scene as presented prior to the cut and accommodated by the rules of continuity editing to ensure the objects of interest retain their spatial locations after the cut. The visual disruption of the cut and associated changes in visual attention and eye movements distort temporal perception. For existence constancy to be perceived across a cut, these distortions need to be accommodated by the editing. If such distortions can be found in continuity editing then it can be inferred that editors must be sensitive to the need for spatiotemporal continuity and the distortions they under go during film perception. Looking for evidence of these distortions in continuity editing and investigating the combinatorial effects of different types of temporal distortions found in film will be the topic of the final experiment of this thesis.

Chapter 6: Accommodating Expectations

6.1 Experiment 2: Introduction

Previous chapters of this thesis have identified systematic shifts of attention that can be used to limit viewer's awareness of a cut (see chapters 3 and 4). The assumed benefit of such continuity of attention is that it leads to perceived existence constancy of the focal-object[112]. However, no evidence of the perceptual consequences of such attentional continuity has yet been seen. The previous chapter decomposed continuity and identified its components: object, spatial, and temporal. In doing so it also acknowledged the perceptual distortions associated with these components. Of critical importance for the continued perception of an object is spatiotemporal continuity. If the object is perceived where and when it was expected then it will be perceived as continuing to exist (Kahneman et al., 1992; Pylyshyn, 1989). However, both spatial and temporal expectations are imprecise, subject to distortion under different viewing conditions and specifically affected by the allocation of attention. If an editor is aiming to achieve continuity, as well as manipulating attention in a way that hides the cut they must also accommodate these perceptual distortions. This chapter will attempt to find evidence for such accommodation.

[112] The object to which attention is focussed.

6.1.1 Separating Spatial and Temporal Expectations

Spatiotemporal expectations are based on the projection of an object's motion prior to a visual disruption forwards over the duration of the disruption (Michotte, 1991). If an object travelling at a constant velocity moves behind an occluder it is expected to continue moving in the same direction and at the same speed (Hirsch, 1982). If either the location at which the object reappears is considerably different to the expected location or the object takes considerably longer to reappear than expected, the object will be perceived as having changed identity during occlusion (Flombaum & Scholl, in press; Nevarez & Scholl, in preparation; Spelke et al., 1995).

In order for the viewer to predict when and where the object will reappear they need to be able to keep track of how long the object has been occluded. This requires an internal temporal referent i.e. an internal clock (Treisman, 1963). Each "beat" of this clock can be used to extrapolate a current position of the object based on its direction of travel and velocity prior to occlusion. Evidence that such an extrapolation of object motion occurs during occlusion has been shown in the tracking behaviour of monkeys (Churchland et al., 2003) and the neuronal activation of adults (Olson et al., 2004).

The problem with using the internal clock to predict when an occluded object will reappear is that the reliability of the clock is dependent on attention allocation (see (Zakay & Block, 1996) and 5.4). If the period of occlusion is short enough that the viewer can continue to track the hypothetical position of the object during its occlusion (as seen in Churchland et al., 2003) their estimate of when it will reappear should be quite accurate as no change in attention has occurred. However, as soon as attention is shifted during the object's occlusion, either overtly, covertly, internally, or externally, the viewer's perception of time will distort and their ability to predict when the object will reappear will deteriorate. Given that viewers have exhibited a tendency to saccade to the occluder's edge in anticipation of the object's reappearance (Johnson et al., 2003) the majority of estimates concerning when the

object will reappear will be subject to attentional distortion. The result of this inaccuracy appears to be a degree of tolerance of when the object reappears. Instead of the viewers expecting the object to reappear at a precise time point they accept its appearance much sooner or slightly later than expected (Nevarez & Scholl, in preparation; Spelke et al., 1995).

However, the existence of the Chronostasis effect shows that a deviation of as little as 50ms from the expected duration of an object across a saccade can be detected (Yarrow et al., 2001b). This apparent difference between temporal sensitivity during occlusions compared with saccades might actually be a product of the stimuli rather than the viewing conditions. Chronostasis experiments are typically conducted with static stimuli such as a numerical counter where as occlusion experiments require the focal object to move behind the occluder. One of the side-effects observed in experiment 1 (see 4.4) was the viewer's apparent inability to accurately perceive the initial position of the moving object after the cut. This was attributed to the *Fröhlich Effect:* when a moving object suddenly appears on the screen, whether in clear-view of from behind an occluder, the object's initial position is typically misperceived further along its path (Fröhlich, 1923). This distortion has been attributed to the covert reallocation of attention to the moving object after its onset (Müsseler & Aschersleben, 1998). The sudden onset of an object has been shown to reliably capture attention (Boot et al., in press; Brockmole & Henderson, in press; Folk et al., 1994; Franconeri & Simons, 2003; Theeuwes, 1994; Theeuwes et al., 1998; Yantis & Jonides, 1984). This covert shift of attention creates a period of perceptual blindness (Van der Heijden, 1992). This is believed to be 'filled-in' based on the object's position after the attentional shift by which time the object has changed position creating the *Fröhlich Effect* (Müsseler & Aschersleben, 1998).

This inability to accurately perceive a moving object's spatial location after an attentional shift combined with the distorted perception of duration caused by the shift, both covert (Block et al., 1980 and overt Yarrow et al., 2001b) explains why viewer's are unable to accurately predict when an object should appear from behind an occluder (Nevarez & Scholl, in preparation; Spelke et al., 1995). The perceived

spatial distortions seem to compound the temporal distortions and result in seemingly inaccurate spatiotemporal expectations. However, there is currently no clear understanding of interaction between spatial and temporal distortions and whether the associated expectations are as imprecise as they appear.

To begin to understand the relationship between the perceptual distortions of space and time and their resulting affect on the perception of spatiotemporal continuity either space or time must be manipulated whilst the effect on the other factor is monitored. Controlling spatial expectations whilst monitoring the distortions of perceived time is relatively simple: a static object with a temporal component (e.g. a flashing light or change in colour) could be suddenly shifted (creating a spatial discontinuity). This sudden movement would result in a saccadic eye movement to continue following the object. As the shift in attention associated with a saccade is known to lead to the extension of perceived duration (Yarrow et al., 2001b) the spatial discontinuity could be said to cause a temporal distortion. The key component in this process is the attentional shift but the cause of the shift is the spatial discontinuity.

Creating similar attentional shifts by controlling temporal expectations is more difficult. As soon as there is a shift of attention, evidence suggests that there will be a corresponding distortion of perceived time (Zakay & Block, 1996). Shifts in attention also result in momentary distortions of perceived space (Lappe et al., 2000; Ross, Morrone, & Burr, 1997; Ross et al., 2001) but these are only usually noticeable if the attentional target is moving (Freyd & Finke, 1984; Fröhlich, 1923; Müsseler & Aschersleben, 1998). In order to control the degree of temporal distortion the attentional shift would have to be made systematic and reliable. For example, the rhythmical presentation of a saccade target could be used to build up temporal expectations of when an object would appear. Once the viewer had settled into the rhythm the duration of their saccades and the corresponding degree of temporal distortion may become constant. Their temporal expectations could then be violated by presenting the saccade target earlier than expected (creating a temporal discontinuity). The sudden unexpected onset of the target might trigger an

involuntary saccade, overriding the planned voluntary saccade. As the size of the saccade is still the same the temporal distortions may remain the same and any distortion of perceived space could then be directly attributed to the involuntary attention shift.

However, without first knowing how such a change in attention from voluntary to involuntary effects temporal perception the possibility that the spatial distortions are caused by temporal distortions could not be ruled out. A possible explanation of the *Fröhlich Effect* could be that the perceptual extension of the saccade duration leads the viewer to expect the moving object to be further along its path than it actually is. The spatial distortion could then be seen as a compensation for the temporal distortion not a direct result of the saccadic eye movement. Spatial distortions such as the *Fröhlich Effect* can be removed by using a static stimuli but the potential for temporal distortions will exist whenever an attentional shift occurs (Zakay & Block, 1996). Therefore, to investigate how time and space perception are affected by changes in attention and, ultimately, combine to form spatiotemporal expectations, spatial expectations must first be controlled and temporal expectations assessed in isolation. This will be the method used during this study.

6.1.2 The range of viewing conditions created by editing

To investigate spatiotemporal expectations across cuts the spatial and temporal relationship between two shots needs to be made explicit. In chapter 5 the dimensions of continuity were identified as object, spatial, and temporal. For every cut the relationship between the shots either side of the cut can be expressed in terms of continuity or discontinuity in each of these three dimensions. Crossing all variations of these three dimensions creates 8 categories of cut (see Table 2). For example, a match-action cut requires all three dimensions to have continuity across the cut (category 1). If an action is incorrectly matched across a cut, even though the location of the object and its identity remain the same, temporal continuity cannot be

assumed (Bordwell et al., 1985). The cut will be perceived as an Ellipsis of an undefined duration (category 5).

No.	Temporal	Spatial	Object	Cut Categories
1	continuity	continuity	continuity	Match-action
2	continuity	continuity	discontinuity	POV or Reverse
3	continuity	discontinuity	continuity	Matched-Exit/Entrance
4	continuity	discontinuity	discontinuity	Crosscutting
5	discontinuity	continuity	continuity	Ellipsis or Repetition
6	discontinuity	continuity	discontinuity	Montage
7	discontinuity	discontinuity	continuity	Following or Flashback
8	discontinuity	discontinuity	discontinuity	Establishing shot

Table 6-1: The three dimensions of continuity and the resulting categories of cut.

These dimensions of continuity accurately represent the ways that spatial and temporal continuity are discussed in relation to film narrative (Bordwell et al., 1985; Bordwell & Thompson, 2001). For a viewer to understand spatial, temporal, and object relationships at this level they must perceive the contents of the shots and identify referents that allow the deduction of discontinuities. For example an object's changed location relative to a landmark (indicating a spatial discontinuity), a change in daylight (temporal discontinuity), or identity of the actor performing an action (object discontinuity). However, continuity editing actively tries to avoid the viewer's conscious reconstruction of the relationship between shots (Bordwell et al., 1985). Instead, continuity editing implies continuity within the film-world by controlling how the film-world is presented. The clearest example of this is the temporal implication of a dissolve compared that of a cut. A cut creates an instantaneous transition from one visual scene to another. As there is no time between the presentation of the first and second scenes there is nothing in the transition to imply that any time has been omitted between the scenes. In order for a viewer to detect an ellipsis they would need to deduce the time difference by comparing the contents of the two shots. By comparison, a dissolve is an artificial blurring of the end of one visual scene and the beginning of the next. As the

transition has a duration the implication is that time has passed between the two shots (Lindgren, 1948). This uncertainty of the temporal relationship between the presentation of the shots has been found to be a useful technique for removing any expectations of temporal continuity the viewer might have had (Bordwell et al., 1985).

This distinction between continuity of the film's appearance on the 2D screen and the continuity of the depicted 3D scene has previously been acknowledged by editing theorists (Block, 2001; Bordwell et al., 1985; Bordwell & Thompson, 2001; Murch, 2001). Out of six criteria to be considered when deciding how to cut, Walter Murch placed 3D spatial continuity of the depicted scene as the least important (Murch, 2001)[113]. He commented that this highlights a failing of the film schools for focussing too heavily on teaching how to ensure 3D spatial continuity (i.e. the 180° Rule) when in practice this criteria is relatively unimportant (Murch, 2001; page 17). However, as indicated by the discussion of the 180° Rule in section 5.3.3 correct application of the rules actually seems to ensure some of the other criteria considered more important by Murch: eye-trace (the location and movement of the viewer's attention within the screen) and the 2D composition of the screen. Bruce Block attributed similar importance to "eye-trace" for controlling viewer's experience of a cut (Block, 2001). Evidence of a similar intuition into the importance of controlling viewer's attention across cuts has been found as early as 1920 (see Bordwell et al., 1985, pages 235-236). It appears that the distinction between continuity of the depicted 3D scene and continuity of viewer's attention across the 2D screen is an integral part of continuity editing.

[113] The six criteria were (in order of importance): emotion, story, rhythm, eye-trace, 2D spatial continuity of the screen, and 3D spatial continuity of the depicted action. Murch believes that continuity of the most important features can obscure continuity errors of the least important features.

6.1.2.1 Defining 2D and 3D continuity

Both spatial and temporal continuity can be expressed in terms of 2D and 3D relationships between the visual information either side of a cut. The 3D level is identical to that previously used to describe the constituents of continuity (see 5.1.1). The 2D level refers to the way that this 3D scene is presented on the 2D screen and, specifically, how attention is allocated to it.

Three-Dimensional (3D)

- **3D-object** = the identity and appearance of an object within the depicted 3D scene. For 3D-object continuity to be perceived, scene-space and film-time continuity must be assumed (see 5.1.1).

- **Scene-space** = the relative position of 3D-objects within the depicted 3D scene. A scene-space discontinuity occurs if this relationship changes across a cut. This is often the level of continuity violated by a spatial continuity error e.g. when an object shifts relative to other objects within the scene across a supposedly temporally continuous cut (see 5.1.1).

- **Film-time** = the temporal relationship between the scenes depicted across a cut. A film-time discontinuity occurs when the new scene does not follow on from the previous scene e.g. the hands on a clock suddenly jumping forwards across a match-action cut.

Two-Dimensional (2D)

- **Focal-object** = the abstract object to which attention is first allocated after the cut. For focal-object continuity to be perceived the object does not need to keep the same identity across a cut it just has to be present on the screen immediately after the cut so that attention can move directly to it. A focal-object discontinuity will be experienced if the focal-object is not present on the screen immediately after the cut or the new shot is composed in such a way that no individual focal-object is immediately apparent. This

representation of 2D object continuity as an abstract target for attention is derived from existence constancy (see 5.3.1) and eye tracking evidence (see 3.4.2.2).

- **Screen-space** = the position of the focal-object on the screen. A screen-space discontinuity occurs when the focal-object shifts position on the screen across the cut. It is important to note that the screen-space relationship is between the focal objects either side of the cut irrespective of their identity. Identity is the concern of the 3D level of continuity.

- **Real-time** = the temporal relationship between the last appearance of the focal-object prior to the edit and its first appearance afterwards. A dissolve or fade is the obvious technique used to create a real-time discontinuity but having the focal-object leave the screen or take time entering the screen after a cut is also a common technique (see Katz, 1991 pages 155-156).

This chapter is principally concerned with the investigation of spatiotemporal continuity, as such both 3D-object and focal-object will be excluded from the current investigation. As object information cannot be excluded completely without also removing spatial and temporal information (a film-time discontinuity cannot be perceived without evidence of the elapsed time on an object) instead it will be controlled. The identity of the focal object and its presence on the screen immediately after a cut will be kept constant. Future studies should focus on 3D and focal-object expectations across cuts but for now the spatiotemporal expectations that form the basis of existence constancy will be investigated instead.

Representing the 2D relationships between shots in terms of the focussing and shifting of attention is important for understanding the development of spatiotemporal expectations across a cut. As has already been seen in this chapter (section 6.1.1), the allocation of attention during the development of spatiotemporal expectations affects the final form of these expectations. The 2D relationships express how attention is shifted across a cut so, by reference to what is already

known about spatial and temporal distortions, changes in 2D continuity can be used to predict how space and time are perceived across a cut. The classic interpretation of why the 180° Rule is so important is that it ensures the accurate perception of 3D spatial continuity (Bordwell & Thompson, 2001; Katz, 1991; Reisz & Millar, 1953). However, the viewer never has direct access to the 3D relationships between shots; their access is always mediated by the 2D level. The decisions made when choosing how to frame each shot and then position it relative to another shot dictate how the 3D scene will be presented to the viewer i.e. the 2D relationships. Given the rapid and automatic allocation of attention across a cut (see 3.4.2.2), these compositional decisions also control how attention is initially distributed across the cut[114]. For the 180° Rule to ensure the accurate perception of the 3D spatial relationships between shots the rule would need to create 2D relationships that are perceived as being equal to the 3D spatial relationships. For example, if a 3D-object maintains its scene-spatial location (i.e. its position within the 3D space of the scene) the desired perception is of spatial continuity. The 180° Rule could ensure this by creating a 2D relationship that is also perceived as spatially continuous e.g. the focal-object does not change screen-spatial position across a cut. This is exactly what is seen when the 180° Rule is used to compose a series of shots (see 5.3.1).

Controlling how temporal relationships are perceived across cuts is more complex. No manipulations of the 2D temporal relationship (i.e. real-time) are possible as the cut ensures 2D temporal continuity and the absence of motion means that the focal object cannot move out or on to the screen[115]. This leaves only modification of the 2D spatial relationships (screen-space) and their associated shifts of attention to affect the perceived temporal relationships between shots. Each shift in attention has been shown to have an affect on the perceived duration of the attentional shift

[114] The 2D relationships only control attention immediately after the cut as it is at this point that the sudden onset of new objects captures attention. After the new shot is established the distribution of attention across the screen becomes more voluntary. This is seen as an increase in variance between where different viewers focus their gaze (see 3.4.2.2).

[115] Motion is excluded from this discussion as the spatial distortions associated with perceiving a moving object were deemed too dependent on temporal distortions.

(Yarrow et al., 2001b; Zakay & Block, 1996). If an attentional shift occurs at the beginning of a shot the perceived starting time of an event depicted in the shot my deviate from the event's actual starting time (i.e. its film-time). Such a deviation would need to be accommodated if the film-time were to be perceived as continuous across the cut. To identify how this accommodation can occur a variety of different types of cut will be dissected into their 2D and 3D constituents. The assumption of this chapter is that the accommodations required to create the perception of temporal continuity will match the prescriptions of the continuity editing rules.

6.1.2.2 Jump Cut

Figure 6-1: Two consecutive shots from Dancer in the Dark (Lars von Trier, 2000).

The best place to start classifying cuts in terms of their 2D and 3D constituents is with the type of cut viewed as the worst type of discontinuity: the Jump Cut. A Jump Cut is "An elliptical cut[116] that appears to be an interruption of a single shot. Either the figures seem to change instantly against a constant background, or the background changes instantly while the figures remain constant" (Prunes, Raine, & Litch, 2002). Jump cuts are anathema to the continuity style of editing as they involve a spontaneous, unmotivated change of shot where the spatial and temporal relationship between the new shot and the old is undisclosed. Because of their opposing nature to the conventional form of editing they have often been associated with experimental, avant-garde or radical film making movements. The most significant of these were the French Nouvelle Vague in the 1960's lead by Jean Luc Goddard and Dogme 95 lead by Lars von Trier (amongst others; see www.dogme95.dk).

[116] A cut involving the removal of a period of time.

An example of a jump cut taken from Lars von Trier's Dancer in the Dark can be seen in Figure 6-1[117]. The key factors that make this cut a jump cut are that the focal-object, Bjork's character at the centre of the screen does not move across the cut (i.e. screen-spatial continuity) whilst the camera has shifted position. This causes a spontaneous change of the background scene and, in this example, an apparent change of the focal-object's scene-spatial location (notice how the character to Bjork's left changes between shots). Such a spontaneous change of camera position would only be acceptable according to the continuity editing rules if the cut was matched to an action. In this example the focal object and the soundtrack do not motivate the cut in anyway. This absence of motivation for the cut is also critical for a cut to be classified as a jump cut[118]. Also, a Jump Cut must, as the name suggests, be a cut rather than a fade or dissolve. If the edit involves a real-time discontinuity all assumptions of film-temporal and scene-spatial continuity will be abandoned. It is this assumption of spatial and temporal continuity of action across a cut that is violated by the Jump Cut (Bordwell et al., 1985).

In terms of the 3D and 2D relationships across the cut a Jump Cut can be classified as:

Type of Cut	Scene-Spatial	Film-Time	Screen-Spatial	Real-Time
Jump Cut	discontinuity	discontinuity	continuity	continuity

This classification allows the identification of the possible effect of a Jump Cut on attention. Given that the viewer will probably be fixating the character at the centre of the screen (the focal-object) and they do not expect the cut, the spontaneous

[117] Dancer in the Dark is not technically a Dogme 95 film although it does comtain a lot of the techniques and aesthetics of the Dogme Manifesto.

[118] This is the belief of the author of this thesis. An absence of motivation is believed to be critical to the creation of a jump cut as if the cut was motivated the viewer would either expect the cut, minimising its disruptive effects, or change their attention in a way that would hide the cut (see chapter 3).

change of the background scene has the potential to capture attention (see 2.3). Experiments in time perception have shown that if a viewer is judging the duration of an event and they are covertly distracted they will perceive the event as having a shorter duration than without the distraction (Block et al., 1980; Zakay, 1992). The period of distraction is not accounted for when judging the overall duration within which it occurred. Therefore, if a visual event is presented across a Jump Cut the new shot will be perceived as starting later than it actually started. In other words, the viewer will perceive a temporal discontinuity even though presentation of the event was temporally continuous (see Figure 6-2). Until the temporal continuity of the depicted scene (the film-time) can be deduced (e.g. by comparing temporal referents to those recalled from the previous shot) the viewer is likely to assume that their perceived temporal discontinuity reflects a film-time discontinuity.

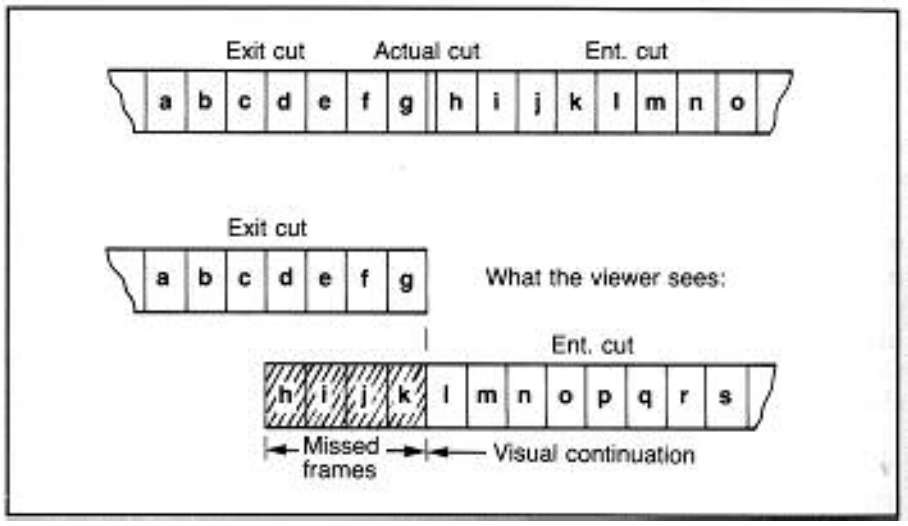


Figure 6-2: The "loss" of 4 frames of visual information due to distraction by the sudden background change across a Jump Cut. 4 frames is just an estimate. The actual degree of distraction is not yet known. Taken from (Dmytryk, 1986) pg 439.

To accommodate this perceived temporal discontinuity the depicted visual event would have to be presented for longer. This would ensure that once the distracting effects of the Jump Cut were over and attention had returned to the focal-object the time point at which the event was perceived as beginning followed on from the last time point depicted prior to the cut. If an action was presented across the cut the same effect could be achieved by replaying the last few frames of action from the previous shot (see Figure 6-3).

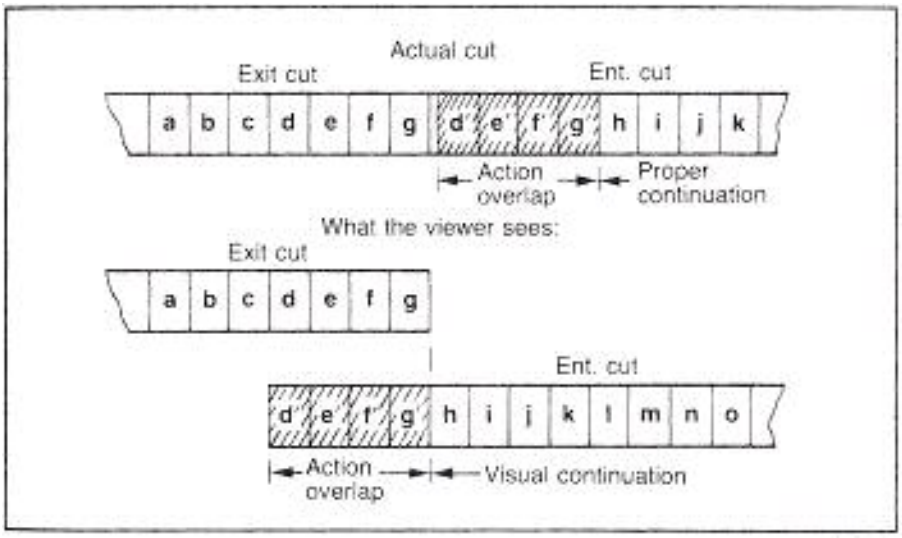


Figure 6-3: Creating action continuity by replaying 4 frames of the action after a Jump Cut. Taken from (Dmytryk, 1986) pg 440.

This method of creating perceived temporal continuity by presenting a temporally discontinuous visual scene seems to have some support from editors. Joseph Anderson, film editor and cognitive film theorist, suggests that "the action (not the actual film) should be overlapped approximately two frames when making the splice[119]" (Anderson, 1996; pg 100). He believes this is necessary as:

> "In a motion picture, a new shot simply overrides (masks) the processing of the last couple of frames of the old shot. Editors know, from previous experience, that cuts where the action is perfectly matched do not work very well, and whether they understand masking or not, they ensure that there will be no gaps in the action presented to our visual system by simply overlapping the action by approximately two frames at the cut. (They proceed by trial and error to find the exact number of frames of overlap needed at a particular cut.)" (Anderson, 1996; pg 103)

Anderson recognises that this default overlap of two frames may vary depending on the cut but he implies that some degree of temporal overlap will be necessary for all match-action cuts. Whilst the Jump Cuts discussed above are not match-action cuts (in fact the inclusion of an action motivating the cut would stop them being Jump

[119] i.e. the cementing together of one strip of celluloid to another.

Cuts) the compatibility between Anderson's prescription and the temporal overlap hypothesised as a way to create perceived temporal continuity is striking. However, Anderson's explanation of the need for the overlap as being due to visual masking cannot be applied to Jump Cuts as masking requires the visual information presented at fixation to be suddenly replaced (Breitmeyer, 1984). As the focal-object does not change position on the screen across the cut the visual information at fixation should not change enough to cause visual masking. The need for part of an action to be overlapped across a Jump Cut is not because the end of the previous shot has been masked but because attention is distracted during the start of the new shot[120].

Whilst the need for a visual event to be presented as longer than its actual duration so that it can be perceived as temporally continuous seems to have a theoretical and practical basis it is yet to be shown empirically. This will be one of the objectives of the current empirical study.

> **<u>Distraction Hypothesis:</u> When the background changes unexpectedly whilst the focal object is fixated the subsequent visual duration will be perceived as shorter than the same duration without background change.**

For the sudden change in background caused by the Jump Cut to cause distraction the viewer must not expect the change. If the change is cued in someway, such as by matching an action across the cut, the viewer will adopt an attentional set that focuses their attention on the focal-object and limits the degree that the peripheral change captures attention (see chapter 3). Whilst motion has been excluded from this study to ensure the effects of attention on temporal perception are unconfounded there still exist other ways to cue a cut. For example, the onset of the cut can be made predictable by tying it to some rhythmical quality in the depicted seen or object. If the viewer is pre-informed that a Jump Cut will occur they will be able to anticipate the timing and exact form of the cut. This should limit the degree that the

[120] At least this is the assumed explanation for the need for temporal overlap chosen for this current study.

background change distracts them from the focal object allowing them to accurately perceive time across the cut.

<u>Resistance to Change Hypothesis:</u> An expected background change will have no effect on the perceived duration of a subsequent visual event.

6.1.2.3 Stop-motion

Figure 6-4: Two shots either side of a Stop-Motion cut. The focal object changes position on the screen across the cut whilst the background remains the same.

In the Jump Cut the background was seen to spontaneously change whilst the focal object remained in the same position on the screen. If the opposite change is made, the background kept constant whilst the focal-object changes position, a different type of cut is created: a stop-motion cut. To create this cut the location of the focal-object must change within the depicted 3D space across the cut (a scene-space discontinuity). The background is kept constant by maintaining the camera position used to film the shot prior to the cut. This results in the focal-object shifting across the screen creating a screen-spatial discontinuity equivalent to the scene-space discontinuity. Given that the object has spontaneously changed position within the 3D scene the film-time cannot be continuous. However, the fact that the object is still present on the screen indicates a real-time continuity. These properties are clearly mapped out below:

Type of Cut	Scene-Spatial	Film-Time	Screen-Spatial	Real-Time
Stop-motion	discontinuity	discontinuity	discontinuity	continuity

The perceptual consequence of a stop-motion cut is apparent motion: the perception of motion from the alternating presentation of two static objects. Apparent motion is the mechanism by which all filmed media functions. Without our ability to perceive rapidly presented static visual scenes as moving there would be no film, TV, or animation. Film is recorded by exposing the frames of the film at the same rate that they will eventually be projected. Each frame records a moment of an action and then when these moments are projected in quick succession (with 30-200ms between each frame) the impression of motion is recreated (Nichols & Lederman, 1985). The precise reason why apparent motion (or stroboscopic motion as it is technically known) occurs is not known but it is thought to be a side effect of the neuropsychology of real motion perception (see Rock, 1985; and Anderson & Anderson, 1985 for a discussion).

Essentially, a spatial and temporal discontinuity exists between each frame of a film but we are unable to perceive these discontinuities as the frames are presented so rapidly that they are perceived as continuous. The spatial and temporal discontinuity involved in the stop-motion cut described above is different to these between-frame discontinuities as it does not represent an incremental change in the object over time. Instead the focal-object suddenly and unexpectedly relocates. This may still result in apparent motion (depending on such factors as the saliency of the focal object, and the distance moved across the cut; Korte, 1915) but given the size of the relocation and the absence of the duration required to normally perceive such a relocation it will probably be perceived as a 3D spatiotemporal discontinuity.

Apparent motion has been shown to be a reliable attractor of attention (Folk et al., 1994). The sudden relocation of the focal-object across a stop-motion cut is likely to capture attention resulting in an involuntary saccadic eye movement to the object. According to the chronostasis effect, such a saccade will be perceived as having a duration longer than the period of time actually taken to move the eyes (by about

50ms) (Yarrow et al., 2001b). This perceptual extension of the saccades duration occurs both when the saccade is voluntary and involuntary (i.e. the result of unexpected attentional capture) (Yarrow et al., 2004). Therefore, if the focal-object is associated with a visual event of a set duration the perceptual extension caused by the saccade at the start of the event will result in the overall duration of the event as being perceived as longer than a similar event viewed under fixation.

> **<u>Chronostasis Hypothesis:</u> When an object suddenly relocates across the screen at the beginning of a visual event, the event will be perceived as ~50ms longer than the same event viewed under fixation. Whether the relocation is expected or not, the perceived extension of duration will be the same.**

To create the impression that the visual event was presented for the same duration as the fixated event the presented duration would have to be shortened by about 50ms. This ellipsis would accommodate the over perception of time during the saccade and ensure that once perceptual sensitivity had returned after the saccade the first time point to be perceived would follow on from the last time point depicted prior to the cut (see Figure 6-5).

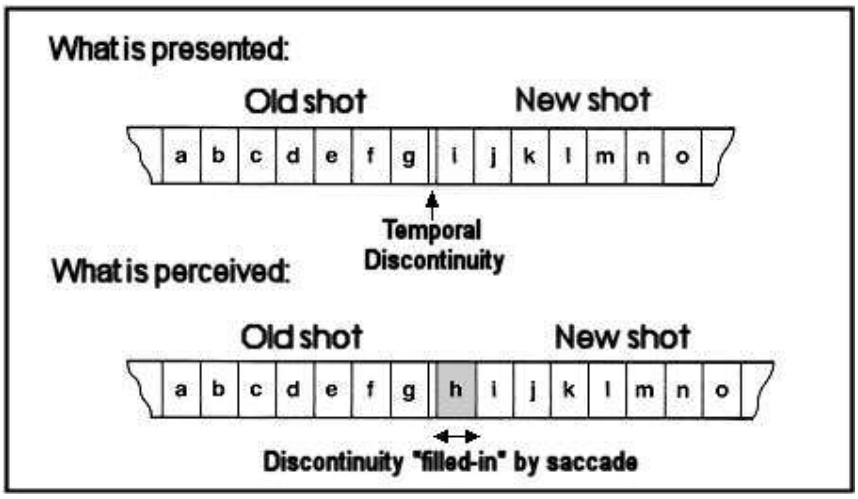


Figure 6-5: The creation of perceived temporal continuity from presented discontinuity across a Stop-Motion cut.

Unlike the Jump Cut, no evidence can be found in the editing literature that editors omit time across stop-motion cuts to create the perception of temporal continuity. In fact, the only detailed discussion of this type of cut suggests that, similar to the Jump Cut, time should be overlapped across the cut (Dmytryk, 1986).

> "….the player rises from his chair in the close shot, and the full shot continues his movement. Here the cut would probably come six to ten frames *after* the start of the action in the close shot[121]. The longer shot would pick up his movement at nearly the same spot. Exact matching of position[122], however, might not result in the smoothest cut….. Often an action overlap of three to five frames is desirable." (Dmytryk, 1986; page 436).

Dmytryk explains that this overlap of action is required "Because [the viewer's] vision has been clouded, or diverted, by the apparently awkward movement of [the viewer's] eyes across the screen" (Dmytryk, 1986; pg 438). To illustrate his point, Dmytryk uses a matched-exit/entrance cut similar to that used in the first experiment (see chapter 4): "The cut to the second scene should be made from three to five frames ahead of the point at which [the actor's] eyes re-enter the frame at the opposite side of the screen" (Dmytryk, 1986; pg 437). He correctly believes that the viewer cannot perceive visual information during a saccadic eye movement and that this saccade would take about a fifth of a second to complete; the equivalent of about 5 frames at 24fps. To cover this 200ms period of "blindness", he suggests that the last 3-5 frames of the previous cut should be replayed immediately after the edit[123]. These duplicated frames will cover the time that the viewer's eyes are in motion so that by the time their eyes have reached the opposite screen edge and they have recovered the ability to perceive the visual information the frame they see is that

---

[121] The techniques using the onset of an action to hide a cut was previously seen in section 3.3.

[122] The position Dmytryk is referring to is the 3D scene-spatial location of the object not its position on the screen.

[123] The diagrams Dmytryk uses to illustrate this overlap have already been used in section 6.1.2.2 to illustrate how distraction can be accommodated (Figure 6-2 and Figure 6-3). Dmytryk never attributed the need for an overlap to distraction but his advice for how perceived temporal continuity should be created across a cut match the conclusions of section 6.1.2.2 better than this current section.

which continues on from the frame they saw before their eye movement. This method for creating perceived temporal continuity from presented discontinuity is similar to that used to accommodate distraction in section 6.1.2.2 (Figure 6-2 and Figure 6-3 are actually taken from Dmytryk, 1986).

Dmytryk's understanding of the perceptual consequences of saccadic eye movements is partially correct. He is right to identify that we cannot perceive during a saccade but he does not appreciate that this perceptual "hole" must automatically be "filled in" by our perceptual system to avoid every saccade being perceived as a visual discontinuity. If we perceived nothing during a saccade, as suggested by Dmytryk, our perception of the world would be made up of static views of space separated by momentary blanks as the eyes move to the next point of fixation. This is emphatically not what we perceive. Instead, as shown by Yarrow et al (Yarrow et al., 2001b), the extrapolation of the visual information found after a saccade back over the saccade's duration over-compensates for the period of perceptual blindness. This active "filling in" should make Dmytryk's suggested compensation redundant. However, Dmytryk's suggested compensation is not based primarily on his incomplete knowledge of the perception consequences of saccadic eye movements; it is based on years of experience as an editor. As such the need for a temporal overlap should not be dismissed outright. One possible instance that may require such an overlap would be if the distracting effect of a Jump Cut were combined with the saccade of a Stop-Motion Cut. Such a combination could be caused by a Reverse-Angle Cut. This will be discussed in the next section.

One last aspect of the Stop-motion cut that should be considered before moving on is the possibility that masking occurs as the focal-object relocates. In the previous section (6.1.2.2), Joseph Anderson's explanation of the need for overlapped action across a match-action cut (Anderson, 1996) was dismissed as in a Jump Cut the focal object is collocated across the cut. Masking only occurs when the visual information presented at fixation suddenly changes (Breitmeyer, 1984). In a Stop-motion cut the visual information at fixation changes from that of the focal-object to the background revealed as the focal-object relocates. If this sudden appearance of the background

masked the last few milliseconds of the previous shot, an action presented across the cut would have to be overlapped to compensate for the masking, just as suggested by Anderson (Anderson, 1996). The chronostasis effect would still occur as attention is captured by the focal-object's relocation but the effect this has on perceived temporal continuity may be modified by the masking. If a Stop-Motion cut is made to a shot depicting an object against a black background the duration of the shot might be perceived as shorter than the same shot with a background. The sudden appearance of the background may mask the end of the previous shot, perceptually extending the start point of the second shot back into the first. No such masking should occur when there is no background. The existence of a masking effect will be investigated in the current study by looking for variations in the chronostasis effect across different background conditions. This is formalised as a Masking Hypothesis:

> **<u>Masking Hypothesis:</u> When the relocation of a focal object at the beginning of a visual event reveals a detailed background, the perceived duration of the event will be longer than the same event presented without a background.**

6.1.2.4 Reverse-Angle Cut

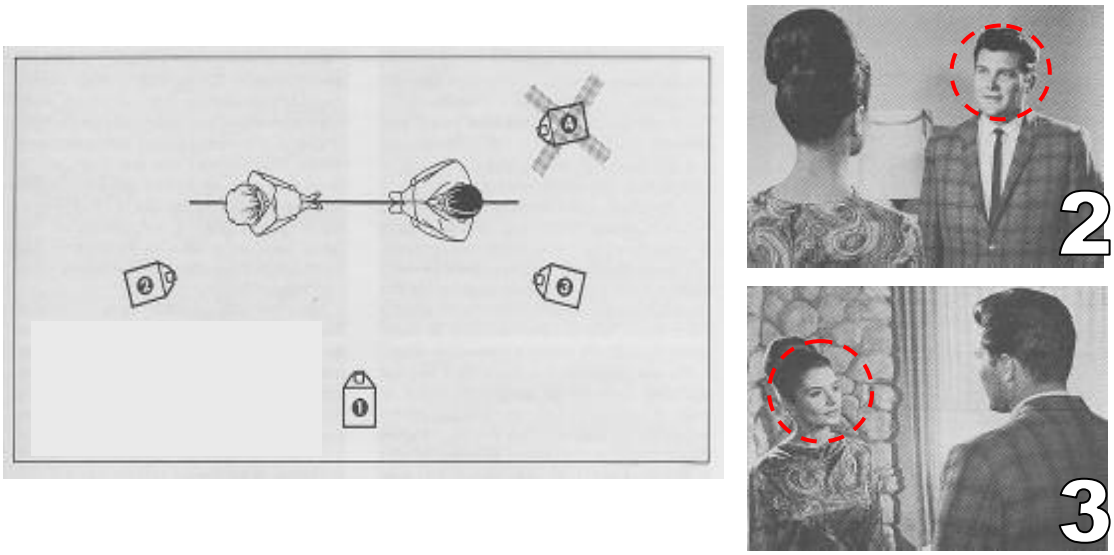


Figure 6-6: The 180° Rule and its application for creating a Reverse-Angle cut. The camera is moved from position 2 around the circumference of an (abstract) 180° arc

surrounding the conversational partners to position 3. This preserves the character's location on the screen and allows the new focal-object (the woman in shot 3) to be clearly seen. The dashed circle represents the focal-object in each shot.

Reverse-Angle editing is most commonly seen in dialogue scenes. When the conversation shifts from one character to another the camera typically needs to be relocated so that the new speaker can be clearly seen. The 180° Rule specifies how such a scene should be shot and edited together (see Figure 5-3). In the first shot (shot 2 in Figure 5-3) the camera is positioned on the left of the scene pointing at the man. When the conversation shifts to the woman the viewer's attention will shift across the screen to the source of the speech (see 5.3)[124]. As only the back of the woman's head is visible in shot 2 the camera needs to be relocated so that the woman's face is in view. The camera needs to at least roughly reverse the angle in which it is pointing (hence the name). The 180° Rule specifies that the best way to accomplish this change in viewpoint is to move the camera along the circumference of the abstract 180° arc that surrounds the conversational partners (see Figure 5-3). By doing this the position of each character is preserved on the screen and the new focal-object (the woman) can be clearly seen.

A Reverse-Angle cut maintains the objects' positions within the 3D space (i.e. scene-spatial continuity) and the relative position of each of these objects on the screen. However, given that the cut coincides with an attentional shift from one object to another the position of the focal-object on the screen changes (in the first shot the focal-object is the man at screen-right, in the second shot it is the woman at screen-

[124] Reverse-angle shots do not always require attention to shift across the screen. Quite often the focal-object after the cut will be collocated with the focal-object before the cut. To distinguish a collocated reverse-angle cut from a Jump Cut the cut would need to be anticipated. This anticipation might allow the capturing effects of the background change to be minimised as predicted in the Resistence to Change Hypothesis (see section 6.1.2.2) or it might lead to some form of attentional withdrawal to such as an eye blink or attentional blink (see 3.3). The effect both of these types of blink have on time perception is currently not known. Investigating their affect on time perception would require closer control over the cues usually used in Reverse-Angle editing to lead attention (such as changes in gaze, off-screen speech, etc; see 3.2.2). An investigation of the effect of such cues is not the intention of this study. Therefore, it will be assumed that the only attentional behaviour occurring across a collocated cut is a change in attentional set as predicted in the Resistence to Change Hypothesis.

left). Therefore, a reverse-angle cut can be interpreted as a screen-spatial discontinuity. A full listing of the cut's properties can be seen below:

Type of Cut	Scene-Spatial	Film-Time	Screen-Spatial	Real-Time
Reverse-Angle	continuity	continuity	discontinuity	continuity

Reverse-angle cuts are very common in the continuity style of filmmaking (Bordwell et al., 1985). Reverse-angle describes a range of different types of cut: Over-the-Shoulder (OTS: such as the shots in Figure 5-3), Point-of-view shots, or shots with matching eyelines (similar to OTS but without the presence of a character's shoulder). All of these shots involve a sudden change in background and a shift of attention from one focal object to another.

The affect of reverse-angle cuts on perceived temporal continuity is not currently known. Evidence of the Chronostasis effect suggests that the shift in attention should lead to a perceived temporal extension (Yarrow et al., 2001b). However, so far all chronostasis experiments have been performed under controlled psychophysical conditions with no background information to distract from the time estimation task (Yarrow et al., 2001b; Yarrow et al., in press; Yarrow et al., 2004). Similarly, experiments investigating the effect of distraction on time perception have been performed without precise control over eye movements (Block et al., 1980; Zakay, 1989; Zakay, 1992; Zakay & Block, 1996). Unlike in the stop-motion example, the sudden relocation of the focal-object can not be guaranteed to capture attention as the Reverse-Angle cut also changes the background scene. If attention was first covertly attracted by some property of the background change the viewer might first perceive a shortening of the perceived duration (see (Zakay & Block, 1996). Only after control of attention had returned after this period of distraction could attention be then voluntarily shifted to the focal-object. This saccadic eye movement would then lead to a lengthening of the perceived duration (Yarrow et al., 2001b). If the initial distraction lasted for ~200ms (~5 frames at 24fps) the chronostasis effect would only compensate for ~50ms of this (~ 1 frame) leaving the perceived time between the cut and the recovery of full attention after the saccade as ~150ms (~4 frames) shorter

than the actual time. This absence of 4 frames could be filled by Dmytryk's suggested 3-5 frame overlap. Given the absence of contradictory evidence and Dmytryk's insistence on the validity of this overlap the current experiment will test whether empirical support can be found for the need for such an overlap to create perceived temporal continuity.

> **Delayed Saccade Hypothesis: When an unexpected background change occurs as the focal-object relocates perceived durations will be shorter than without the background change but longer than without the relocation.**

This effect is conditional on the viewer not expecting the focal-object relocation. If they expect the relocation they will adopt an attentional set specific to this relocation and begin to prepare the saccadic eye movement before the cut occurs (see 3.2). This should make them resilient to attentional capture by the background change. As a result their perception of time across the saccade should only be subject to distortion by the chronostasis effect (i.e. in line with the Chronostasis Hypothesis).

6.1.3 Summary

Whilst the three types of cuts described above do not cover all possible cuts their differences are sufficient to highlight the ways in which editing influences temporal perception, namely the distribution of attention across the cut. As attention changes so does our ability to accurately perceive time (Zakay & Block, 1996). If attention is distracted during an event the perceived duration of that event shortens (Block et al., 1980; Zakay, 1989; Zakay, 1992). If a saccade is performed during the event the perceived duration lengthens (Yarrow et al., 2001b). However, in film the majority of cuts involve some combination of background change and focal-object relocation. Editor's have suggested that, for most cuts, the viewer will be blind to a period of time immediately after the cut (Anderson, 1996; Dmytryk, 1986). This period must be accommodated by replaying 2-5 frames of the action if the temporal continuity is to be perceived across the cut. Overlapping action across a cut seems to be an editing

convention yet no empirical evidence exists that can explain why it would result in the perception of temporal continuity. Generating such evidence will be the main aim of this empirical study.

6.2 Experiment 2: Methodology

6.2.1 Experimental Design

To investigate how the combination of distraction and saccadic eye movements occurring across cuts affects the perception of temporal continuity an experimental paradigm must be used that addresses certain concerns:

- The paradigm must allow the precise control of background change and focal-object relocation across a cut.
- It must permit accurate measurement of perceived time.
- Visual events must be presented that have clear duration. These events will be used to detect temporal discontinuities across a cut.
- However, these visual events cannot contain any motion (see section 6.1.1).
- The visual scenes used must be richly detailed and comparable to those presented during a normal film.

A modification of Yarrow et al's (Yarrow et al., 2001b) paradigm was chosen as it fulfilled all these requirements. Subjects were instructed to focus on a letter which changed every 1000ms[125]. This change of letter was the only action in the visual scene. All letters were presented for exactly 1000ms except one, 'E', which varied. The subject's task was to judge whether the duration of the target letter, 'E' was longer or shorter than the other letters. A Modified Binary Search procedure (MOBS: Tyrrel and Owens, 1988; see section 6.2.2 for more details) was used to control the presentation duration of the target over multiple trials and home in on a duration perceived by the subject as being equal to the duration of the other letters (i.e. 1000ms). The product of the MOBS is referred to as the matched estimate. If a

[125] 1000ms was chosen as the comparison duration as this was the duration used in the original Yarrow et al. (2001) study and it is an easy duration for subjects to judge (people are used to counting in 1 second intervals).

subject perceives the target duration as being longer than the comparison duration, for the two durations to be perceived as equal the target duration will have to be made shorter. This will result in a matched estimate that is shorter than the comparison durations. If the subject perceives the target duration as longer, a shorter matched estimate will be required for the target duration to be perceived as equal to the comparison durations.

This duration estimation task was performed under a variety of viewing conditions designed to replicate previous effects (Zakay, 1992; Yarrow, Haggard, et al. 2001) and identify any interactions between them. Three **Independent variables** were used to create these conditions:

Degree of Target Relocation (referred to as 'Saccade'). As the target letter changed from the previous letter to the target of the time estimation task, 'E', the letter would either remain in the same position (0° relocation) or instantaneously relocate across the screen (this time point is referred to as the "cut"). Two degrees of relocation were used, 20° and 40°, measured as the visual angle between the target's positions before and after the cut (as viewed by the subject). See conditions 1, 2, and 3 in Figure 6-7. This variable allowed the reproduction of the Chronostasis effect (Yarrow, Haggard, et al. 2001).

Degree of Peripheral Distraction (referred to as 'Background'). This variable had three levels: none, static photo, and changing photo. The control condition, "none", presented the target against a black background with nothing to distract the subject from the time estimation task. The "static photo" condition presented the target within a photorealistic scene full of peripheral details (see condition 4 in Figure 6-7). This "static photo" remained in the same position on the screen across the cut irrespective of target letter relocation. The "changing photo" condition began with the same photo background as the "static photo" condition but at the cut the background instantly changed to a different photorealistic scene. This photo was shot within the same locale as the static photo but there was no overlap (this was done to minimise any apparent motion effect as the photos switched).

Predictability of Cut Type (referred to as 'Blocking'). The two variables 'Saccade' and 'Background', each with three levels, were crossed to create 9 experimental conditions (see Figure 6-7). These conditions were either presented randomly during the course of the experiment (Random) or grouped together into blocks of the same condition (Blocked). When conditions were blocked subjects would know in advance how the cut would affect the background or object location. When the conditions were randomised the viewer would be unable to anticipate the cut. This variable was included as expectation has been shown to affect attentional capture (Simons, 2000) and temporal perception (Zakay & Block, 1996). Specifically, when attention is focussed on an event, the perceived duration of the event is longer than when attention is distracted (Block et al., 1980; Zakay, 1992) or divided between stimuli (Zakay, 1989; see 5.4.2 or Zakay & Block, 1996). In order for the effect of expectation to be made explicit in this study matched estimated for the two presentation conditions (Blocked and Random) will be compared. The predicted effect of each presentation condition can be formalised as two new hypotheses:

> **Concentration Hypothesis: When editing conditions are predictable matched estimates will be significantly shorter than the actual value.**
>
> **Divided Attention Hypothesis: When editing conditions are unpredictable matched estimates will be significantly longer than the actual value.**

6.2.1.1 Experimental Conditions

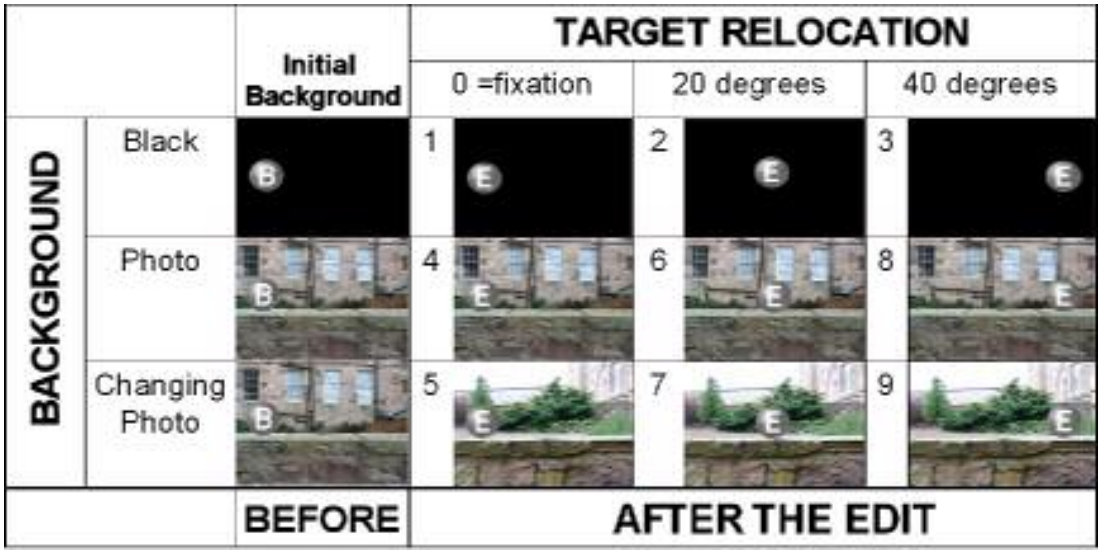


Figure 6-7: Nine experimental conditions.

Condition 1 functioned as the control for all other conditions as it contained no peripheral distraction, changes, or relocation. Matched estimates produced under this condition should provide a baseline for the duration perceived by subjects as being equal to 1000ms under normal viewing conditions. Conditions 2 and 3 isolated object relocations from peripheral distraction to indicate the effect of object relocation alone (and therefore, the corresponding saccadic eye movements). Conditions 4 and 5 modified the degree of peripheral distraction whilst controlling object relocation to identify the individual effects of each degree of distraction compared to the baseline (condition 1). Condition 5 can be seen as a replication of the Jump Cuts discussed in section 6.1.2.2. Conditions 6 and 8 relocate the object across the cut whilst keeping the background constant replicating the Stop-Motion Cuts discussed in section 6.1.2.3. Conditions 7 and 9 vary background and object relocation replicating the Reverse-Angle Cuts discussed in section 6.1.2.4.

6.2.2 Modified Binary Search (MOBS)

To derive estimates for durations perceived by subjects as being equal to 1000ms (referred to as matched estimates) a Modified Binary Search (MOBS: Tyrrel and Owens, 1988) procedure was used. This procedure forced subjects to identify a target duration as being either longer or shorter than comparison durations[126]. Their responses are used, over numerous trials (6-26 in this study) to home in on a duration perceived as 1000ms. The MOBS used in this study (slightly modified from the original) functioned according to the following steps:

1. A range of possible durations was chosen. This range was bound by an upper and lower boundary (1200ms and 800ms respectively in this study). These boundaries are lists with the current boundary at the head and previous boundaries behind it. Either the upper or lower boundary list was updated after every trial.
2. For the first trial a random value between the upper and the lower boundaries was tested[127]. For every other trial the midpoint of this range would be chosen as the next duration to be tested. Subjects are then forced to identify the presented duration as being either longer or shorter than the comparison durations.
3. If the subject responds 'longer', the value is added to the head of the upper boundary list. If they respond 'shorter', it is added to the head of the lower boundary list. The process then returns to step 2.

---

[126] Subjects are not permitted to identify the target as being equal in duration to the comparisons or choose not to respond. Technically, they are required to make a 'two alternative forced choice' (2AFC) response.

[127] This was a modification of MOBS made for this study. It was decided that the midpoint of the range should not be the first to be tested as this would always be 1000ms. Forcing the subjects to identify this as either longer or shorter than 1000ms would create a new range where the matched estimate was very close to one boundary. This would make it very difficult for subjects to correctly identify the boundary and may, therefore require the boundary to be repeatedly confirmed. By randomly choosing the first value, there is more chance of the matched estimate residing somewhere in the middle of the range. This randomness also makes the three repetitions performed for each condition different, adding more confidence in their eventual average matched estimate.

4. When two consecutive responses are identical the validity of the opposite boundary is tested. For example, if two 'longer' responses had been given this could be due to the fact that the lower boundary is also perceived as 'longer' than 1000ms and was erroneously placed as a lower boundary. If the subject identifies the lower boundary as now being longer than 1000ms it will be removed from the boundary list and the previous lower boundary reinstated. This widens the range and reorients it, hopefully, with the actual matched estimate within it.
5. This process (iteratively halving the range and testing its boundaries) continues until two pre-selected criteria are met:
   - six reversals have occurred (alternating confirmations that the matched estimate is 'shorter' than the upper boundary and 'longer' than the lower boundary) and,
   - the final distance between upper and lower boundaries is less than 5% of the original range (5% of 400ms = 20ms). Once these criteria are met the midpoint of the final range is taken as the matched estimate.
6. Three repetitions of this MOBS procedure are performed for each condition. This improves the accuracy of the procedure. The final matched estimate is the mean of these three estimates.

A Modified Binary Search procedure was chosen as the method to derive matched estimates as it was previously used by Yarrow et al (2001) in their original Chronostasis experiment and has been shown to be efficient and precise (Tyrrel and Owens, 1988). The procedure provides precise estimates of transient states (such as subjective duration estimates) with fewer stimulus presentations than conventional staircase methods[128]. It is also more robust when subjects are uncertain in their responses and incorrectly identify boundaries. Existing binary search techniques[129] would be unable to "backtrack" and recheck the boundaries (Tyrrel and Owens,

[128] A staircase method is similar to the MOBS technique but instead of halving the range after every trial the upper and lower boundaries are incrementally reduced by a set step size. This makes the reduction in range after each trial much smaller than MOBS and, therefore, requires more trials to reach an estimate.

[129] Traditional binary search techniques halve the range after each trial.

1988). In this study, MOBS proved very quick and reliable in producing matched estimates within 26 presentations of each condition. If the subject was presented a condition 26 times and the two criteria for termination had not been met, it was shown during piloting that a matched estimate would not be found as the subject's responses were too inconsistent. In this instance the MOBS was aborted and a null matched estimate recorded.

6.2.3 Experimental Details

6.2.3.1 Subjects

20 Edinburgh University students (10 female; aged 19-33 years, mean of 24 years) took part in this study. Subjects who admitted to have being diagnosed with an attentional disorder such as attentional defect hyperactivity disorder (ADHD) were excluded from the study. All subjects had normal or corrected to normal vision. Subjects were recruited from throughout the university and paid £5 for their time. Subjects signed a consent form prior to the experiment acknowledging there understanding of goals of the experiment, their voluntary participation, and their option to withdraw at any point. The experiment was designed according to the British Psychological Society's ethics guidelines.

6.2.3.2 Apparatus

The experiment was presented on a 19" Dell M992 CRT monitor with a resolution of 800 x 600 pixels and a refresh rate of 85Hz. This was connected to an Intel Pentium 4 3GHz PC with 512MB RAM and a Matrox Millennium G550 display adapter. The PC operating system was Microsoft Windows XP. A Canon Powershot S30 digital camera was used to produce the photographic stimuli. The initial image resolution was 1024 x 768 pixels. This was later reduced to the experiment resolution of

800x600 pixels[130]. Macromedia Director (Macromedia Director MX 2004, 2004) was used to build the experimental software and generate the stimuli.

6.2.3.3 Stimuli

The experimental stimuli used in this study consisted of a series of letters (D, B, E, R, F, and P) presented against a variety of backgrounds. The backgrounds were either black or one of two photographs of a garden scene. The same photographs were used throughout the whole experiment to ensure that the level of peripheral distraction was kept constant[131]. The letters were printed in bold white on to a 3D grey sphere (see examples in Figure 6-7). This ensured that the clarity of the letters was always a constant. If the letters had been placed against the different backgrounds without the grey surround the relative contrast would have varied. The grey sphere was positioned along the centre line of the screen which, when presented against a photographic background, made it appear as if it sat on a wall within the scene. The lighting within the scene was projected on to the sphere so that it looked like it was part of the scene and didn't visually 'jar' with its surroundings.

Stimuli were generated during runtime by the experimental software. Macromedia Director (Macromedia Director MX 2004, 2004) was used to construct the entire experiment and within it a 3D scene was constructed into which a variety of photographic backgrounds and a 3D sphere could be placed. The backgrounds, the "skin" of the sphere[132] and its location within the scene could be changed instantaneously. This allowed the software to quickly and precisely generate the

---

[130] The experimental software required smaller resolution images so that it could maintain the precise millisecond control of changes between images. During prototyping, the efficiency of the experiment was optimised so that the lag time between the sending of the command for an image to change and the image's actual appearance was <5ms. This would increase as demands on the systems CPU increased so close control over background applications had to be made during the experiment.

[131] It was accepted that the repeated presentation of the same background might lead the subject to become habituated to it decreasing the degree of peripheral distraction over the course of the experiment. To limit this effect conditions were presented in a random order half of the time .

[132] The "skin" refers to a bitmap image depicting the desired white letter against a grey background which is stretched around the sphere. Without a skin the sphere would be transparent.

stimuli required by each experimental condition whilst varying the presentation duration of the target letter according to the MOBS procedure.

As in the matched-exit/entrance experiment, the experimental software also controlled the presentation order of experimental conditions, recorded all subject input, system events, pre-processed the data and output it as tab-delimited files ready for statistical analysis. On top of this it also allowed the rules of the MOBS to be explicitly formalised and the rich visual stimuli generated accordingly. This would not have been possible if the stimuli had to be pre-rendered before the experiment.

6.2.3.4 Presentation Specification

The experiment was presented at a resolution of 800x600 (with an actual screen width of 36.3cm) and viewed at a constant distance of 33cm. This viewing distance was controlled using a chinrest. At this distance the screen subtended a visual angle of 57.30°. The sphere had a diameter of 150 pixels (11.70°) and each letter a width of about 50 pixels (3.91°). The sphere always began a trial 133 pixels from the left screen edge. A 20° relocation of the sphere positioned it at the centre of the screen whilst a 40° relocation positioned it 133 pixels from the right screen edge. These distances were always the same irrespective of degree of peripheral distraction. The stimuli were presented at a rate of 85 Hz to match the display refresh rate. This meant the screen was updated every 11.67ms. This created a potential discrepancy between the presentation durations suggested by the MOBS and the actual durations achieved. To compensate for this, actual presentation durations were recorded during runtime and used to update the MOBS boundaries. In this way the matched estimates generated by the MOBS were an accurate representation of the actual presentation durations[133].

[133] For example, the MOBS might chose to present a duration of 950ms but due to the refresh rate of the monitor and the time taken to generate the animation the actual duration presented might be 955ms (the software was designed to ensure that the actual duration never deviated more than 5ms either side of the intended value). Instead of updating the MOBS boundaries with the 950ms value the actual value presented is used (955ms) as this is what the subject responded to.

6.2.3.5 Procedure

The experiment was split across two sessions separated by exactly 24 hours. This introduced another within-subjects independent variable which will be referred to as Session that has two levels: 1st and 2nd. The two sessions were used to present each subject both levels of the 'Blocking' variable: Blocked and Random (see section 6.2.1). During one session experimental conditions were presented in blocks within which the condition remained the same until the MOBS terminated. In the other session the presentation order of conditions was randomised so that subjects could not predict what was going to happen at the cut (whether the object would relocate or the background change). These two presentation conditions were intended to highlight the effects of expectancy on subjective durations. The order of these two sessions was counterbalanced across subjects. A period of 24 hours between the two sessions was chosen as it ensured that time of day was consistent across sessions, all subjects had the same opportunity for rest between the sessions, and any techniques they had developed during the first session would also be used in the second.

At the start of each session demographic information was recorded: age, gender, handedness. The session then began with a series of instructions that described their task and the two keys they were to use to identify the target letter as being longer or shorter than the comparison durations. These keys were "S" for shorter and "L" for longer on a UK QWERTY keyboard. For half the subjects the left/right relationship of the keys was reversed. This was intended to counteract any handedness bias[134]. Subjects were instructed to respond as quickly as possible after the target duration had ended but to wait for later comparison durations if they were unsure of a response. The next trial would not be presented until a response was made.

Subjects were first given there randomly selected practice trials. After practice the main experiment would begin. If the session's conditions were blocked each block of

---

[134] As it happened all subjects used in this study were right handed and the results indicated that the positioning of the keys had no effect on matched estimates (see 6.3.1.5).

the experiment would begin with an example of the experimental condition of that block. This primed the subjects to expect the viewing conditions without having to "warm up" to them during the block. This was important as each response within the block was used by the MOBS to generate the matched estimate and inconsistent responses could jeopardise the MOBS accuracy. The block would continue presenting the same experimental condition until the MOBS terminates or had reached 26 trials.

If the session's conditions were to be randomised all nine experimental conditions would be presented within the same large block. Once a trial had been presented the subject's response was incorporated into the MOBS for that trial's experimental condition. The experimental condition presented in the next trial would then be randomly chosen from all MOBS left running. Once an experimental condition's MOBS terminated it would be removed from the list of conditions to be presented. Gradually all MOBS would terminate and the block would complete with a matched estimate for each experimental condition.

When all MOBS had terminated subjects would be forced to take a 3 minute break. Subjects were encouraged to move away from the screen and rest their eyes[135]. After the rest period the second repetition of all the experimental conditions began. This was identical to the first set. After all the MOBS of this set had terminated a second 3 minute rest period was enforced and then the final set of repetitions were presented. After the experiment had terminated subjects were asked to fill in a brief paper-based questionnaire and then paid £5 for their participation.

Each trial of the experiment lasted about 7 seconds with a variable rest period immediately afterwards[136]. The average duration of a MOBS was ~2.5 minutes. With 9 conditions in each set, a set would take, on average 22.5 minutes. With the two 3

---

[135] Due to the constant concentration required by the experiment and the close viewing condition enforced by the chin rest the experiment was very tiring on the eyes.

[136] In practice, subjects tended to speed past this rest period and continue on to the next trial, speeding up the experiment.

minute rest periods the entire length of the experiment should have been around 73.5 minutes. However, subjects could shorten this by making very rapid and consistent responses. The average duration of the actual experiment was around 50 minutes.

6.2.3.6 Dependent Variables

As well as recording the matched estimates for each repetition of the experimental conditions within the experiment various other data were also recorded.

The first extra piece of data recorded was the **total number of repetitions** required for each MOBS to terminate. This indicates the consistency of subjects' duration judgements: if they are very inconsistent the boundaries will repeatedly need to be checked which will increase the number of repetitions required to settle on a matched estimate.

The second set of data recorded was response data for every target duration tested during an experimental condition's MOBS. For every duration tested during a MOBS two pieces of information were recorded: how long after the target letter finished being presented the subject made a response (**response time**), and the type of response (**response type**: 0= "shorter", 1="longer"). These data were placed in bins of similar presentation durations. There were 65 bins covering 100-1800ms, decreasing in size as the durations they represented approached 1000ms (see Appendix C for full details). Each experimental condition had three sets of bins. The first set represented response times within each bin averaged across all three MOBS run for that experimental condition. The second set represented the average response types within each bin. The third set represented the total number of presentations made within each bin across all three MOBS.

The intention of these data was to map responses across the entire range of presentation durations to detect any signs of indecision: blocks of durations that subjects were unable to consistently respond to or to which they took longer to respond. It was hoped that this data could be used to map a "zone of tolerance"

surrounding the matched estimate within which any duration would be perceived as being indistinguishable from 1000ms.

6.2.3.7 Questionnaire

After the subject had finished the experiment they were asked to fill in a quick paper-based questionnaire. The questionnaire was intended to detect any strategies the subjects were using which might have biased their results. Most questions required a short unconstrained written answer. These responses were later grouped together and coded in a way that permitted statistical analysis.

Question 1: "How difficult did you find the experiment?"

Subjects responded to this question using a 5-point Likert scale with "Very Difficult" as 1 and "Very Easy" as 5. If the subject was answering the questionnaire after the second session they were also asked to compare the difficulty of this session to the first session. Their answers were made on a 3-point scale: "Easier" (1), "Same" (2), "Harder" (3).

The rest of the questions required subjects to write a short unconstrained answer. The intention and relationship between the questions and the expected effects of this study should be clear.

Question 2: Was there anything about the task you found particularly difficult?

Question 3: Did any of the animations make the task easier or harder? Just describe the animations as you remember them.

Question 4: Do you think there were any reasons why you personally found the experiment more difficult than it might have been for other people e.g. tiredness, poor eyesight, time of day, lack of concentration, etc?

Question 5: Did you use any strategies to help you with the task?

Question 6: Did you develop a systematic pattern of answers or think that you figured out the way the experiment chose to change the durations each trial? What was it?

Question 7: Did you press 'Longer' more than 'Shorter' or vice versa? Was there a reason for this?

Question 8: Did you find it hard to remember which key was which?

6.3 Experiment 2: Results

The primary type of data generated by this study was match estimates: the duration of the target letter 'E' perceived by subjects as being equal to 1000ms. Other data, such as the number of repetitions required for each MOBS to terminate, reaction times, and response types for each trial, as well as questionnaire answers, were also recorded. However, matched estimates will be the primary data used during this analysis. The secondary types of data will only be used when deemed suitable.

Analysis of the results will begin with general statistics for the matched estimates averaged across all 9 experimental conditions. This will allow the detection of any general trends in the data before closer, more directed analysis is performed for each hypothesis. The affect of individual differences on matched estimates will also be analysed. These general statistics will then be followed by specific analyses of the hypotheses identified in sections 6.1.2 and 6.2.1 and their associated experimental conditions.

6.3.1 General Statistics

6.3.1.1 Matched Estimates

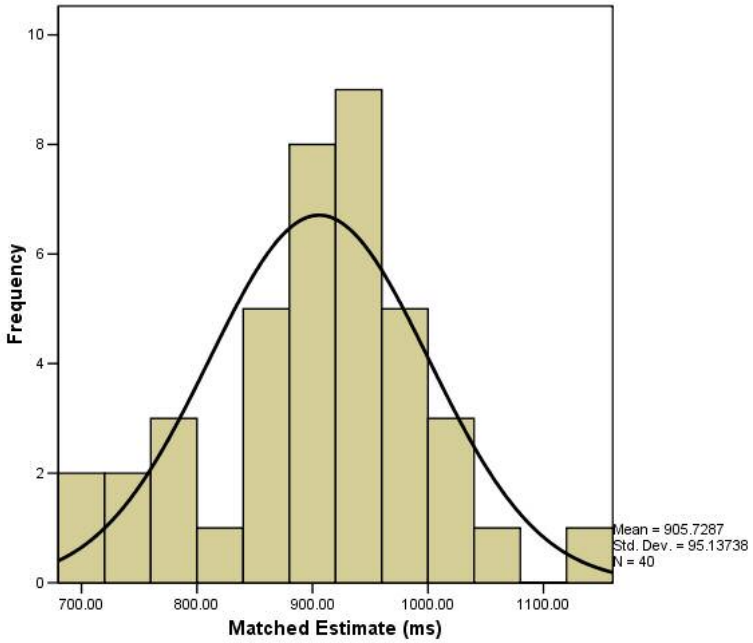


Figure 6-8: Distribution of matched estimates (ms) across all subjects and conditions. Curve represents the normal distribution.

The experimental design identified presentation durations at which a target letter 'E' was perceived as being equal to 1000ms (i.e. matched estimates). If subjects were able to perform this task accurately the average matched estimate across all viewing conditions should be equal to 1000ms. As can be seen from the distribution of matched estimates depicted in Figure 6-8, the mode matched estimate does not lie at 1000ms. The mean matched estimate is 905.73ms (min. 696.56ms, max. 1146.15ms, standard deviation 95.14). This is significantly lower than the expected value of 1000ms (One-sample t-test, t=-6.267, df=39, p<.000). This indicates that, as predicted by the hypotheses, the viewing conditions affect the matched estimate. The precise cause of this deviation will be investigated over the course of this analysis.

The average matched estimates are distributed normally across the entire subject set (the distribution is not significantly skewed: skewness value, -0.174, is not more than

twice the standard error skew, 0.374). This means that parametric tests can be used when analysing the matched estimates.

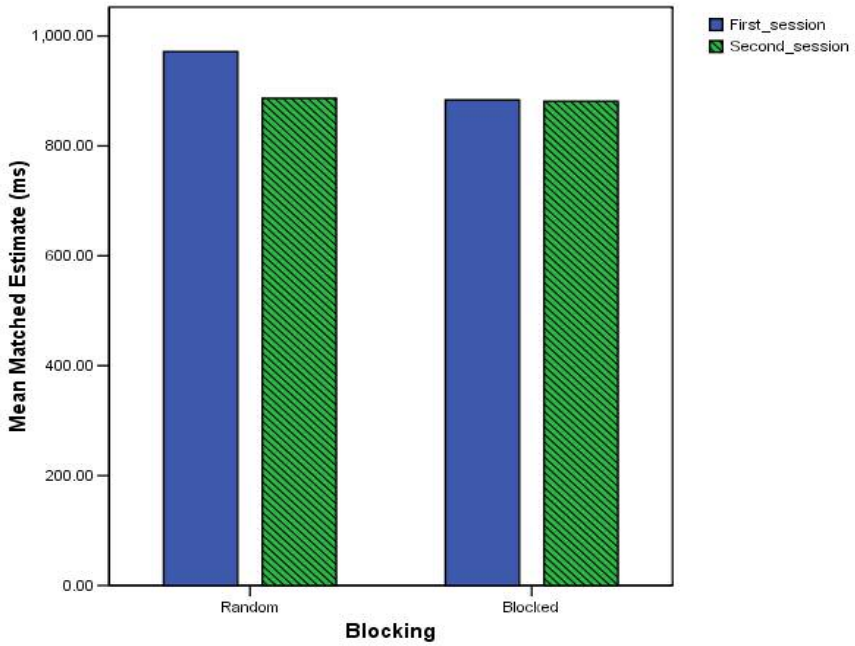


Figure 6-9: Mean matched estimates (ms) across the four presentation conditions.

The significant deviation of the mean matched estimate from the expected value (1000ms) remains if the data is split according to Blocking (random presentation of experimental conditions: t=-3.467, df=19, p<.01; or blocked: t=-5.22, df=19, p<.000) and Session (first day: t=-3.138, df=19, p<.01; second day: t=-6.243, df=19, p<.000). However, when the matched estimates are split into the four types of experimental session (Random-1st, Blocked-2nd, Blocked-1st, Random-2nd; see Figure 6-9) we find different mean matched estimates. If experimental conditions are randomised in the first session (left solid-blue bar in Figure 6-9), matched estimates do not differ significantly from 1000ms (mean=971.52ms: t=-1.091, df=0, p=.304, not sig.). By comparison, blocking in either the first (left striped-green bar in Figure 6-9: mean=883.64: t=-3.449, df=9, p<.01) or second session (right striped-green bar in Figure 6-9: mean=886.64: t=-4.32, df=9, p<.01) leads to significantly shorter matched estimates as does random in the second session (right solid-blue bar in Figure 6-9: mean=881.11: t=-4.280, df=9, p<.01). This indicates a large influence of presentation order on the matched estimates. This will be investigated in more detail

under the Concentration and Divided Attention Hypotheses (section 6.3.2.1). This difference between presentation orders must be considered when analysing the influence of any other experimental conditions.

6.3.1.2 Individual Differences

6.3.1.3 Gender Differences

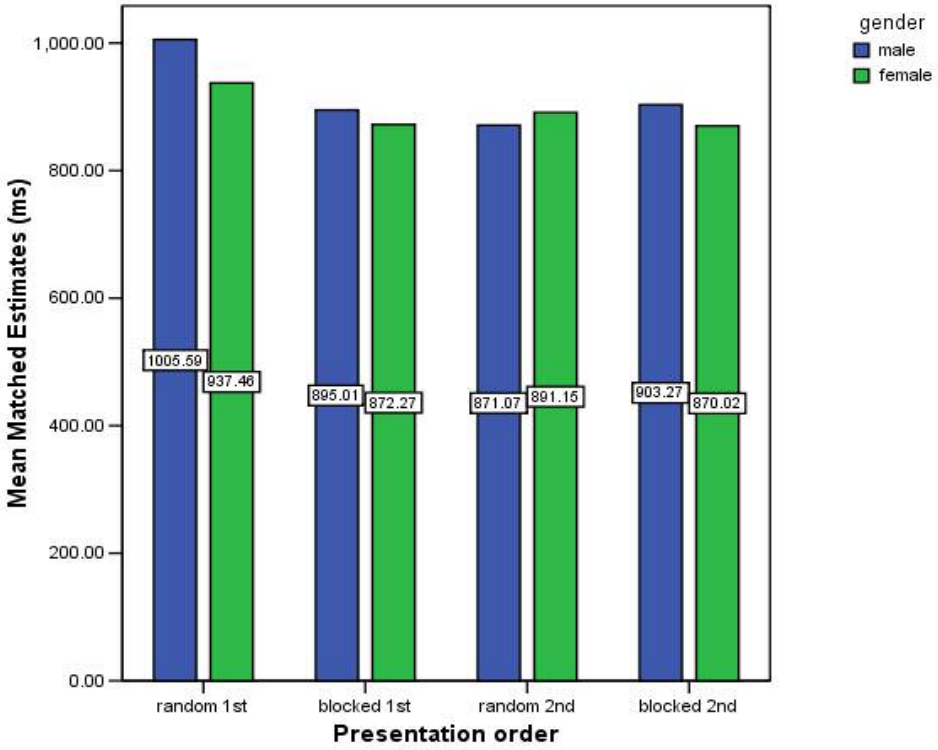


Figure 6-10: Gender differences (left bar=male, right bar=female) in matched estimates (ms) across the four presentation conditions.

The number of males and females was balanced in this study (10 each). Averaged across all conditions, Females have lower mean matched estimates than males (892.72ms compared to 918.74ms). However, this difference is not significant (t=.862, df=38, p=.394, not sig). As can be seen from Figure 6-10, males produce higher matched estimates than females when conditions are randomised in the first and second sessions and blocked in the first session. Females produce higher matched estimates when conditions are randomised in the second session. However,

none of these differences are significantly different[137]. Therefore, males and females can be grouped together for the rest for the rest of the analysis.

6.3.1.4 Age Differences

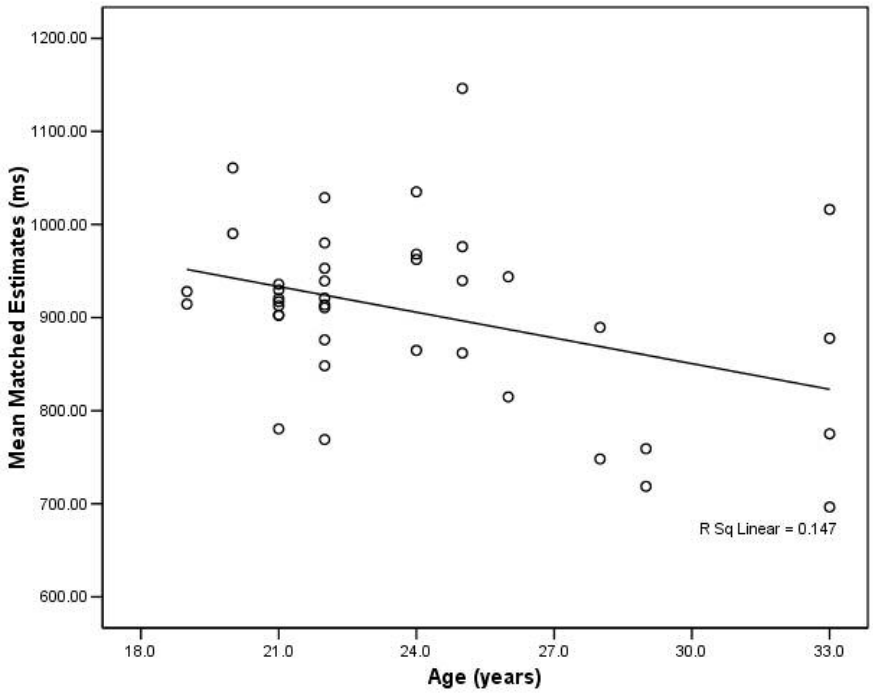


Figure 6-11: Correlation between mean matched estimates (ms) and age (years).

The mean age of participants was 24 and ranged from 19 to 33 years. Figure 6-11 illustrates the relationship between age and matched estimates. There appears to be a trend for older subjects to produce longer matched estimates compared to younger subjects. This would suggest that age is positively related to the perceived duration of visual events. Such a relationship would result in a negative correlation between matched estimates and age. Performing a Pearson's Correlation between age and mean matched estimates (Figure 6-11) shows a weak negative correlation ($r=-0.383$, $p<.05$). However, given the small number of subjects and the uneven distribution of ages (most subjects were aged between 21 and 26) more subjects across a broader range of ages would be needed to place much confidence in this correlation. No

[137] Independent samples t-tests performed between gender groups split across four presentation conditions. All analyses failed to show significance at the $p<.05$ level.

effect of presentation order on this correlation can be found as the number of subjects is too few (only 10 per order spread across the age range).

6.3.1.5 Response Key Placement and Handedness

The MOBS required subjects to choose either 'Shorter' or 'Longer' after the letter 'E' was presented. One of these keys was pressed with the left hand and the other the right. To protect against a handedness bias the left-right relationship of the keys was balanced across subjects. If subjects used their favoured hand more frequently this would mean that half of right-handed subjects would press 'Longer' more often and half 'Shorter'. When the matched estimates from these two groups were averaged together any handedness bias would be eliminated.

To investigate if such a bias existed, matched estimates were split across the two forms of key placement. When the 'Longer' key was pressed with the right hand estimates were larger (916.98ms) than when the 'Longer' key was on the left (894.48ms). This difference is not significant ($t=.744$, $df=38$, $p=0.462$, not sig.). The slight bias towards longer estimates when the 'Longer' key is on the right cannot be attributed to handedness bias as all of the subjects used in this study were right handed. If subjects were pressing the 'Longer' key most frequently because it was positioned under their favoured hand this should have pushed matched estimates lower, not higher as was seen.

The other dependent variable useful for judging subject preference for key placement is the total number of repetitions. This is the number of times each MOBS presented a condition before a matched estimate was found. The greater the number of repetitions the more inconsistent the subject's responses. Looking at the number of repetitions for the two key placements, when the 'Longer' key was on the left significantly more repetitions were required compared to when 'Longer' was on the right (16.83 vs. 15.59: $t=-2.426$, $df=38$, $p<.05$). This is consistent with subject's informal feedback stating that they found it odd when 'Longer' was on the left as they expected 'Shorter' to come before 'Longer' when working left-to-right across

the keyboard. However, in response to the 'Keys' question on the questionnaire (question 7), 25% of subjects said that they pressed 'Longer' most often when it was on the right and 45% when it was on the left. The majority of the remaining subjects said they pressed both keys equally. This suggests that there was a preference for pressing 'Longer' more than 'Shorter' irrespective of its position but that subjects found it easier to respond consistently when 'Longer' was on the right. Fortunately, the absence of a significant difference between matched estimates for the two key placements indicates that MOBS was able to compensate for this bias and produce accurate matched estimates.

6.3.2 Hypotheses

The analysis of individual differences indicates that gender, age, and key placement have no effect on matched estimates. As such, all subjects will be grouped together in subsequent analyses. The general statistics indicated that the overall matched estimate is significantly lower than 1000ms and this appears to be due to differences between Blocking and Session groups. These differences will be investigated in the next section.

6.3.2.1 Concentration and Divided Attention Hypotheses

Before investigating the affect on matches estimates of each experimental condition the affect of presentation condition (Blocking and Session) need to be identified. The predicted effect of Blocking has previously been formalised in two hypotheses:

Hypothesis 1:

> Concentration Hypothesis: When editing conditions are predictable matched estimates will be significantly shorter than the actual value.

Hypothesis 2:

> Divided Attention Hypothesis: When editing conditions are unpredictable matched estimates will be significantly longer than the actual value.

These two hypotheses predict how the two presentation conditions, Blocked (Hypothesis 1) and Random (Hypothesis 2), affect matched estimates under the control condition (condition 1; see Figure 6-12). We have already seen in section 6.3.1.1 that matched estimates are significantly smaller than 1000ms under both blocked and random conditions when averaged across all experimental conditions. It has also been shown that this difference changes depending on the session (see Figure 6-9). However, as these matched estimates are averaged across all experimental conditions the effect may not just be due to presentation order. To remove all other confabulating factors (Saccade and Background) we need to look just at the matched estimates under condition 1.

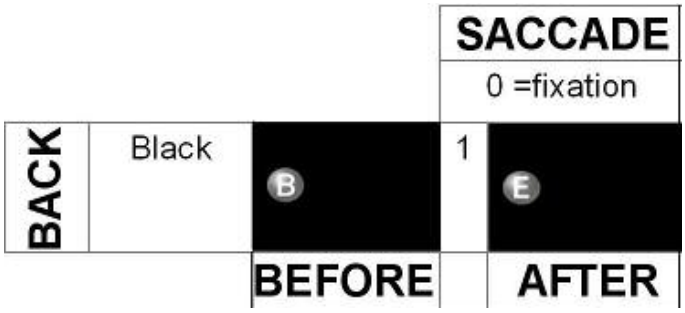


Figure 6-12: The Control condition (condition 1). The onset of the target letter 'E' isn't accompanied by a target relocation or background change.

Across both sessions, the mean matched estimates for condition 1 when blocked is 901.66ms (s.d.=97.77) and when presented randomised, 952.62ms (s.d.=103.62). The blocked estimate is significantly smaller than 1000ms ($t=-4.498$, $df=19$, $p<.000$) where as the random estimate is not ($t=-2.045$, $df=19$, $p=.055$, not sig. to $p<.05$). These results support hypothesis 1 but not hypothesis 2. Estimates under random presentation conditions were expected to be longer than 1000ms which is clearly not the case given these results. In fact, the absence of significant deviation from 1000ms seems to suggest that the random presentation conditions actually aids the subject in making an accurate estimate. These matched estimates also differ significantly from each other ($t=-2.562$, $df=19$, $p<.05$).

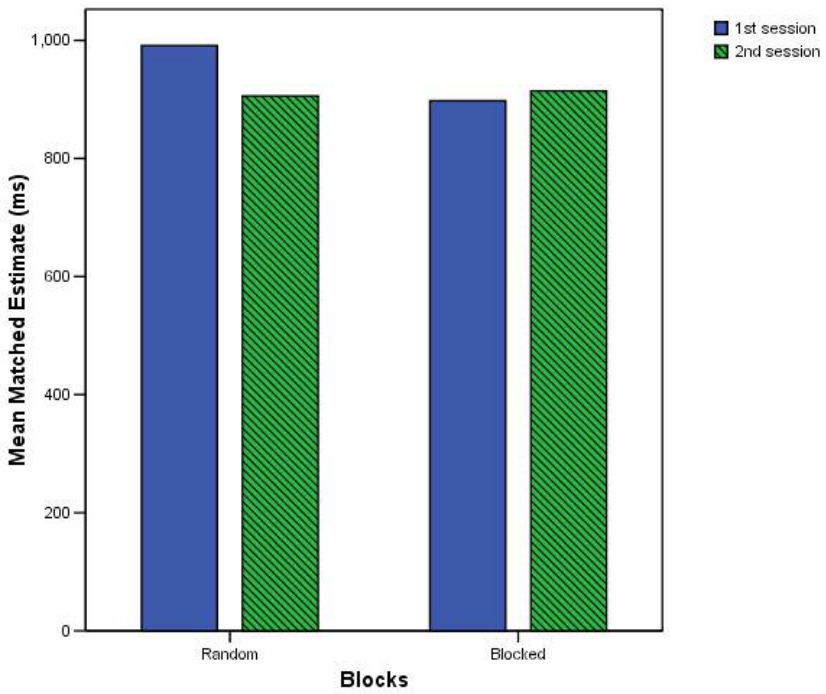


Figure 6-13: Mean matched estimates (ms) for condition 1 split across the two Blocking and Sessions conditions: left solid-blue bar= Random 1st, left striped-green=Random 2nd, right solid-blue=Blocked 1st, right striped-green = Blocked 2nd.

Figure 6-13 shows the mean matched estimates for condition 1 split across the four presentation conditions (left solid-blue bar= Random 1st, left striped-green=Random 2nd, right solid-blue=Blocked 1st, right striped-green = Blocked 2nd). When condition 1 is presented randomly in the first session matched estimates are very close to 1000ms (Random 1st: mean = 990.98ms, s.d.=111.03: t=-.257, df=9, p=.803, not sig. to p<.05). By comparison, blocking in the first session produces matched estimates significantly lower than 1000ms (Blocked 1st: mean=897.60ms, s.d.=104.61: t=-3.096, df=9, p<.05). This difference between the random and blocking groups in this session is significant (t=1.936, df=18, p<.05, on-tailed). This again supports hypothesis 1 and, whilst the random condition does not produce estimates significantly larger than 1000ms, they are significantly larger than in the blocked condition which does indicate a trend in the direction predicted by hypothesis 2.

This advantage due to random presentation disappears when used in the second session (Random 2nd: mean=914.25ms, s.d.=84.08: t=-3.225, df=9, p<.01 compared to 1000ms). The matched estimate drops down to a similar level as the blocked group

(Blocked 2nd: mean=905.72 s.d.=95.93: t=-3.108, df=9, p<.05) and there is no significant difference between them (t=.212, df=18, p= .835, not sig. to <.05). Given that there is a clear advantage of randomising the experimental conditions in the first session it seems odd that the same advantage would not continue into the second session. The difference between the two random estimates is large and only narrowly escapes being significant (t=1.742, df=18, p=.099, not sig. to .05 two-tailed). At this stage of the analysis no explanation presents itself for why the random presentation loses its accuracy in the second session. It is possible that an examination of the other experimental conditions may provide an explanation.

This analysis of matched estimates across the four presentation groups indicates that both Blocking and Session affect matched estimates. Given the purity of the time estimation task under condition 1 (i.e. no distractions) the matched estimates produced under this condition should be used as a baseline comparison for all subsequent analyses. 1000ms cannot be used as the baseline as, under some Blocking conditions and Sessions, subjects do not perceive the letter 'E' presented for 1000ms as being equal to the comparison letters presented for 1000ms.

6.3.2.2 Chronostasis Hypothesis

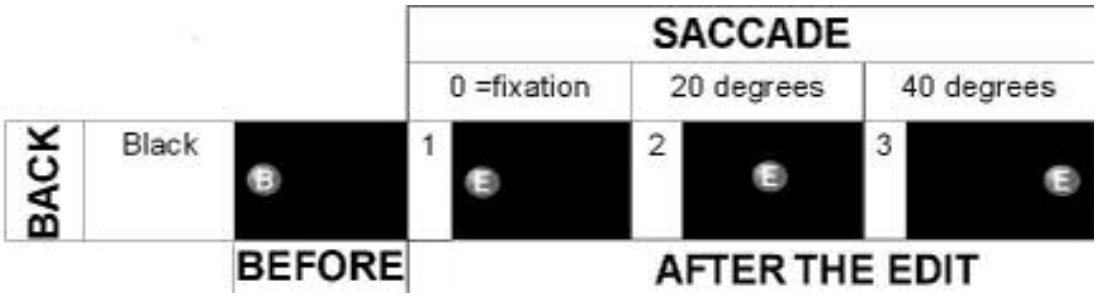


Figure 6-14: The experimental conditions (1, 2, and 3) used to test the effect of Saccade on matched estimates (i.e. the Chronostasis Hypothesis).

In section 6.1.2.3 a type of cut, referred to as Stop-Motion, was identified in which the only change occurring to the screen across the cut was the relocation of the focal object. This relocation is believed to capture attention and result in an involuntary saccadic eye movement. It was predicted that this saccade would result in the perceptual extension of the saccade's duration (the Chronostasis effect; (Yarrow et

al., 2001b). A hypothesis was formulated to test this prediction: **"***When an object suddenly relocates across the screen at the beginning of a visual event, the event will be perceived as ~50ms longer than the same event viewed under fixation. Whether the relocation is expected or not, the perceived extension of duration will be the same"*. This hypothesis can now be reformulated with respect to the current experimental design:

Hypothesis 3:

> Chronostasis Hypothesis: Duration estimates will be significantly shorter (~50ms) than the Control (Condition 1) when accompanied by eye movements (conditions 2 and 3). There will be no significant difference between the lengths of the eye movements or Blocking.

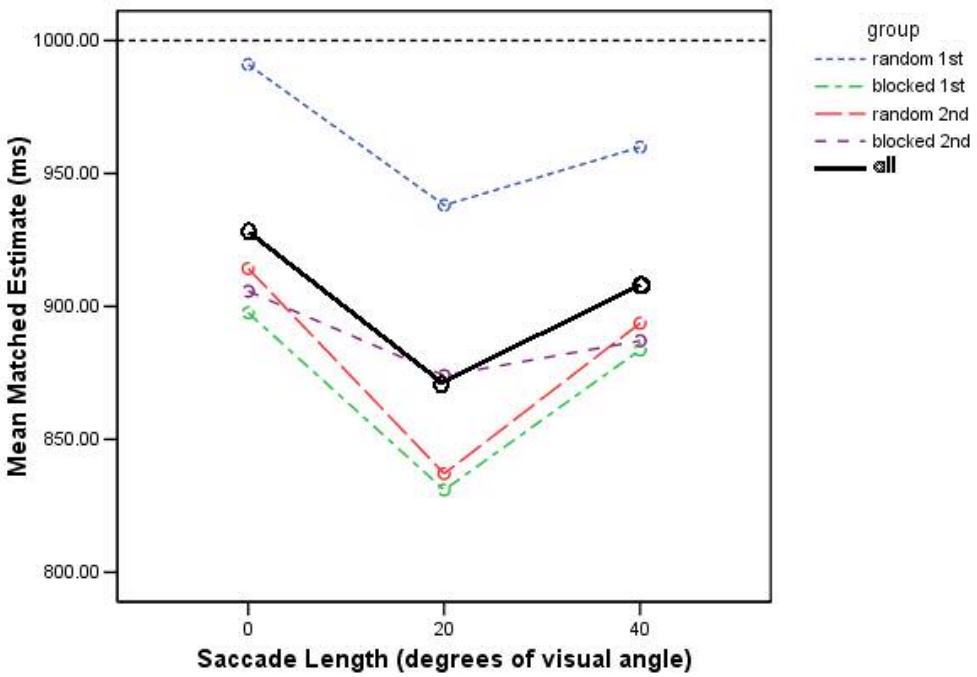


Figure 6-15: Mean matched estimates for the three saccade lengths (0°=condition 1, 20°=condition 2, 40°=condition 3). Black line represents the average across all presentations conditions. The other lines represent the four presentation conditions.

Across all presentation conditions (solid black line in Figure 6-15) there was a main effect of saccade length (Repeated-measures ANOVA with Greenhouse-Geisser correction: F=11.959, df=1.922, p<.000). The mean matched estimate for condition 1 was 927.14ms, condition 2 was 870.03ms, and condition 3 was 906.07ms. The

57.11ms average decrease in matched estimates between 0 and 20° saccades is exactly as was predicted by hypothesis 3 and in agreement with the degree of Chronostasis previously reported (Yarrow et al., 2001a). This decrease in matched estimates is significant (pair-wise comparison: $p<.000$, two-tailed). The difference between matched estimates for 0° and 40° saccades is still significant ($p<.05$, one-tailed) but the decrease is smaller than expected (mean difference=21.07ms). This also means that estimates for 40° saccades are significantly higher than for 20° saccades ($p<.01$, two-tailed) which does not support hypothesis 3.

When the estimates are split according to presentation conditions (Blocking + Session) the same decrease in estimates with saccades can be seen (see Figure 6-15). When conditions 1, 2, and 3 are presented randomly in the first session (dotted blue line in Figure 6-15) there is no main effect of Saccade (GG: $F=2.173$, $df=1.906$, $p=.146$, not sig. to .05 one-tailed) but 20° does have significantly shorter matched estimates than 0° (mean diff.=52.93ms, $p<.05$, one-tailed). No other differences are significant.

When conditions are blocked in the first session (green dot-dash line in Figure 6-15) there is a main effect of Saccade (GG: $F=3.937$, $df=1.831$, $p<.05$) which can be attributed to 20° having significantly shorter matched estimates compared to 0° (mean diff.=-66.63ms, $p<.05$) and 40° (mean diff=-52.69, $p<.05$, one-tailed). No other differences are significant.

When conditions are presented randomly in the second session (dashed red line in Figure 6-15) there is a main effect of Saccade (GG: $F=5.359$, $df=1.679$, $p<.05$) which is due to 20° Saccade producing significantly shorter matched estimates compared to 0° (mean diff.=-77.15ms, $p<.05$) and 40° (men diff.=-56.57ms, $p<.05$). No other differences are significant.

Finally, when conditions are blocked in the second session (small dash purple line in Figure 6-15) there is no main effect of Saccade (GG: $F=1.423$, $df=1.723$, $p=0.268$,

not sig. to .05) and no significant differences between matched estimates for Saccade conditions.

This analysis shows that under most presentation conditions (random 1st, blocked 1st, and random 2nd) durations containing 20° saccades lead to matched estimates significantly shorter than Control just as predicted in hypothesis 3. However, hypothesis 3 also predicted that there would be a significant decrease for 40° saccades and no significant difference between estimates for 20° and 40°. None of the presentation conditions show a significant decrease for 40° saccades and only random 1st and blocked 2nd show no difference between 20° and 40°. It is interesting to note that these two presentation conditions refer to the same subjects. This suggests that the order in which the presentation conditions were used across the sessions might be important.

To test whether subject group (random 1st and blocked 2nd vs. blocked 1st and random 2nd) affected matched estimates the data was split into the two subject groups. A repeated-measures ANOVA shows that there is no main effect of subject group (GG: F=2.048, df=1, p=.170, not sig to .05) and this is probably due to the matched estimates for blocked-2nd (first subject group; small dash purple line in Figure 6-15) being very similar to both blocked-1st (green dot-dash line in Figure 6-15) and random-2nd (dashed red line in Figure 6-15). Both of these presentation conditions come from the second subject group. Within the first subject group (random 1st and blocked 2nd) there is a main effect of Session (GG: F=5.814, df=1, p<.05) which isn't seen within the second subject group (blocked 1st and random 2nd: F=.225, df=1, p=.647, not sig.). This effect of Session indicates that the estimates from the two sessions cannot be averaged together. This was expected as random and blocked presentations should lead to different averages (hypotheses 1 and 2). However, the absence of this difference in the second subject group is highly unexpected (see section 6.3.2.1. for a previous example of this absence).

.

6.3.2.3 Resistance, Anticipation, and Distraction Hypotheses

In section 6.1.2.2 a type of cut, referred to as Jump Cut, was identified in which the only change occurring across the cut was the sudden change of background. It was predicted that this sudden change would result in attention being covertly distracted from the time estimation task (i.e. the focal-object being fixated). Such covert distraction has previously been shown to result in a perceived shortening of durations (Block et al., 1980; Zakay, 1992; Zakay, Nitzan, & Glicksohn, 1983a). Distraction only occurs when the change isn't expected (see chapter 3. These predictions were formalised as two hypotheses: "*Distraction Hypothesis: When the background changes unexpectedly whilst the focal object is fixated the subsequent visual duration will be perceived as shorter than the same duration without background change"* and *"Resistance to Change Hypothesis: An expected background change will have no effect on the perceived duration of a subsequent visual event"*. These hypotheses can now be reformulated with respect to the current experimental design:

Hypothesis 4:

> Resistance to Change Hypothesis 1: Matched estimates will not differ significantly from Control (condition 1) when the target is presented against a constant (condition 4) or changing background (condition 5) and the type of cut is predictable.

The Distraction Hypothesis must be split into two hypotheses as the experimental design introduces an extra no-Saccade viewing condition, condition 4 (constant Background; see Figure 6-16). As already seen in section 6.3.2.1, the unpredictable presentation condition causes subjects to divide their attention across the screen in anticipation of a change. The presentation of a background scene (condition 4) provides a higher degree of potential distraction for this divided attention compared to a black background (condition 1). Therefore, it is predicted that matched estimates will be slightly shorter than the Control when a constant background is presented and the viewing conditions are unpredictable.

Hypothesis 5:

> Anticipation of Change Hypothesis: When the target is presented against a constant background (condition 4) matched estimates will be significantly longer than without a background (condition 1) if the viewing conditions are unpredictable.

Hypothesis 6:

> Distraction Hypothesis: When the background changes unexpectedly (condition 5) matched estimates will be significantly longer than without a background (condition 1) or with a constant background (condition 4).

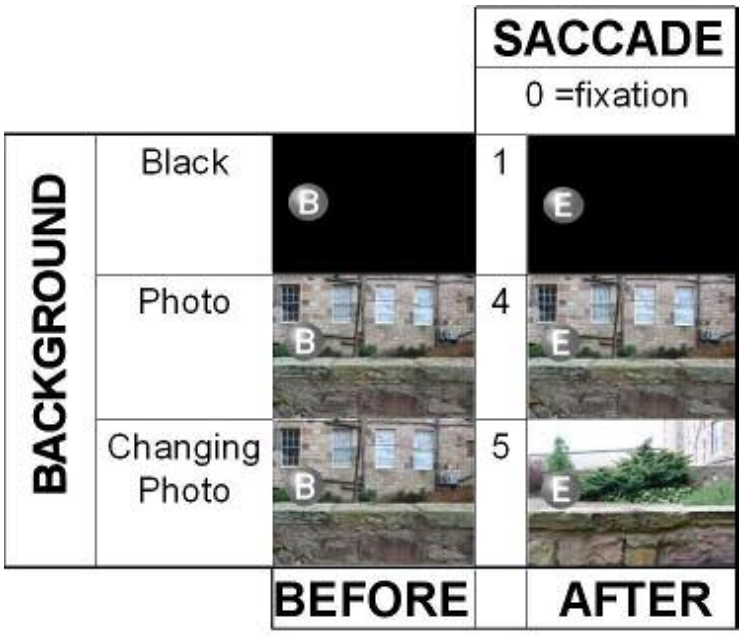


Figure 6-16: The experimental conditions (1, 4, and 5) used to test the effect of Background on matched estimates under fixation.

.

As can be seen from Figure 6-17 (next page; solid black line) when matched estimates are averaged across all subjects they are very similar: condition 1 = 927.14ms, condition 4 = 931.18ms, condition 5 = 934.81ms. There is a slight increase in estimates with background (condition 4) and change (condition 5) as predicted by hypotheses 5 and 6 but a repeated-measures ANOVA shows that this is not significant (GG: F=.349, df=1.383, p=.630, not sig. to .05). The same analysis indicates that there is a significant interaction between the Session in which the conditions are tested and Subject Group (GG: F=8.786, df=1, p<.01). This interaction

can be attributed to presentation condition random-1st having significantly higher matched estimates than all other presentation conditions (see Figure 6-17, dotted blue line). This indicates, as suggested by hypotheses 4, 5, and 6, that there are different effects for each presentation condition.

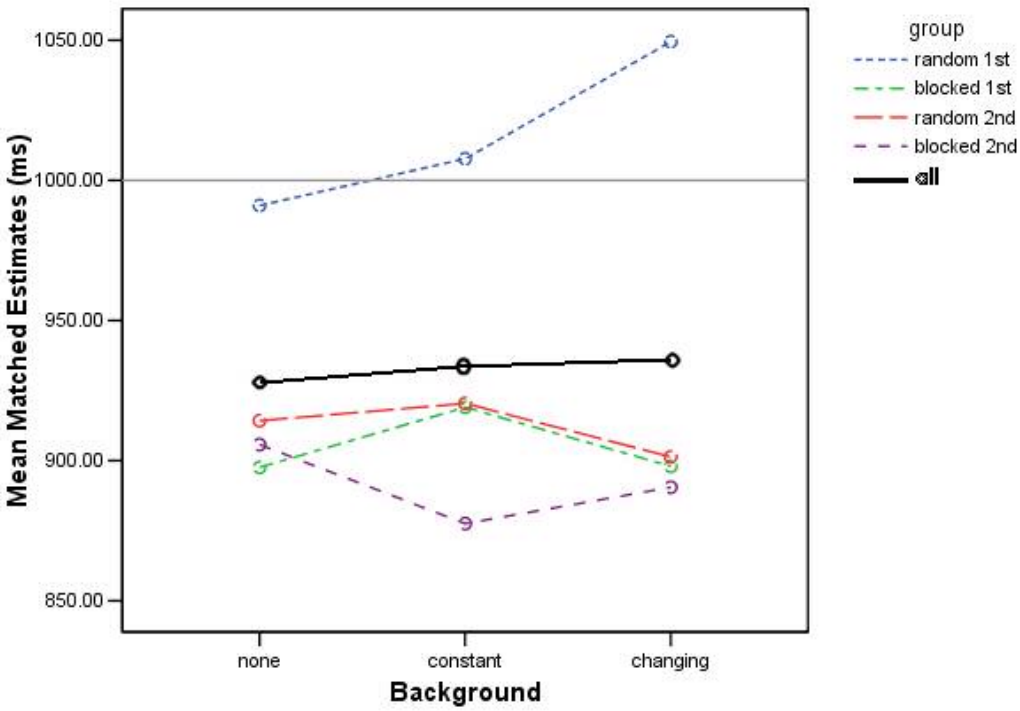


Figure 6-17: Mean matched estimates (y-axis: ms)for the three background conditions (x-axis: none, constant, and changing). Black line represents averages across all subjects. The other four lines represent averages for the four presentation conditions.

When conditions 1 (no background), 4 (constant) and 5 (change) are presented randomly in the first session (Figure 6-17, dotted blue line) there is a main effect of Background (GG: F=10.491, df=1.770, p<.01). This can be attributed to condition 5 (mean=1049.43ms) having significantly higher matched estimates than condition 1 (mean=990.99ms, p<.01) and condition 4 (mean=1007.70ms, p<.05). This supports hypothesis 6 as the changing background increases matched estimates as expected. However, these results do not support hypothesis 5 as condition 4 (constant background) is not significantly higher than condition 1 (no background: mean diff.=16.72ms, p=.160, not sig. to .05).

When the conditions are blocked in the first session there is no main effect of Background (GG: F=.520, df=1.871, p=.592 not sig to .05) and no pair-wise

differences between conditions: means = 897.60ms (condition 1), 919.05ms (condition 4), 897.97ms (condition 5). This confirms hypothesis 4.

When the conditions are randomised in the second session there is no main effect of Background (GG: F=.405, df=1.461, p=.613 not sig to .05) and no pair-wise differences between conditions: means = 914.25ms (condition 1), 920.40ms (condition 4), 901.33ms (condition 5). These results contradict hypotheses 5 and 6 as higher matched estimates were expected under Random viewing conditions.

When the conditions are blocked in the second session there is no main effect of Background (GG: F=1.829, df=1.620, p=.198 not sig to .05) and only one pair-wise difference: condition 1 has significantly higher matched estimates (mean = 905.72ms) compared to condition 4 (mean = 877.55ms, p<.05). The absence of a significant difference between conditions 4 and 5 (mean = 890.50ms) supports hypothesis 4 but the significantly higher matched estimates for condition 1 are highly unexpected.

6.3.2.4 Masking Hypothesis

In section 6.1.2.3 two hypotheses were proposed that predicted how the sudden relocation of the focal object across a cut might affect time perception. The first considered only the effect of the associated saccadic eye (hypothesis 3, the Chronostasis Hypothesis) and has been analysed in section 6.3.2.2. To examine the 'pure' affect of saccadic eye movements on matched estimates only the no-Background conditions (conditions 1, 2, and 3) were used in this analysis. The second hypothesis from section 6.1.2.3 acknowledged that the Chronostasis effect might be modified by a masking effect. If the Stop-motion cut contained a background, the sudden uncovering of the background behind the relocated target might mask the last few milliseconds of the previous shot. In doing this, the masking would artificially extend the perceived duration of the new shot. This hypothesis was formalised as: *"When the relocation of a focal object at the beginning of a visual event reveals a detailed background, the perceived duration of the event will be*

longer than the same event presented without a background". In light of the current experimental design this can be reformulated as:

Hypothesis 7:

> Masking Hypothesis: When an unexpected target relocation occurs against a constant background (conditions 6 and 8) matched estimates will be significantly shorter than estimates when the same relocation occurs without any background (conditions 2 and 3).

No such effect is expected when the conditions are blocked as subjects expect the sudden uncovering of the background and resist masking by preparing the required saccade before the cut occurs.

Hypothesis 8:

> Resistance to Change Hypothesis 2: When an expected target relocation occurs against a constant background (conditions 6 and 8) matched estimates will not differ significantly from estimates made under similar target relocations without a background (conditions 2 and 3).

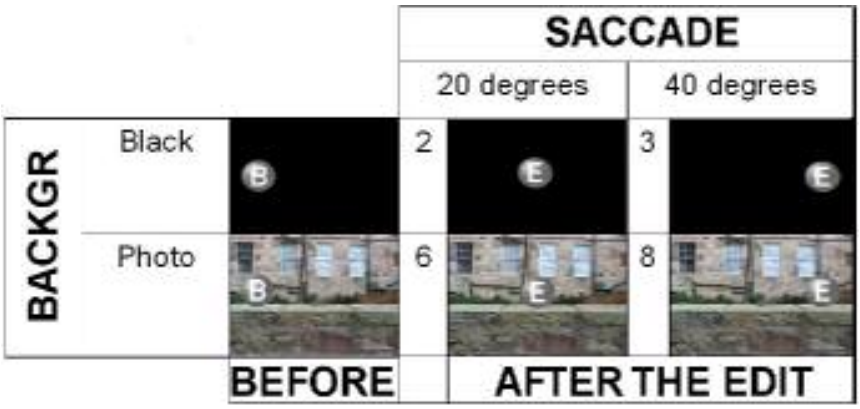


Figure 6-18: Experimental conditions used to investigate the Masking Hypothesis. Matched estimates for saccades performed against backgrounds (conditions 6 and 8) are compared to the same saccades without a background (conditions 1 and 2).

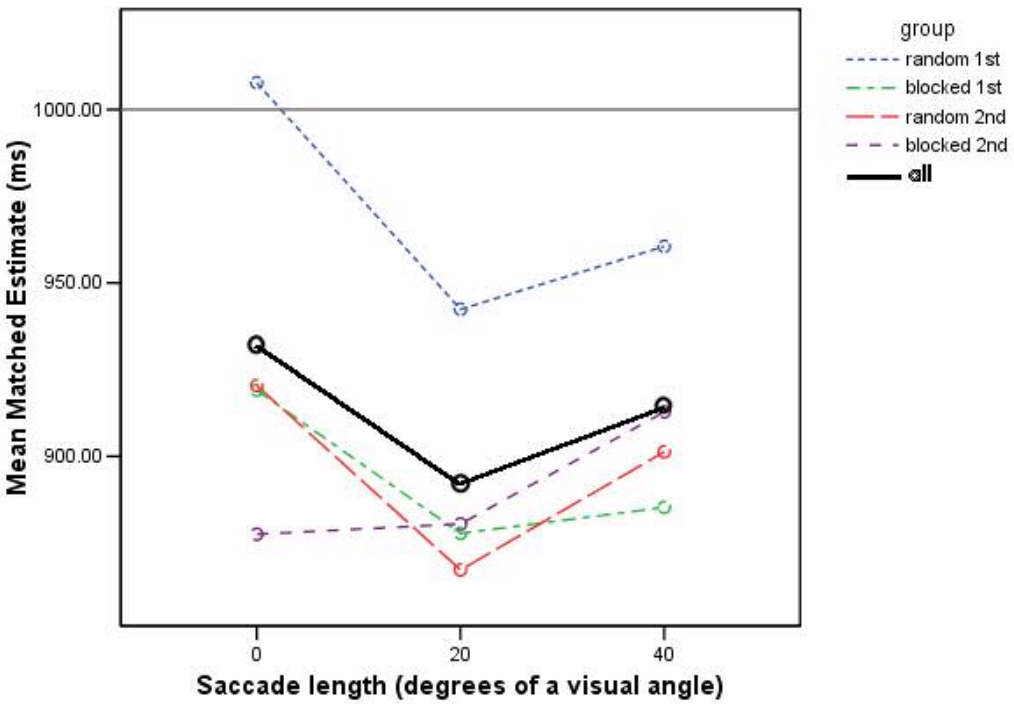


Figure 6-19: Mean matched estimates (ms) for the three saccade lengths performed against a constant background (0°=condition 4, 20°=condition 6, 40°=condition 8). Solid black line represents the average across all subjects. The four other lines represent estimates within each presentation condition.

To test the Masking Hypothesis, the matched estimates for conditions under which a saccade was performed against a constant background (conditions 6 and 8) should be compared to conditions with the same saccade but without any background (conditions 2 and 3; see Figure 6-18). However, given that any masking effect is believed to function in combination with the Chronostasis effect a test to see if the Chronostasis effect still exists should be performed.

A repeated-measures ANOVA across conditions 4, 6 and 8 for all subjects (solid black line in Figure 6-19) shows a main effect of Saccade (GG: $F=4.043$, $df=1.582$, $p<.05$) and this can mostly be attributed to condition 6 (20° saccade, mean=892ms) which has significantly shorter matched estimates than condition 4 (0°, mean=931.18ms, $p<.05$) and condition 8 (40°, mean=915ms, $p<.05$, one-tailed). This decrease in matched estimates after 20° saccade again confirms the Chronostasis effect (hypothesis 3) but 40° fails to do so (as was previously seen in 6.3.2.2).

The repeated-measures ANOVA shows no between-subjects effect of presentation group (F=1.694, df=3, p=.186, not sig. to .05) but random-1st does produce significantly higher matched estimates compared to blocked-1st (mean diff.=76.16ms, p<.05, one-tailed), random-2nd (mean diff.=73.88ms, p<.05, one-tailed) and blocked-2nd (mean diff.=79.92ms, p<.05, one-tailed). This again emphasises the importance of splitting the data into presentation groups.

Within the random-1st group there is a main effect of Saccade (GG: F=3.89, df=1.452, p<.05, one-tailed) and, just as predicted by hypothesis 3, this is due to **both** 20° and 40° saccades producing significantly lower matched estimates compared with Control: 0° mean =1007.70ms, 20° mean =942.40ms (sig. lower than 0°: p<.05, one-tailed), 40° mean = 960.53ms (sig. lower than 0°: p<.05, one-tailed). When conditions are presented blocked in the first session there is no main effect of Saccade (GG: F=1.991, df=1.203, p=.187, not sig.) but the matched estimate for 40° saccades is significantly shorter than 0° as predicted by hypothesis 3 (mean diff=33.75, p<.01). When the conditions are presented randomly in the second session there is a main effect of Saccade (GG: F=3.011, df=1.865, p<.05, one-tailed) which can be attributed to 20° saccade producing significantly shorter matched estimates than 0° (mean diff.=53.133, p<.05). When conditions are blocked in the second session there is no main effect of Saccade (GG: F=.504, df=1.324, p=.542, not sig.) or pair-wise differences.

These results indicate that the Chronostasis effect exists for constant Background conditions (conditions 6, and 8) when presented randomly in the first session and there is some signs of a similar effect in blocked-1st and random-2nd as well. Next, these estimates need to be compared to estimates produced without any background under the same Saccade conditions (conditions 1, 2, and 3) to identify any affect of masking.

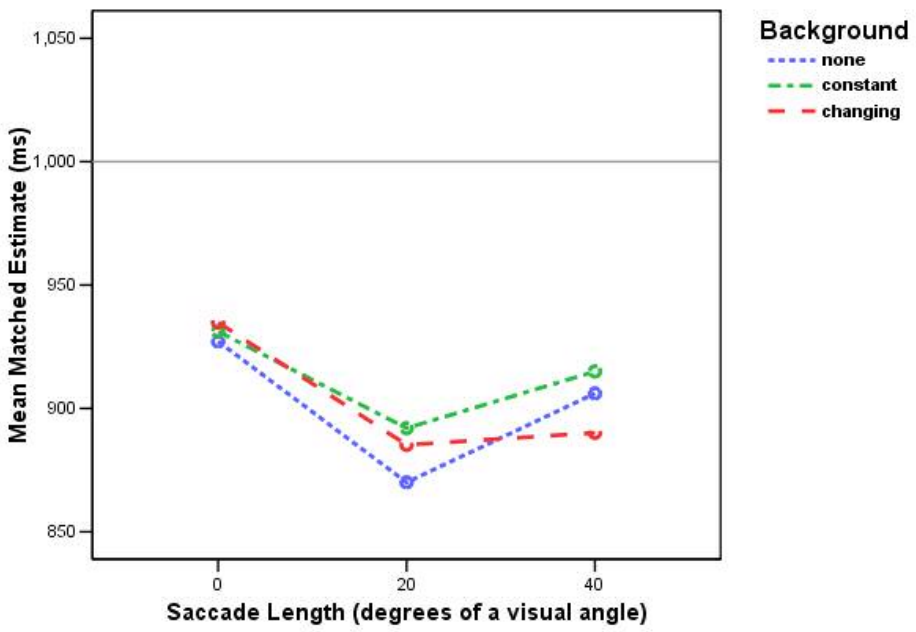


Figure 6-20: Effect of Saccade (x-axis: degrees of visual angle) on Matched Estimates (y-axis: ms) across the three Background conditions (none = blue-dotted, constant=green-dot-dash, changing=red-dashed) averaged across all subjects.

When matched estimates for constant-Background conditions 4, 6, and 8 are averaged across all subjects and compared to matched estimates for the no-Background conditions 1, 2, and 3 (see Figure 6-18) there is a main effect of Background (GG: F=3.579, df=1, p<.05, one-tailed). Matched estimates with constant background are always higher than without. This difference is visible in Figure 6-20: green dot-dash line = constant Background, blue dotted line = no Background. However, within each Saccade length there is no significant difference between the Background conditions. This trend for the constant-Background conditions to produce higher matched estimates is completely opposite what was predicted under the Masking Hypothesis (hypothesis 7). This could be caused by averaging across the four presentation conditions. Random presentation conditions were predicted to produce shorter matched estimates (hypothesis 7) and Blocked conditions the same matched estimates as the no-Background conditions (hypothesis 8).

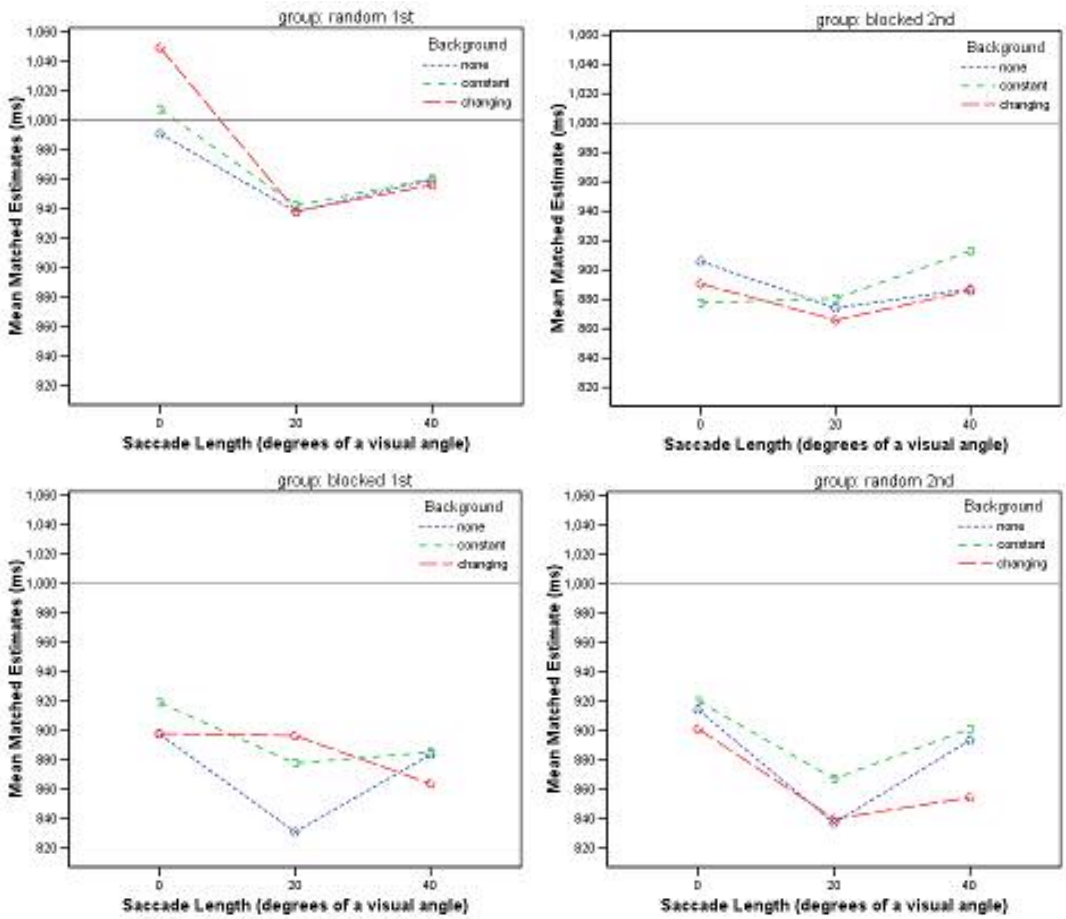


Figure 6-21: Effect of Saccade (x-axis: degrees of visual angle) on Matched Estimates (y-axis: ms). Three lines represent different Background conditions. Each graph represents a different presentation condition. Top row represents data from the first subject group (Random 1st, Blocked 2nd) and bottom row the second subject group (Blocked 1st, Random 2nd).

Figure 6-21 illustrates the difference between matched estimates across Saccade and Background conditions split into the four presentation groups. When conditions are presented randomly in the first session (top left graph in Figure 6-21) there is no overall significant difference between estimates made with or without a background (GG: F=.640, df=1, p=.444, not sig.) and no significant differences for each Saccade length. This invalidates the Masking Hypothesis (hypothesis 7).

In the Random-2nd presentation group there is a main effect of Background (bottom right graph in Figure 6-21: GG: F=4.296, df=1, p<.05, one-tailed) but this is caused by the constant-Background conditions having higher matched estimates than the no-Background conditions. This difference is in the opposite direction to that predicted

by the Masking Hypothesis. However, none of the differences within Saccade lengths are significant.

For both of the Blocked presentation groups (bottom left and top right graph in Figure 6-21) there is no main effect of Background (Blocked 1st: GG, F=1.935, df=1, p=.198, not sig.; Blocked 2nd: GG, F=.009, df=1, p=.927, not sig.). This confirms the 2nd Resistance to Change Hypothesis (hypothesis 8).

6.3.2.5 Delayed Saccade Hypothesis

In section 6.1.2.4 a common type of cut, referred to as a Reverse-Angle Cut, was identified in which both the background detail and the position of the focal-object changed across the cut. It was hypothesised that the unexpected background change would attract attention and delay the onset of the saccade. As predicted that this initial distraction would shorten perceived duration (as predicted by the Distraction Hypothesis) and then the saccade would lengthen the perceived duration of the visual event occurring immediately after the cut (as predicted by the Chronostasis Hypothesis). This combined effect was formalised as the Delayed Saccade Hypothesis: "*When an unexpected background change occurs as the focal-object relocates perceived durations will be shorter than without the background change but longer than without the relocation.*

In light of the current experimental design this hypothesis can now be formalised in terms of the experimental conditions. The initial distraction, as shown in section 6.3.2.3, should lead to the shortening of the letter's perceived duration. This will require matched estimate longer than the Control (condition 1) to create perceived continuity. However, this shortening of perceived duration will then be partially offset by the perceptual extension occurring during the saccade. This will mean that matched estimates shorter than the same peripheral distraction condition without a saccade (condition 5) will be required to create perceived continuity.

Hypothesis 9:

> Delayed Saccade Hypothesis: When an unexpected background change occurs as the target relocates (conditions 7 and 9), matched estimates will be significantly longer than conditions 2 and 3 due to peripheral distraction but significantly shorter than condition 5 due to the over estimation occurring during the saccade.

However, this initial distraction will only occur if the background change is unexpected. If the change is expected the likelihood of attention capture will decrease and matched estimates should not change from the same saccade conditions without background change (conditions 2, 3, 4, and 6).

Hypothesis 10:

> Resistance to Change Hypothesis 3: When an expected background change occurs as the focal-object relocates (conditions 7 and 9), matched estimates will not differ significantly from estimates made under similar target relocations without a change (conditions 2, 3, 6, and 8).

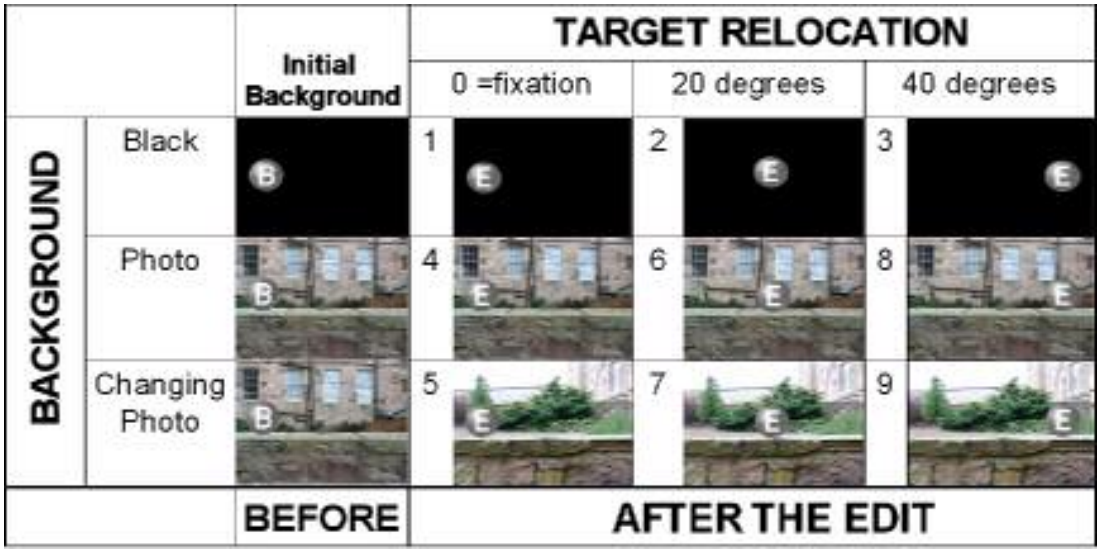


Figure 6-22: Experimental conditions used to investigate the Delayed Saccade Hypothesis. Matched estimates for saccades performed against changing backgrounds (conditions 7 and 9) were compared to the same saccades without a background change (conditions 2, 3, 6, and 8).

To test whether the initial background change affects matched estimates after saccades, conditions 7 and 9 must be compared to the same Saccade conditions without a Background (conditions 2 and 3) and with a constant background (conditions 6 and 8). However, the initial distraction should only serve to delay the onset of the saccade, not affect the saccade's effect on duration perception. Therefore, the first stage of analysis will be to check if the chronostasis effect still exists across the changing-Background conditions (5, 7, and 9).

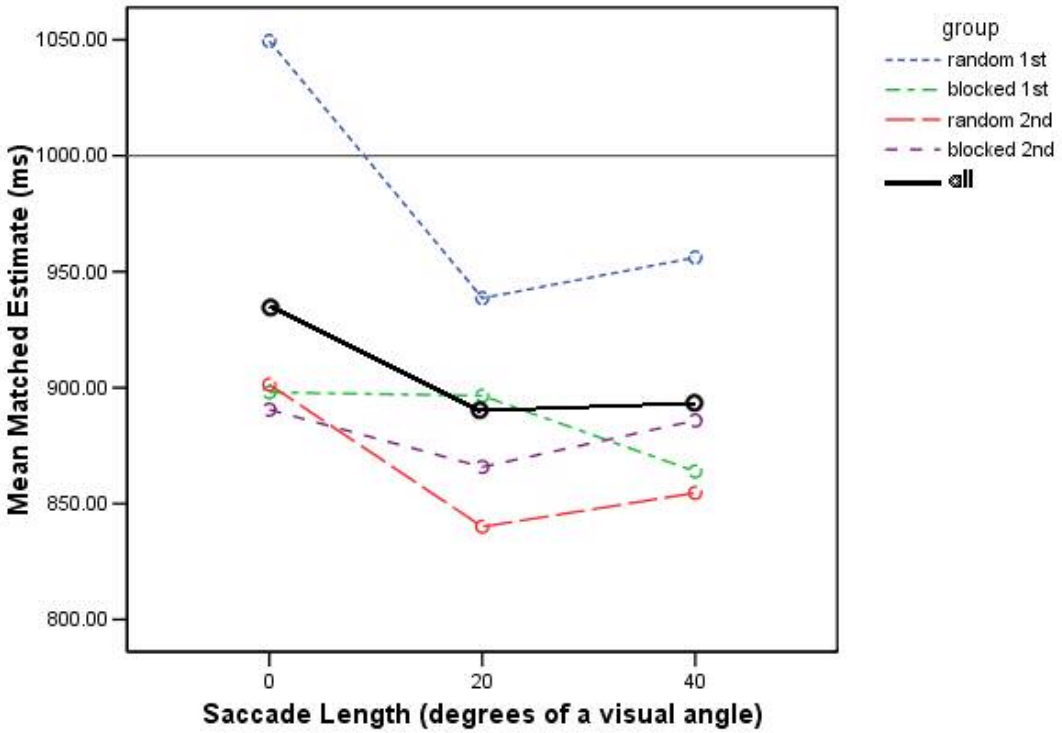


Figure 6-23: Mean matched estimates (ms) for the three Saccade lengths with a Background change (conditions 5, 7, and 9). Solid black line represents the average across all subjects. The four other lines represent estimates within each presentation condition.

Performing a repeated-measures ANOVA across the three degrees of Saccade (conditions 5, 7, and 9) averaged across all subjects (see solid black line in Figure 6-23) a main effect of Saccade is observed (GG: F=13.170, df=1.574, p<.000) with 20° and 40° Saccades both producing significantly shorter matched estimates (20°: mean diff. = 49.56ms, p<.000; 40°: mean diff.=44.70ms, p<.01). No other differences are significant. This confirms that the Chronostasis still exists when the onset of the saccade is accompanied by a peripheral change. In fact, conditions 7 and

9 show better support for the Chronostasis Hypothesis than conditions 2, 3, 6, and 8 did (see 6.3.2.2 and 6.3.2.4). The Chronostasis Hypothesis predicted that both saccades of 20° (conditions 2, 6, and 7) and 40° (conditions 3, 8, and 9) would show significantly lower matched estimates than the same Background conditions without a saccade (conditions 1, 4 and 5 respectively). However, in the other Background conditions only 20° saccades have shown the predicted significant decrease in matched estimates (see 6.3.2.2 and 6.3.2.4). Now, with a changing Background, full support for the Chronostasis Hypothesis is found.

As can be seen from Figure 6-23, when the data is split across presentation groups the chronostasis effect becomes somewhat less pronounced in the blocked (blue dotted and red large-dashed lines) but more pronounced in the random presentation groups (green dot-dashed and purple small-dashed lines). When conditions are presented randomly in the first session (blue dotted line in Figure 6-23) there is a main effect of Saccade (GG: F=16.449, df=1.681, p<.000) and this can be attributed to both 20° and 40° Saccades producing matched estimates significantly lower than 0° (20°: mean diff.=110.85ms, p<.000; 40°: mean diff.=93.30, p<.01). Matched estimates for 20° are not significantly different to 40°. When conditions are randomised in the second session (red large-dashed line in Figure 6-23) there is a main effect of Saccade (GG: F=4.214, df=1,747, p<.05) which can be attributed to condition 7 (20° mean = 840.05ms) having significantly lower matched estimates than condition 5 (0° mean = 901.33ms, p<.05). Also, matched estimates are almost significantly lower for condition 9 (40° mean = 854.60ms) compared to condition 5 (p<.053, one-tailed). These results indicate that matched estimates for both of the random presentation groups support the Chronostasis Hypothesis.

When conditions are blocked in the first session (green dot-dashed in Figure 6-23) there is no main effect of Saccade (GG: F=1.846, df=1.311, p=.202, not sig.) and no significant between-Saccade condition differences. When conditions are blocked in the second session (purple small-dashed line in Figure 6-23) there is no main effect of Saccade (GG: F = .564, df=1.385, p=.521, not sig.) and no significant between-Saccade condition differences. This lack of Chronostasis effect when the conditions

are blocked can clearly be seen in Figure 6-23 (green dot-dashed and small dashed purple line). In comparison to the Random-1st group (blue dotted line) the two Blocked lines are virtually flat across the three saccade lengths. As expectation was not expected to change the Chronostasis effect this evidence is in opposition to the Chronostasis Hypothesis.

Now that the existence of the Chronostasis effect has been established (at least in the random presentation groups) the effect of Background change must be identified. This will be accomplished by comparing matched estimates for the changing-Background conditions (5, 7, and 9; red dashed line in Figure 6-20 on page 293) to the no-Background conditions (1, 2 and 3; blue dotted line in Figure 6-20). A repeated-measures ANOVA comparing matched estimates for these conditions shows no main effect of Background (GG: $F=.087$, $df=1$, $p=.771$, not sig.). There is also no significant difference between Background conditions within each Saccade condition. Matched estimates for the changing Background conditions (red dashed line) are generally longer than estimates without a background (blue dotted line) except for after 40° saccades which produce longer estimates without a background (none of these differences are significant). This result is contrary to the predictions of the Delayed Saccade Hypothesis.

If the matched estimates are divided into the four presentation groups the Random-1st group shows a main effect of Background (GG: $F=5.549$, $df=1$, $p<.05$). However, this can be solely attributed to the large difference between condition 1 and 5 ($t=-3.880$, $df=9$, $p<.01$). The matched estimates for 20° and 40° saccades across the two Background conditions are almost identical (see the difference between the large-dashed red and dotted blue lines in the top-left graph of Figure 6-21). This evidence invalidates the Delayed Saccade Hypothesis. Instead, the clear effect of distraction on duration perception observable when the letter is static (condition 5) seems to be overridden by the over-perception of duration occurring during the saccade (conditions 7 and 9).

When the conditions are blocked in the 1st session there is no main effect of Background (GG: F=.734, df=1, p=.414, not sig.). As can clearly be seen from Figure 6-21 (bottom-left graph) this is due to conditions 7 and 9 (large-dashed red line) not matching the pattern of estimates of conditions 2 and 3 (dotted blue line). Whilst, condition 7 does produce significantly higher matched estimates compared with the same Saccade condition without a background (condition 2: t=-2.282, df=9, p<.05) conditions 5 and 9 produce very similar matched estimates to 1 and 3 (respectively). This indicates that there is no clear effect of Background change when the presentation conditions are predictable. This supports the 3rd Resistance to Change hypothesis.

When the conditions are blocked in the 2nd session there is no main effect of Background (GG: F=.470, df=1, p=.510, not sig.) and no within-Saccade differences. Similar to the Blocked-1st presentation group, this evidence indicates that an expected background change prior to a saccade has no effect on duration perception. This supports the 3rd Resistance to Change hypothesis.

When conditions are randomised in the 2nd session there is no main effect of Background (GG:F=1.100, df=1, p=.322, not sig.) and there are no within-Saccade differences. This indicates that the unexpected background change has no effect on duration perception when the change also triggers a saccadic eye movement. This invalidates the Delayed Saccade hypothesis. Combined with the same absence of distraction in the Random-1st presentation condition and the clear effect of Saccade for both Random presentation groups, it appears that accompanying the onset of a saccade with an unexpected peripheral change actually improves the Chronostasis effect.

6.3.2.6 Looking for Tolerances

Throughout the course of this analysis matched estimates have been regarded as a single, specific presentation duration that subjects perceived as being equal to 1000ms under specific experimental conditions. This absolute value allowed different conditions to be compared so that any effect of the independent variables

(Saccade and Background) could be extracted. However, in reality the matched estimate is actually just the midpoint of a range of presentation durations that the subject may perceive as being equal to 1000ms. This range has previously been referred to as the "Zone of Tolerance" (see section 6.2.3.6). As can be seen from Figure 6-24, subject's judgements range from "definitely shorter" for low presentation durations (probability of subject selecting "longer" is close to 0.0), through a period of indecision around the matched estimate (~0.5), to definitely longer for high presentation durations (~1.0). This period of indecision is important as it indicates how imprecise subject's judgements are under different conditions. Some conditions may make it difficult for subjects to form clear judgements about the letter's duration whilst others may aid duration perception making the "zone of tolerance" much smaller. If this "zone of tolerance" could be identified it would provide editors with a more realistic measure of viewers' ability to distinguish between presented durations. This would allow them to vary the degree of temporal discontinuities incorporated across a cut whilst still having the confidence that the cut will be perceived as temporally continuous.

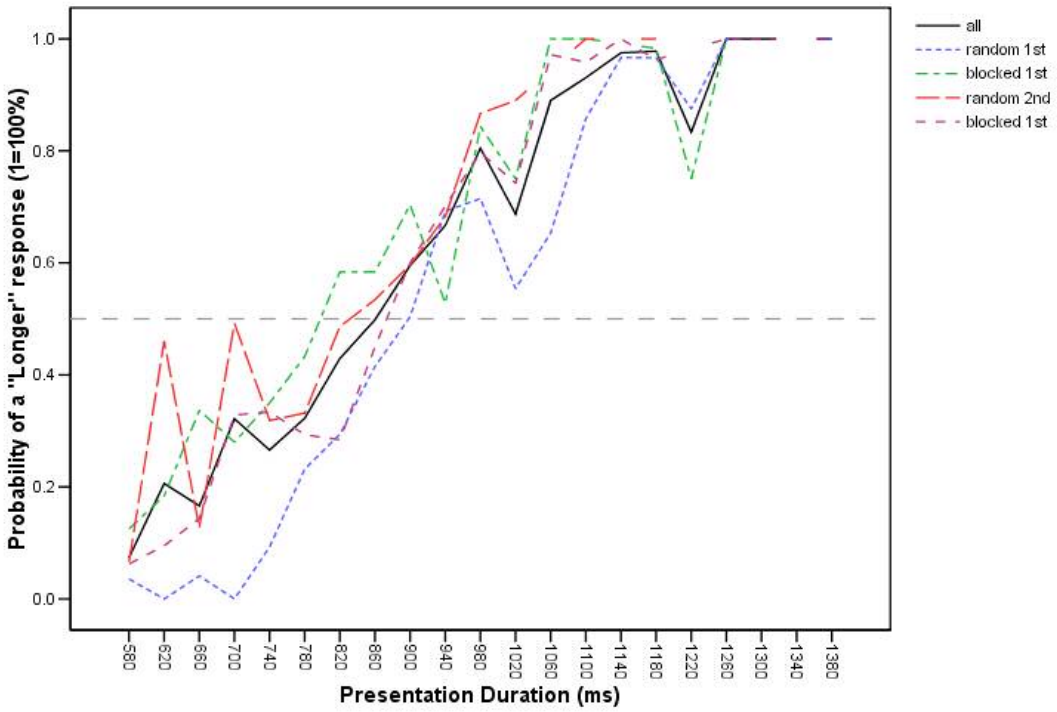


Figure 6-24: Probability of a "Longer" response (1=100%) for each presentation duration. Solid black line represents the average probabilities across all subjects. Other lines represent the four presentation groups.

To calculate this "zone of tolerance" the difficulty exhibited by the subjects in making a duration judgement (reaction time) and the statistical confidence in the matched estimate were used (confidence intervals).

6.3.2.6.1 Reaction Times

Reaction time is an accepted measure of the difficulty or cognitive load required to process stimuli (e.g. Lang & Basil, 1998). It was expected that, as presentation durations approached the matched estimate subjects would take longer to respond as they were unable to identify the duration as longer or shorter. When reaction times are mapped across the entire range of possible durations, it is expected that there would be a noticeable peak in reaction times around the matched estimate.

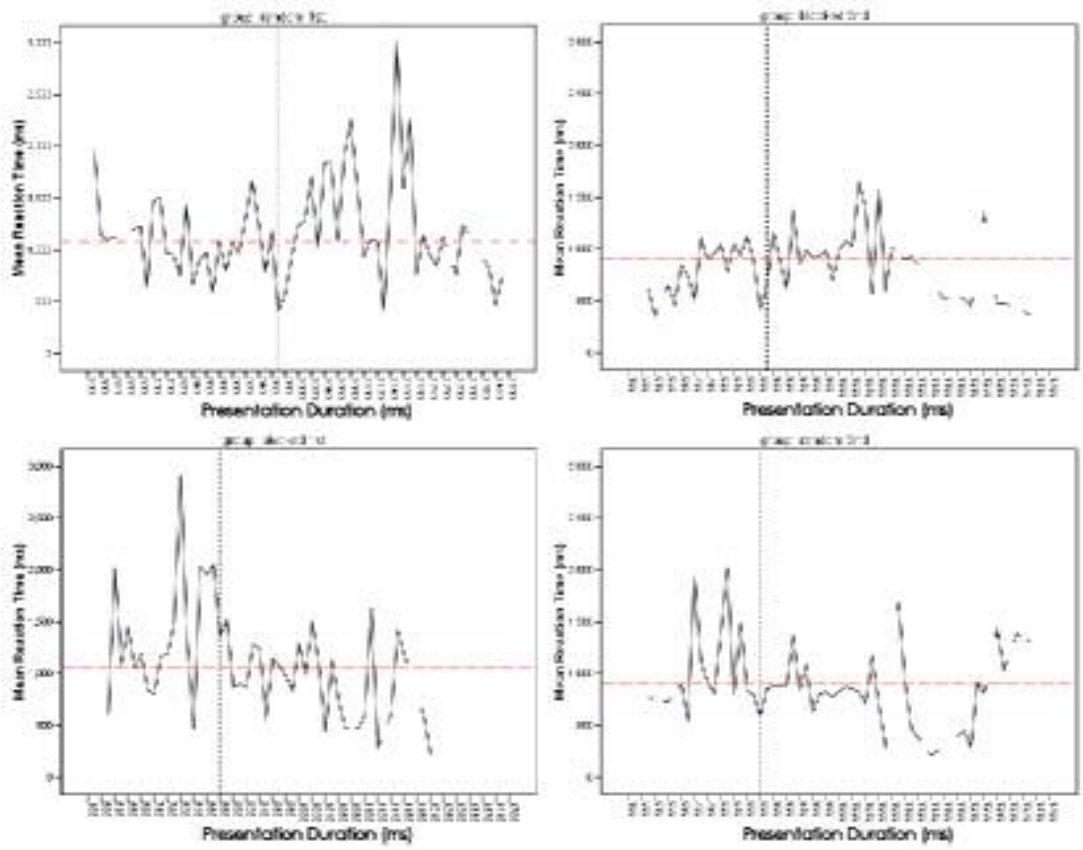


Figure 6-25: Mean reaction times (y-axis: ms) across the range of presentation durations (x-axis: ms) under condition 1 (0° Saccade + no Background). Each graph represents a presentation condition: top left = random 1st, top right=blocked 2nd, bottom left=blocked 1st, bottom right=random 2nd. Dotted black vertical line denotes the matched estimate and the dashed red horizantal line denotes the mean reaction time for that presentation condition.

As can be seen in Figure 6-25, when reaction times are mapped across the range of presentation durations within condition 1 (0° Saccade + no Background) for each presentation group there is no peak in reaction times around the matched estimates (dotted black vertical line). There is very little deviation of reaction times from the mean (dashed red horizontal line) and what deviation there is doesn't seem to follow any predictable pattern. This complete absence of any effect on reaction time of proximity to matched estimate seems to suggest that reaction time is not a useful indicator of "tolerance" in this study. One possible reason for this may be an error in the experimental instructions. If reaction time was to be used to detect tolerance it must be assumed that subjects are responding as soon as they have formulated an opinion on the target letter's duration. If they are waiting after this decision point before they respond their response data will be uninformative. In the instructions to this experiment subjects were instructed to respond as soon as they had decided on the letter E's relative duration but that they may base this judgement on letters presented both before and after the 'E' (see section 6.2.3.5). In hindsight, it can now be seen that these instructions are contradictory as either the subject responds as soon after the offset of the letter 'E' as possible, basing their judgement on letters presented before the 'E', or they continue to compare the 'E' to subsequent durations and respond once they have seen enough. This means that there is no consistent point from which reaction times can be measured. Therefore, reaction time data does not represent a subject's difficulty with judging the E's relative duration and cannot be used to gauge tolerance.

To fix this design error the instructions would have to be more specific about which durations the subject is expected to compare. This could be the duration immediately prior to 'E' and then after 'E' no letter would be presented, forcing the subject to respond as quickly as possible. Alternatively, they could be instructed to compare 'E' to the letter presented immediately after 'E' and then reaction times would be measured from this letter's offset.

Whilst solving the reaction time problem, these designs might actually alter the way that different experimental conditions affect duration perception. For example, when

subjects are told to use the letter before 'E' as a comparison, if masking occurs at the onset of 'E' this could perceptually extend the duration of 'E' back into the previous duration, shrinking the previous duration. This might mean that the letter 'E' has to be presented for less time to match the duration of the previous letter as this has also been distorted. By comparison, if subjects are instructed to compare 'E' to the next letter, masking at the onset of 'E' would not effect the comparison duration and 'E' would need to be presented for longer to match the comparison letter. This difference would provide an interesting test of masking that wasn't possible in the current design[138].

6.3.2.6.2 Confidence Interval

Given that the reaction time is not suitable for indicating a "tolerance" range other data must be used. Luckily there exists a simple statistical measure that can be applied to the matched estimates to generate this range: confidence intervals.

Previously, matched estimates have been calculated as "point estimates" of the population mean (i.e. representative of all subjects within a group). It is assumed, that the sample used in this study (N=20) is representative of the entire population and, therefore, the mean of the sample (i.e. the matched estimate) should also be the mean of the population. However, there will always be a degree of error in this estimate due to differences between samples. The population mean could occur at any point along the entire range of values used in the sample but it is most probable that it occurs close to the sample mean. To express this error and represent the mean as a range of possible values rather than a point estimate, the mean's confidence interval can be calculated. Confidence intervals (typically) represent a range of values within which there is a 95% probability that the population mean will occur. This interval is calculated using the standard error of the sample mean which is based

[138] In this study, the apparent absence of masking is inconclusive as it is not known whether the letter before or after 'E' was being used as the comparison. It could be the case that a mixture of the two tactics has produced a matched estimate which appears to show no signs of masking but is actually an average of the influence of masking and distraction.

on the standard deviation of the distribution and the sample size. As the standard deviation (i.e. the spread of the distribution) decreases or the sample size increases (getting closer to the population size) the standard error decreases. This makes the size of the 95% confidence interval shrink, making the sample mean closer to the population mean.

In most statistical tests, 95% confidence that a value does not differ from another value is taken as near certainty. Any value falling within this 95% confidence interval is seen as not deviating significantly from the expected value. If this idea is now applied to matched estimates, it becomes clear that, whilst the matched estimate represents a point estimate for the presentation duration perceived by a population as being equal to 1000ms there is actually a 95% confidence interval within which other presentation durations would also be perceived as being equal to 1000ms. If this confidence interval is calculated for each experimental condition it can be interpreted as a conservative "zone of tolerance". If this zone was extended beyond the 95% confidence interval the probability that the zone incorporated durations not perceived as equal to 1000ms would increase. Therefore, the standard 95% confidence interval will be used to identify a "zone of tolerance" for each of the experimental conditions in this study.

Rank	Condition	Presentation Group	Confidence Interval (ms)
1	9	random 1st	81.34457944
2	6	random 1st	107.6285167
3	3	blocked 2nd	114.6537736
4	8	random 1st	118.2980084
5	1	random 2nd	120.2961372
6	3	random 1st	122.3592818
7	9	blocked 2nd	128.6529504
8	2	blocked 2nd	129.6305514
9	5	random 2nd	129.9475686
10	4	random 2nd	131.6555462

Table 6-2: The ten smallest confidence intervals (ms).

Confidence intervals were calculated for each experimental condition and presentation group (see ranked table of the ten smallest intervals, Table 6-2). The entire table of confidence intervals, means, and their rank according to interval size can be seen in Appendix D[139]. This table indicates which conditions and groups were the most precise in their duration estimates. If subjects were able to identify that presentation durations close to the matched estimate were not equal to 1000ms the distribution of their responses would be very narrow. This would lead to small confidence intervals. As can be seen from Table 6-2, the condition with the smallest confidence interval (by a large degree) is the random 1st presentation group under condition 9 (changing Background + 40° Saccade). This condition is the most complex as it involves object relocation and background change. It might have been expected that condition 1 would produce the smallest confidence interval as all subjects had to do was judge the letter's duration without any distractions. However, this condition only figures once in the top ten (rank 5, random 2nd) where as condition 9 features twice. This indicates that subjects are more consistent in their duration judgements when the duration begins with an eye movement. In fact, only one of the first eight conditions with the shortest confidence intervals does not contain an eye movement.

The clearest result that can be seen from Table 6-2 is the apparent advantage of random presentation. Seven out of the first ten conditions are presented randomly and four of these in the first session. This can clearly be seen when confidence intervals are represented as error bars (Figure 6-26). Random 1st produces smaller confidence intervals on average (131.11ms) compared with the other presentation groups (blocked 1st = 167.14ms, random 2nd = 140.58, blocked 2nd = 142.30). This yet again confirms the consistency and clarity of the effects associated with random presentation in the first session.

---

[139] Appendix D also contains confidence intervals for data averaged across presentation groups ('all' in the 2nd column) and experimental conditions ('all' in the 1st column). As these averages are taken from larger samples than the individual conditions and groups their confidence intervals are smaller. This ranks the averages higher in terms of the size of their confidence interval.

In summary, confidence intervals are a good measure of subject's tolerance for deviation about the matched estimates. There is a 95% probability that a duration lying within the confidence interval will be perceived as being equal to 1000ms. The smaller the confidence interval, the more consistent subject's responses are indicating that their viewing behaviours are also similar. As between and within subject differences increase, confidence in the matched estimate as a representative mean for the whole population decreases, increasing the confidence interval. The clearest example of this is in condition 5 which has the largest confidence interval across all subjects probably due to the unpredictable effect of distraction (see Appendix D).

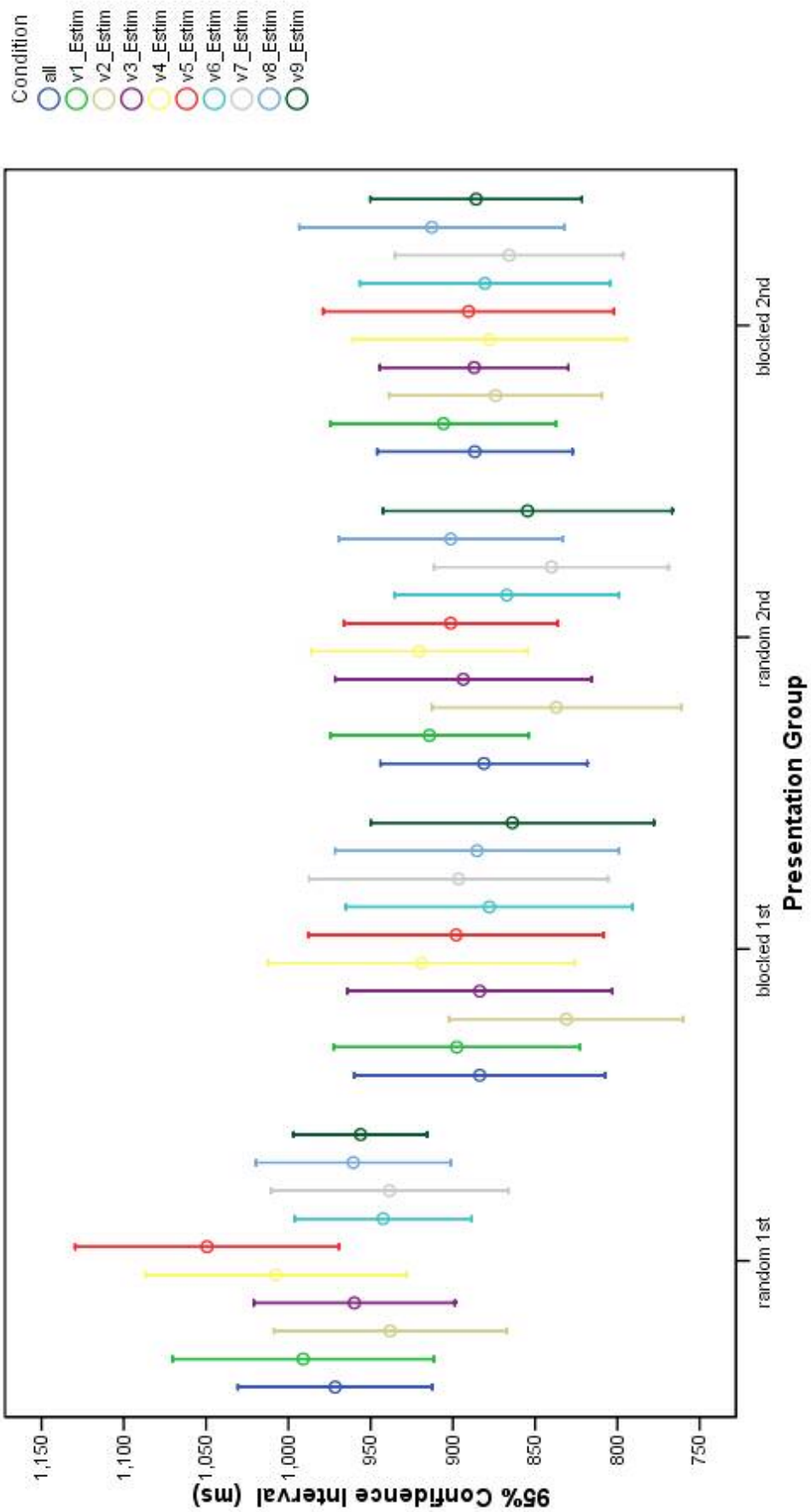


Figure 6-26: 95% Confidence Intervals (ms) for each condition (different lines) and presentation groups (x-axis). Bars represent the confidence interval about the mean (i.e. the matched estimate = circle at the centre of each bar).

6.4 Experiment 2: Discussion

This study investigated the individual and combined effects of peripheral distraction and saccadic eye movements on duration perception. Ten hypotheses were formally tested under various viewing conditions. The results pertaining to each hypothesis will be discussed in turn. A slight alteration to the order of the hypotheses in the Results section (6.3.2) will be used to aid logical development. The Chronostasis Hypothesis (Hypothesis 3) will be discussed after the Resistance (H4), Anticipation (H5), and Distraction Hypotheses (H6) so that the chronostasis effect across all Background conditions can be discussed. The rest of the hypotheses will be discussed in order.

6.4.1 Concentration and Divided Attention Hypotheses

The first analysis established baseline matched estimates for the duration perceived as 1000ms without any target relocation (Saccade) or background change (Background). This value was then used as the baseline value to which all subsequence matched estimates were compared.

Hypotheses 1 (Concentration Hypothesis) stated that when the viewing conditions were predictable matched estimates would be significantly shorter than 1000ms. This hypothesis was supported by existing evidence from time estimation studies (Block et al., 1980; Zakay, 1989). Analysis of matched estimates under the blocked control condition in this study (condition 1, no Saccade + no Background) indicated that matched estimates were significantly shorter than 1000ms as predicted by hypothesis 1 (mean difference = 98.34ms). Significantly shorter matched estimates were also recorded when the data was split across the two experimental sessions: mean difference for first session was 102.4ms and for the second session 94.28ms. When matched estimates were averaged across all experimental conditions (not just condition 1) the Blocked presentation condition still produced significantly shorter matched estimates (see section 6.3.1.1). This evidence confirms hypothesis 1: when

the editing conditions are expected a visual duration presented after the cut is perceived as longer than its actual duration.

Hypothesis 2 (Divided Attention Hypothesis) predicted that when the viewing conditions could not be predicted (the Random presentation condition) matched estimates would be significantly longer than 1000ms as attention would be divided between the time estimation task and preparing for the possible focal-object relocation. This hypothesis was also based on existing time estimation studies (Block et al., 1980; Zakay, 1989). Analysis of the data for condition 1 showed that matched estimates were significantly longer in the Random presentation condition compared with the Blocked condition but that they were not significantly longer than 1000ms. This result does not support hypothesis 2. In fact, matched estimates under the Random condition were very close to 1000ms: 952.62ms across both sessions, 990.98ms in the first session, and 914.25ms in the second session. This difference between matched estimates when conditions were presented randomly in the first and second sessions was seen repeatedly throughout the rest of the analysis. Potential causes of this difference between the Random groups will be discussed in the next section (6.4.2).

Identifying the baseline visual durations perceived by subjects under the Blocked and Random conditions as being equal to 1000ms allows the effect of expectation to be removed from all subsequent analyses. It is assumed that by comparing matched estimates produced under other experimental conditions to these baseline estimates the effect of expectation can be removed and any remaining difference attributed to the particular experimental manipulations.

6.4.2 Resistance, Anticipation, and Distraction Hypotheses

The Resistance (Hypothesis 4), Anticipation (H5), and Distraction Hypotheses (H6) predicted how the effect on matched estimates of either a constant (H5) or changing background (H6) was modified by expectation (H4). Existing studies have shown

that when attention is distracted during a time estimation task, subject's will perceive the elapsed time as being shorter than without the distraction (Block et al., 1980) (Zakay et al., 1983a). In the current study, such distraction was predicted to occur when a constant (H5) or changing (H6) background was presented and the experimental conditions randomised. Both Background conditions should result in longer matched estimates. However, if the distracting event is expected and known to be irrelevant it will not capture attention (Simons, 2000). Therefore, the Resistance to Change hypothesis (H4) predicted that there would be no difference between matched estimates in the changing background condition (condition 5) compared with the constant (4) or no background condition (1).

The results of this study support hypothesis 4 but only for the first session. There is no effect on matched estimates of the presence of a constant or changing background when these conditions are presented blocked in the first session. However, when the conditions are predictable in the second session the condition without any background produces significantly higher matched estimates (905.72ms) than the condition with a background change (877.55ms). This result does not support hypothesis 4 and is in the opposite direction to that predicted by hypothesis 5. A possible explanation could be that subjects are actively ignoring the background when they know that it isn't going to change or the target relocate. This might increase their concentration to the time estimation task resulting in the artificial extension of the perceived duration as is predicted by the Concentration hypothesis. Why this effect of concentration would only be seen in the second blocked presentation group is not known.

Hypothesis 5 is not supported by these results although there is a slight trend in the right direction. When conditions are presented randomly in the first session the constant Background condition does show slightly higher matched estimates compared to the no Background condition (mean difference = 16.71ms). In the second session the difference drops to 6.15ms. These differences are not significant and compared with the increase in matched estimates caused by background change in the first session (mean difference between conditions 1 and 5 = 58.44ms) suggest

that a constant background causes no distraction. This is not surprising considering that the same photograph was always used, subjects were instructed to fixate the letters and ignore the background, and the background was static so it contained no features that may have attracted attention (e.g. moving leaves).

Hypothesis 6 is supported by the results of this study but only during the first session (i.e. subject group Random-1st). When the background changes unexpectedly (condition 5) matched estimates are significantly higher (mean=1049.43ms) than without a change in the background (condition 4: mean=1007.70ms) or any background at all (condition 1: mean=990.99ms). The absence of any distraction effect when conditions are randomised in the second session is highly unexpected. One possible explanation could be that during the first session, in which the experimental conditions were blocked, subjects were able to learn to ignore the background change (this was clearly seen in relation to hypothesis 4). When the experimental conditions are predictable this should be relatively easy but it appears that this group of subjects (Blocked 1st and Random 2nd) are also able to apply the skill to the subsequent random session. Such an ability to build up resistance to the attention capturing effects of a repeated sensory event is known as habituation. Habituation is a widely established attentional phenomenon (see Posner, 1994 for a review) and seems like a valid explanation for the absence of distraction in the second session of this study.

The confirmation of hypotheses 4 and 6, in combination with hypothesis 1 in the previous section (6.3.2.1) provide further supporting evidence for the role of attention in time perception (Block, 1990; Block et al., 1980; Zakay, 1989; Zakay, 1992; Zakay & Block, 1996; Zakay & Block, 1996; Zakay et al., 1983b).

6.4.3 Chronostasis Hypothesis

The experimental design utilised in this study allowed the affect of saccadic eye movements and background distraction on matched estimates to be identified in isolation before examining their combined effect. The first stage of this process was

to check that the modifications to Yarrow et al's (Yarrow et al., 2001b) original experimental paradigm made for this study did not affect the Chronostasis effect. Previous demonstrations of the Chronostasis effect instructed subjects to perform a saccadic eye movement to a target location and only once their eyes were moving was the target of the time estimation task presented (Yarrow et al., 2001b; Yarrow et al., in press; Yarrow et al., 2004). By recording the duration of eye movements, these studies were able to identify that subject's perceived the target after the saccade as beginning about 50ms before the eyes began moving (Yarrow et al., 2001b). The experimental design used in this study used a target relocation to both mark the beginning of the target duration and cue the saccadic eye movement. It was assumed that because the saccade would be performed during the target's presentation duration any perceptual extension would affect the perceived duration of the target. However, there was the possibility that, given that the target had already been perceived prior to the saccade, the perceptual extension of the saccade duration would not occur as this would clash with the last time point represented for the target prior to the saccade. Therefore, it was imperative that the existence of the Chronostasis effect was established in this experimental design before any extra factors were introduced (i.e. Background).

To first examine the affect of saccadic eye movements on matched estimates without any background information conditions, 1, 2 and 3 were used (see Figure 6-7). Comparison of the matched estimates across these three conditions partially confirmed the existence of the Chronostasis effect. Across all subjects durations containing a 20° relocation of the target (condition 2) were perceived as 57.11ms longer than without a target relocation (condition 1). Both randomised and blocked presentations of the experimental conditions across both sessions produce this effect. Hypothesis 3 also predicted that 40° saccades would produce a similar increase in perceived duration (and a corresponding decrease in matched estimates). A significant decrease in matched estimates from 0° to 40° Saccade conditions was found when estimates were averaged across all subjects but the effect was significantly smaller than for 20° saccades and was not found to be significant in any of the individual presentation groups. The consistency of the chronostasis effect

(~50ms) irrespective of saccade length is a key component of the Chronostasis effect as identified by Yarrow et al (Yarrow et al., 2001b). Therefore, the absence of the same size decrease in matched estimates for 40° saccades compared to 20° indicates that our evidence does not fully confirm the Chronostasis Hypothesis.

This tendency for 40° estimates to be longer than 20° is most noticeable for the second subject group (blocked 1st and random 2nd). Figure 6-15 clearly shows a large increase in matches estimates after 40° saccades for the second subject group (large-dashed red and dot-dashed green lines). This could indicate that the subjects in this group were using a different viewing tactic to the other group. For example, given that the experimental conditions were blocked during the first session for these subjects they may have chosen to fixate the position at which the 'E' would appear after it has relocated rather than fixate the letters prior to the 'E' and saccade once the 'E' had appeared. If subjects adopted this viewing strategy the presentation duration of the letter 'E' would not contain a saccade so there would be no chronostasis effect, just as observed for this subject group. The experimental instructions clearly stated that subjects must fixate the letters at all times and only move their eyes after the 'E' had appeared. However, without eye tracking data there is no way to check their actual viewing behaviour.

Once subjects had developed this viewing strategy in the first session they may have applied it to the second session. For this subject group, the experimental conditions were presented randomly during the second session. As the subject would be unable to predict what was going to happen at the cut, there would be less of an obvious benefit for fixating anywhere on the screen other than the location of the initial letters but this does not mean that the subjects did not continue to use some form of alternate viewing strategy. However, the use of an incorrect or variable viewing tactic should have shown up in the number of repetitions required to complete a MOBS (longer when viewing behaviour is more inconsistent) but there are no significant differences in repetitions between subject groups. 30% of subjects did report on the questionnaire that they "defocused their eyes" instead of saccading

when conditions were blocked in the first session[140] however the matched estimates for these subjects do not prove to be outliers so they cannot be excluded from the analysis.

The only other explanation that might explain the absence of chronostasis effect for 40° saccades without any background could be that the wrong "baseline" matched estimate is being used. It was assumed that any effect of concentration on perceived duration would be consistent across all Blocked experimental conditions (see 6.3.2.1). However, when the subject knows that a saccade will be required they may divide their attention between the saccade target and the letters presented before 'E'. As section 6.3.2.1 has shown, divided attention leads to significantly higher matched estimates. If the matched estimate for condition 4 (40° saccades + no Background) under the Blocked viewing condition (see green dot-dashed line in Figure 6-15) were compared to a Random Control, such as condition 1 from the Random 1st subject group (blue dotted line) the difference would be considerably larger and more likely to be significant. However, Yarrow et al (Yarrow et al., 2004) have shown a consistent Chronostasis effect across various types of saccade and degrees of expectation. In their experiments they used similar control conditions as used in the current study. Therefore, the no-Background Saccade conditions (1, 2, and 3) used in this study should exhibit a chronostasis effect and the lack of an effect cannot currently be explained.

The chronostasis effect becomes more apparent once background is added. With a constant background the 20° saccade condition shows a significant decrease in matched estimates (39.18ms) compared with the control (0° saccade) across all subjects. As in the no-Background conditions the decrease in matched estimates for 40° saccades predicted by hypothesis 3 is not seen. When conditions are presented blocked in the first session only 40° saccades show a chronostasis effect and none of the saccade conditions show an effect when conditions are blocked in the second session. When conditions are randomised in the second session only 20° saccades

[140] This technique allows them to judge the onset and offset of 'E' via peripheral change instead of following it with their eyes.

show signs of the chronostasis effect and this can again be attributed to the possible alternate viewing strategy used by the second subject group. The clearest signs of the chronostasis effect are found when conditions are randomised in the first session. Both 20° (mean difference to 0° = 65.3ms) and 40° saccades (mean difference = 47.17ms) show significantly shorter matched estimates compared to the condition without a saccade. This can not be attributed to the control condition (0° Saccade + constant Background) having an abnormally high matched estimate due to the distracting effects of the background as there is no significant difference between this condition and the no Background condition (see 6.4.2).

The best experimental conditions for showing the chronostasis effect across all subjects was when the background changes. Both 20° and 40° saccades exhibited a significant decrease in matched estimates compared with the control condition when accompanied by a background change (20° mean difference from 0°=49.56ms; 40° mean difference=44.70ms). However, when the data is split across the four presentation groups it becomes apparent that this effect is due to the contribution of the Random viewing conditions only. When conditions are blocked in either the first of second session there is no significant decrease in matched estimates across the Saccade conditions. By comparison, matched estimates for the Random 1st group show a significant decrease from 0° Saccade to both 20° (mean difference=110.85ms) and 40° (mean diff.=93.30ms). These differences are much higher than the 50ms predicted by hypothesis 3. This can be attributed to the distracting effect of the unexpected background change in the fixation condition: mean difference between change and no Background condition is 58.44ms. The effect of distraction can be removed by comparing matched estimates for condition 1 (no Background + 0° Saccade) to the changing Background Saccade conditions. This comparison indicates a less significant decrease in matched estimates for 20° saccades (mean difference from 0° saccades = 52.40ms, $p<.05$) and an almost significant decrease for 40° saccades (mean diff.= 34.85, $p=.069$, one tailed). Removing the effect of distraction from the control condition makes the decrease in matched estimates across the Saccade conditions less pronounced but it indicates that there is still support for hypothesis 3 and the existence of the chronostasis effect.

Support for hypothesis 3 under changing-Background conditions is also provided by the Random 2nd group. Matched estimates for this group decrease significantly from 0° to 20° saccades (mean difference = 61.28ms) and almost significantly to 40° saccades (mean difference = 46.73ms). For this group there are no concerns about the distorting effects of distraction on the 0° saccade condition as no significant difference was found between changing-Background and no-Background conditions (see 6.4.2). These results suggest that, even though the Random 2nd group have been previously accused of employing alternate viewing tactics, when then viewing conditions are highly demanding (background change and target relocations) they perform as predicted by the Chronostasis hypothesis.

This analysis of the Chronostasis hypothesis (H3) has shown the clearest evidence of the chronostasis effect for unpredictable viewing conditions with a background. The inconclusive evidence of the Chronostasis effect when conditions are predictable may indicate a higher degree of variability between subjects. When subjects can predict when the target is going to relocate they may choose to perform their saccade before, during or after the target relocation. This variability of saccade start time would affect the accuracy of the MOBS for identifying the matched estimate. This can clearly be seen in the confidence intervals depicted in Figure 6-26. The larger the confidence interval the more variance between matched estimates across subjects. The random presentation groups have the smallest confidence intervals and the blocked groups the largest.

The fact that when the viewing conditions are unpredictable there is greater degree of agreement between the duration perceived as 1000ms indicates that there must be a high degree of conformity between subjects viewing behaviour. When the focal-object suddenly relocates across the screen, irrespective of what is happening to the background, all subjects saccade to the object. Evidence for this attentional synchrony has previously been presented in section 3.4.2.2. This saccade creates a perceived temporal discontinuity which is quite consistent across subjects (Yarrow et al., 2001b). This consistency allows editors to predict and accommodate their

audience's temporal expectations. Exactly how they should do this will be discussed later (section 6.4.6).

6.4.4 Masking Hypothesis

The Masking hypothesis (hypothesis 7) predicted that the sudden uncovering of the background when the target relocates would mask the last few milliseconds of the previous shot. This would artificially extend the perceived duration of the new shot resulting in shorter matched estimates. Masking is a well established phenomenon (Breitmeyer, 1984) but its effect on time perception is not known. Joseph Anderson proposed that masking will occur every time a cut is made and that it would have to be compensated for if temporal continuity were to be perceived (Anderson, 1996). Hypothesis 7 investigated the validity of this proposal. It was also predicted that masking would only occur when the target relocation was unexpected. This was formalised as the Resistance to Change 2 Hypothesis (H8).

The results of this study fail to provide any evidence that masking occurs when a focal-object relocates and reveals a background scene. When saccades are performed against a photographic background, matched estimates tend to be slightly higher than without a background but this is not significant (random-2nd does show a main effect of Background but this does not represent any specific differences). If masking was occurring, matched estimates would be expected to be significantly lower than the same Saccade conditions without a background. As such, this evidence invalidates hypothesis 7 and supports hypothesis 8. Instead of a masking effect, the slightly higher matched estimates in the Background conditions suggest that a small degree of distraction may occur instead (similar to that expected in hypothesis 5).

The lack of support for the Masking Hypothesis in this study does not rule out the possibility that masking will affect time perception under different viewing conditions. In the current experimental design whenever there was the potential for masking it was also accompanied by another highly salient visual event: the focal-object relocation. It appears that the object relocation captured attention which

obscured any signs of masking. If, for example, a cut was made to a shot with no focal-object (a technique known as "Clearing the frame" Katz, 1991) the new shot may be seen to mask the end of the previous shot just as was proposed by Joseph Anderson (Anderson, 1996). If the effect of masking on time perception were to be investigated in a future study this would be the type of stimuli that would need to be used. The mask would have to be presented at fixation (like most masking studies; see Breitmeyer, 1984) and different duration comparison tasks used to measure if the masking effects the end of the last visual event or the beginning of the new event.

One other problem with using this experimental design to investigate masking was that subjects could either compare the target event, 'E', to letters before or after the target. If masking deletes the last few milliseconds of the previous visual event, when a subject tries to match the duration of the target event to that of the masked event the matched estimate would be shorter than when compared to a non-masked event. Whether subjects used the pre or post-target events for comparison was not controlled in this study which could explain some of the variance in the data. The instructions and direction of comparison would have to be improved if this study was to be replicated.

6.4.5 Delayed Saccade Hypothesis

The Delayed Saccade Hypothesis (hypothesis 9) predicted that when an unexpected background change occurred at the same time as the focal object relocation, subjects would experience a period of distraction before they were able to initiate their saccade. This hypothesis was informed by Edward Dmytryk's belief that such a cut leads to the perceived absence of 125-208ms (3-5 frames at 24fps; Dmytryk, 1986). Dmytryk suggests that in order for the viewer to perceive temporal continuity this missing period of the action must be replayed. In the present study, the same overlap of action could be achieved by presenting the target letter 'E' for 1150ms. This extra duration would accommodate the 200ms lost due to distraction and the 50ms gained through chronostasis. However, as has been seen in relation to the Distraction Hypothesis (section 6.4.2), the largest effect distraction has on matched estimates is

58ms (Random 1st). This indicates that it is unlikely that distraction can explain Dmytryk's overlap.

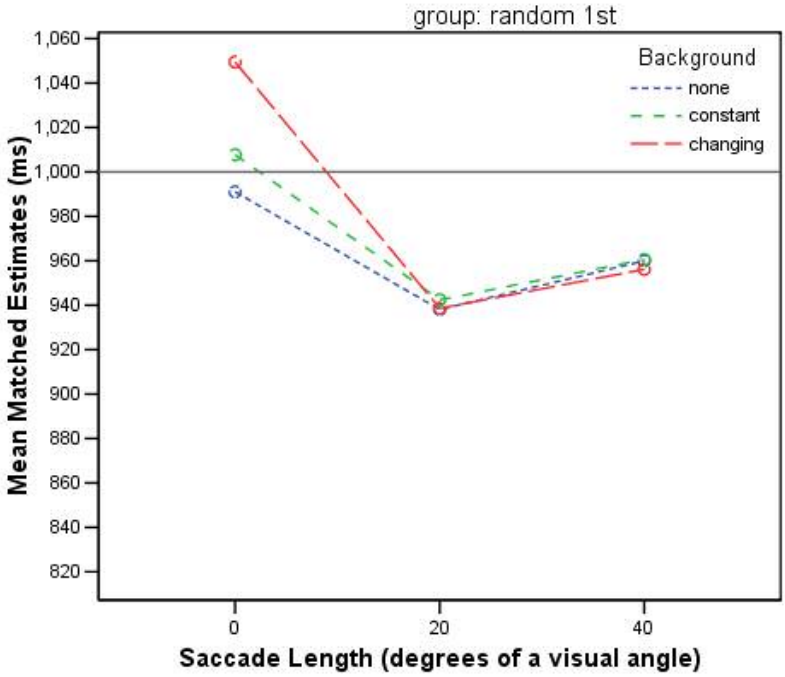


Figure 6-27: Effect of Saccade (x-axis: degrees of visual angle) on Matched Estimates (y-axis: ms). Three lines represent different Background conditions (blue-dotted = no background, green-dot-dash= constant, red-dashed =changing). Data is just for Random 1st group.

The most interesting finding of the whole study is presented in Figure 6-27. According to the Delayed Saccade Hypothesis, matched estimates should show a systematic increase when the background changes in comparison to the no or constant Background conditions. This was predicted to function along side the Chronostasis effect. In a graph like Figure 6-27, this effect of distraction would have been clearly visible as the line representing the changing Background conditions (large-dashed red line) being higher than the other lines whilst still following the same pattern across Saccade conditions. The fact that the actual results show a complete absence of such a difference indicates that an unexpected background has no effect on time perception when co-occurring with focal-object relocation. In other words, the results of this study do not support the Delayed Saccade Hypothesis.

The difference between the effect of background change when the focal-object is static compared with when it moves can be explained as a function of the viewer's

attentional set (see 3.2t). The viewer's task is to judge the duration of the focal object. To perform this task they need to fixate the object at all times. As the viewing conditions are unpredictable the viewer needs to divide their attention between the time estimation task and any sudden changes to the visual scene that might indicate a target relocation. The effect of this divided attention is to allow them to judge durations very accurately (see 6.3.2.1). However, this divided attention also means that they are susceptible to attention capture by a background change during fixation (the effects of which can be seen in Figure 6-27). The viewer's attention is captured by the background change because their attentional set is tuned to direct attention quickly to any relocation of the focal-object. This attentional set also appears to make them sensitive the sudden onset of new objects even when the focal-object is still present (i.e. the background change).

When the focal-object relocates at the same time as a background change the attentional set ensures that the viewer's attention will be captured by the focal-object as, in the stimuli used in this study this was the most salient new object in the visual scene. If the focal-object was hidden or obscured in someway or another object more salient then it was presented this other object may capture attention instead of the focal-object. The effect this unexpected change in object identity would have on perceived temporal continuity is not known.

As predicted by the 3rd Resistance to Change Hypothesis (hypothesis 10) the results of this study also indicated that there was no effect on matched estimates of background change when the experimental conditions were predictable. The viewers do not need to divide their attention when the experimental conditions are blocked as they always known when a saccade will be required. This allows them to prepare for the background change and limit its ability to capture their attention.

The lack of any supporting evidence for Dmytryk's 3-5 frame overlap is both a positive and a negative result. It is a positive result as it indicates that focal-object relocation is a reliable way to capture attention as suggested in chapter 3 and it produces a systematic perception of time across the cut that can be used to create

perceived temporal continuity. It is a negative result as Dmytryk cannot be dismissed as incorrect for he has years of practical experience of editing and he believes that there is a need for a 3-5 frame overlap. The inability of this study to support his hypothesis or explain how it may function may indicate that this study is not representative of the film viewing experience.

The key component that distinguishes the types of cuts discussed by Dmytryk from that used in this study is motion. Dmytryk was specifically referring to match-action cuts when he suggested overlapping 3-5 frames of "action" across the cut. As was discussed in depth in section 3.3, coinciding a cut with the onset of motion at the beginning of a visual event is a very effective way of ensuring that the visual disruption of the cut does not capture attention. The difference between Dmytryk's insight and the results of the present study could be due to the lack of motion in this study. There is some indication that the duration of a visual event may be perceived differently depending on whether the event is "empty", such as an event marked only by the onset and offset of a letter as used in this study, or "filled" by action (Fraisse, 1963). However, it is not known how a sudden change in motion, such as that caused by a cut, might affect time perception.

In the first experiment of this thesis (see 4.4) one of the side-effects of the experimental design was the subjects' apparent inability to accurately perceive the initial position of the moving object after the cut. This was attributed to the *Fröhlich Effect:* when a moving object suddenly appears on the screen, whether in clear-view of from behind an occluder, the object's initial position is typically misperceived further along its path (Fröhlich, 1923). This distortion has been attributed to the reallocation of attention to the moving object after its onset (Müsseler & Aschersleben, 1998). As attention also affects viewers' ability to accurately perceive time (Zakay & Block, 1996) the *Fröhlich Effect* could distort time perception after the saccade in a way that is compatible with Dmytryk's proposed overlap.

In fact, recent evidence suggests that the opposite is true: perceived space is distorted to accommodate the misperception of time associated with the Chronostasis effect.

(Yarrow et al., 2005). In a recent follow up to their original Chronostasis study, Yarrow and colleagues investigated temporal and spatial perception across saccadic eye movements to moving targets (Yarrow et al., 2005). They found that the perceived location of a moving target after a saccade is systematically mislocated forwards along its path by 37ms[141]. Yarrow et al suggest that this mislocalisation cannot be solely attributed to the *Fröhlich Effect* as it is initially smaller (55ms) than would normally be seen with the *Fröhlich Effect* (100ms)[142] and then increases in size over the first 500ms after the saccade. If the subject is asked to identify the position at which the target changes identity within 500ms after the saccade they identify a position 140ms ahead of its actual position. 103ms of this misperception can be attributed to *representational momentum*: the misperception of the last presented position of a moving object forwards along its path (Freyd & Finke, 1984). Yarrow et al attributed the extra 37ms to a perceived adjustment of the objects spatial location compensating for the perceived temporal distortion caused by the saccade (Yarrow et al., 2005). These findings indicate that perceived spatiotemporal coherence takes precedence over purely spatial accuracy. In other words, a coherent spatiotemporal percept will be constructed to match the misperceived duration of the saccade even though the spatial component does not match the actual location of the target.

Yarrow et al's findings suggest that, when presented spatiotemporally continuous motion across a saccade, a viewer will perceive spatiotemporal discontinuity due to the spatial compensation of the temporal distortion caused by the saccade. The viewer perceives the action as jumping forwards in space and time. The only way to present the motion so that it is perceived as being spatially continuous across the saccade is to relocate the moving object back along its path to the same degree as will be later compensated. By presenting a spatial discontinuity, when the perceived spatial location of the object is later mislocalised forwards to compensate for the

---

[141] The mislocalisation is presented as the time between the end of the saccade and the time that the object would have actually been at the location identified.

[142] Yarrow et al used control conditions in which the moving target suddenly appeared at fixation and found that the initial mislocalisations were greater: 100ms.

temporal discontinuity the object will be perceived in the right location. Tentative evidence for this technique was found by Yarrow et al who showed that relocating the moving target backwards by 24ms during the saccade resulted in perceived spatial continuity (Yarrow et al., 2005). However, this finding was not statistically significant so. A more comprehensive repetition of this study is required.

If this technique of creating perceived spatial continuity from presented discontinuities is applied to film editing it begins to resemble Dmytryk's "overlap" technique. When a continuous action is depicted in two shots separated by a cut the moving object can be relocated back in time by replaying frames from the last shot at the beginning of the new shot. An overlap of only one frame (41.67ms at 24fps) would be required to compensate for the chronostasis effect (~50ms). This is less than the 3-5 frame overlap suggested by Dmytryk but it is in the right direction. Future empirical studies should endeavour to find further supporting evidence for this technique of creating perceived spatiotemporal continuity by presenting spatiotemporally discontinuous action.

6.4.6 Formalising Informed Editing Guidelines

The results of this study can now be used to suggest how temporal expectations can be accommodate across different types of cut to create perceived temporal continuity. Only cuts with a background change will be discussed as most real cuts involve a background change.

6.4.6.1 Background change + Focal object collocation

When a cut is unexpected and the focal-object does not change screen position across the cut a Jump Cut will result (see 6.1.2.2). The background change will attract attention which will result in a shortening of perceived duration. To accommodate this distortion the duration of an event presented after the cut must be increased by

58.44ms. This could be achieved by replaying the last 1 or 2 frames from the previous shot as suggested by Anderson (1996) and Dmytryk (1986).

According to the results of this study, an overlap of 58.44ms should result in perceived temporal continuity. However, for any individual subject there is a 95% probability that the actual degree of overlap they will perceive as continuous lies somewhere within the range 138.52ms to -21.64ms (i.e. an ellipsis) (see confidence interval table Figure 6-26). The size of this confidence interval (from 4 frames overlap to 1 frame ellipsis) indicates that the 58.44ms overlap is not representative of the whole population.

Also, it should be acknowledged that the effect of distraction on perceived duration reported in this study is for one particular example of distraction. Other types of stimuli or change of stimuli presented under different viewing conditions may produce different degrees of distraction and different distortions of perceived time. This study only presented one degree of distraction so, whilst a relationship between distraction and the distortion of perceived time can be concluded, it should not be assumed that the same the size of distortion (58.44ms) will be observed with different degrees of distraction. Some evidence does exist indicating that as the degree of semantic relatedness between shots decreases the amount of attention captured by the cut increases (Geiger & Reeves, 1993; Lang et al., 1993). However, as well as semantic factors other, more immediate factors will also affect the degree of attention capture such as contrast in colour, brightness, shape, level of detail, and objects. All of these factors will contribute to how the change of shot captures attention yet it is very difficult to formalise any of these factors as a continuum of change. How can you identify one level of change as having no effect on attention and another as a large effect? Such a continuum would be needed if the effect of a change on perceived duration is to be predicted. The attentional theory of time perception assumes that as the amount of attention allocated to the time estimation task decreases the accuracy of perceived duration will also decrease (Zakay & Block, 1996). The result is a negative relationship between degree of distraction and

perceived duration. However, without a precise measure of the degree of distraction no predictions can be made about the resulting effect on perceived duration.

From an editing perspective this unpredictability of the effect of distraction on perceived duration makes it hard for an editor to accommodate distraction when editing for continuity. Unlike the chronostasis effect which is defined by its regularity (Yarrow et al., 2001b), distraction cannot be precisely accommodated. Both Joseph Anderson and Edward Dmytryk suggest overlaps of the action across a cut which might accommodate the effects of distraction (even though they didn't describe them as such). However, they also both acknowledge that there is no one amount of overlap that works for every cut. The editor must "tweak" each cut until it works (i.e. is perceived as continuous). This unpredictability of distraction could be one reason why Jump Cuts are forbidden within the continuity style of filmmaking.

The other, more immediate reason is that distraction implies attention capture by an unexpected change. As was discussed in the Hiding a Cut chapter (3) expectation forms a conceptual bridge across a cut. The viewer does not become aware of a continuity cut as their attention is directed towards the answer to a perceptual question established by the previous cut. When this expectation is absent the sudden change creates a conceptual "rupture" that the viewer must work to resolve. To understand the relationship between the old and the new shot they must reflect on the content of the old shot, the new shot, and their comparison. The viewer has to compensate for the editing rather than having their perceptual enquiries accommodated by the editing.

The most obvious way to eradicate the negative effects of a Jump Cut is to make the viewer expect the cut. The results of this study have shown that when a background change is expected it causes no distraction and has no effect on perceived duration. Various techniques exist for cuing a cut (see chapter 5) but the most common is a match-on-action. If a match-on-action functioned in the same way as the blocking condition used in this study the action could be presented as temporally continuous and it should be perceived as continuous. However, the results of Yarrow et al

(Yarrow et al., 2005) and the classic *Fröhlich Effect* indicate that some compensation for perceived spatial distortions may be required. The exact form of these distortions needs to be investigated in future studies.

6.4.6.2 Background change + Focal object relocation

A very common type of cut (referred to as Reverse-Angle; see 2.1.3) changes both the background scene and the location of the focal-object across the cut. The findings of this study indicate that when the focal-object relocation is unexpected it will overtly attract attention resulting in a lengthening of perceived duration. The background change will have no effect on perceived duration so perceived temporal continuity can be created by deleting 43.63ms[143] (~1 frame at 24fps).

A lot more confidence can be placed in this recommendation compared to the overlap prescribed for cuts with unexpected background change as there is a lot less variance in the matched estimates. The 95% confidence interval when condition 9 (changing Background + 40° Saccade) is presented randomly is 81.34ms. This is half the size of the confidence interval for condition 5 (changing Background + 0° Saccade). This indicates that for an individual subject the actual degree of ellipsis that will lead to the perception of temporal continuity will lie somewhere between 2.96ms and 84.3ms. This is only a difference of 2 frames. Based on this evidence an editor could remove one frame from the beginning of a new shot when the cut is unexpected and the focal object shifts across the screen and they could be 95% confident that their audience will perceive the resulting film as continuous.

---

[143] This value was derived from the average difference between matched estimates for 20° and 40° Saccade conditions with Background change (conditions 7 and 9) and the control condition with no Background or Saccade (condition 1). The Random 1st subject group were used as they had the most authentic viewing conditions (i.e. no expectation or training).

The two main constraints on this technique for creating perceived continuity are that the focal-object is not moving and that the cut is unexpected. As Yarrow et al (Yarrow et al., 2005) have shown, when the saccade target is moving spatial distortions may occur that compensate for the temporal distortions. This combination of perceived distortions may have to be accommodated by an editor for their cuts to result in perceived continuity.

The cut must also be unexpected as this study has shown that, whilst 20° saccades sometimes show the chronostasis effect when they are voluntary, in general the saccadic behaviour across subjects is too variable for temporal perception to be predicted. However, in this study expectation was based on repetition not an overt form of cuing as is commonly used in film (e.g. shift in gaze, pointing, walking off-screen). The repetition allowed subjects to initiate their saccade whenever they wanted in relation to the cut. In continuity editing, deictic cues are usually positioned just before a cut (see chapter 5). If there is less than 200ms between the cue and the cut, the subject will not be able to complete a saccadic eye movement before the cut has occurred. This would ensure that when they begin perceiving the scene after the saccade the focal-object has already relocated. If the cue occurs less than 80ms before the cut, the sudden relocation of the focal-object may actually facilitate the planning of the saccadic eye movement. Such facilitation is purely hypothetical but it does seem to be compatible with the way in which deictic cues are used in continuity editing.

6.4.6.3 Expected vs. Unexpected

In the previous two sections suggestions about how to create perceived temporal continuity where given in relation to the change occurring across the cut. It was assumed that the composition of the shots could not be changed and the only modification at the editor's disposal is the removal or addition of frames. However, a director also plans for continuity when they are choosing which shots to use before filming a scene. At this point during the film's production there is complete freedom in how the shot can be composed. The only constraint the director has is the action

being filmed[144]. This action might control if the point at which the director envisages a cut is cued or uncued (e.g. does the actor turn their head). As this is the only constraint the director has they need to know how best to compose a shot so that continuity can later be constructed during editing.

If the cut is uncued collocation of the focal-object would result in a Jump Cut. This has been shown to lead to a high degree of variability between how viewers perceive temporal continuity. It would be very hard for the editor to accommodate distraction and ensure temporal continuity was perceived by all viewers.

The solution is to relocate the focal-object across the uncued cut. The target relocation will uniformly capture attention, limiting the degree to which the viewers are aware of the background change, and decreasing the variance between viewer's temporal expectations. The director just has to compose the shot so that the focal-object of the new shot is the most salient object in the shot and, preferably, significant to the viewer. This significance should ensure that the focal-object is compatible with the viewer's attentional set.

If the director has to compose a shot which is cued they are safest collocating the focal-object across the cut. An expected background change shows no sign of attention capture and does not affect perceived temporal continuity. There is also some indication that the existence of the cue may allow the conceptual resolution of the cut (see chapter 5). If the composition of the shot was changed so that the focal-object relocated across the cut the editor may find it more difficult to find a degree of ellipsis that would be perceived as creating temporal continuity for all viewers. The only way that a cue could be used in combination with an object relocation is if the cue occurred immediately before the cut and it was a *push* cue such as a shift in gaze. A *push* cue directs attention towards the target of the cue where as most cues, such as a sudden movement, attract attention to themselves (known as *pull* cues; Langton et

[144] The action can be modified slightly but it must always be in primary service of the script and characterisation. I have never heard of a director instructing an actor to perform a certain action because it will be used to create a continuity cut during editing.

al., 2000). A *pull* cue would lead to perceived spatial discontinuity across the cut as by the time the saccade to the cue had finished the cut would have relocated the source of the cue elsewhere. This explains the convention of collocating matched-actions (Pepperman, 2004).

6.5 Experiment 2: Conclusion

This study had two main goals: 1) to identify the distortions of perceived duration occurring across a range of cuts and 2) show that the accommodations required for these distortions to result in perceived continuity are compatible with the conventions of continuity editing. Previous chapters of this thesis have identified systematic shifts of attention that can be used to limit viewer's awareness of a cut (chapters 3 and 4). The timing of these attentional shifts seem to match what editors would regard as acceptable edit points (chapter 3). Hiding a cut in an attentional shift ensures the assumption of existence constancy across the cut (chapter 4 and 5) but once the attentional shift is over the viewer's spatiotemporal expectations must be satisfied otherwise the viewer will perceive the discontinuity caused by the cut.

Unfortunately, satisfying spatiotemporal expectations is not just a matter of matching an object's location in time and space across a cut. Our ability to accurately perceive time and space is dependent on attention (Zakay & Block, 1996). When attention shifts our spatiotemporal expectations appear to distort (Müsseler & Aschersleben, 1998; Yarrow et al., 2001b; Yarrow et al., 2005). If spatiotemporal continuity is to be perceived across a cut the distortions resulting from the attentional shifts caused by the cut must be accommodated. As the intention of continuity editing is to maintain *"the spectator's illusion of seeing a continuous piece of action"* (Reisz & Millar, 1953; pg 216) there must exist evidence of such accommodation in the continuity editing rules.

There does exist some indication that editors tend to overlap action by 2-5 frames across a match-action cut (Anderson, 1996; Dmytryk, 1986). Such an overlap is compatible with the effect of covert attentional distraction on time perception (Block

et al., 1980). When attention is distracted during an event, the duration of the event is perceived as shorter than without the distraction. If the distraction is caused by a cut an overlap of action across the cut might accommodate the perceived distortion and result in perceived temporal continuity. However, distraction is not the only attentional shift resulting in the distortion of perceived duration. When a saccadic eye movement is performed to a static target the duration of the target will be perceived as longer than the same duration presented during fixation (the Chronostasis effect; Yarrow et al., 2001b). If the chronostasis effect occurs when a focal-object relocates across a cut, the first frame of the new shot would have to be removed so that the distorted temporal expectations are accommodated and result in perceived continuity. No evidence for such ellipsis can be found in the editing literature.

One possible explanation for editors' apparent ignorance of the chronostasis effect is the possibility that the viewing conditions of a cut combine both distraction with chronostasis to distort temporal expectations. Chronostasis has only previously been investigated under abstract experimental conditions with no background information (Yarrow et al., 2001b; Yarrow et al., in press; Yarrow et al., 2004). It was hypothesised that the inclusion of a sudden background change (as would be found in a typical cut) at the same time as the focal-object relocation might result in an absence of perceived time due to distraction followed by an increase in perceived time across the saccade.

This hypothesis and others were investigated in a duration estimation experiment. Matched estimates were derived that represented how long the focal-object had to be presented after the cut for its duration to be perceived as 1000ms. The size and direction of the matched estimates deviation from 1000ms indicated the perceived temporal distortions caused by the attentional shifts across a cut.

The findings of this study indicated that distraction and chronostasis both occur when the changes caused by a cut are unpredictable. If the cut involves a sudden background change without a relocation of the focal-object the viewer's attention will be captured by the change. This results in a perceived shortening of the focal-

object's duration requiring an overlap of one frame across the cut for time to be perceived as continuous. This result is compatible with the overlap suggested by Anderson and Dmytryk. However, as acknowledged by Anderson and Dmytryk, the degree of distraction is not consistent across different changes of stimuli so the precise degree of temporal overlap would need to be varied according to the degree of change. If the background change is expected it will not capture attention or affect perceived duration. Expectation can be controlled by cuing the cut either rhythmically (such as in this study) or through the use of attentional cues such as onsets of motion or shifts of gaze (see chapter 3)[145].

This study also provided evidence of the chronostasis effect across various viewing conditions. The clearest perceptual extension of the duration of a saccade was observed when an unexpected relocation of the focal-object was accompanied by a background change. If the background was kept constant across the cut or no background was used, the chronostasis effect was observed for short saccades (20°) but became less apparent for longer saccades (40°). This pattern was also observed when the editing conditions were predictable. To accommodate the perceived distortion of duration when the saccade is unexpected, an ellipsis of one frame across a cut was suggested. The matched estimates produced under the saccade conditions were significantly more precise than those produced under distraction. This indicates that an unexpected focal-object relocation creates consistency between subject's attentional shifts. This can be attributed to attentional capture. These results indicate that if a cut cannot be cued the focal-object should be relocated across the cut so that the focal-object rather than the background change captures attention.

The absence of any effect of unexpected background change on perceived duration when accompanying a focal-object relocation supports the technique of Reverse-Angle editing associated with the 180° Rule. As has previously been discussed in sections 3.2 and 5.3, the 180° Rule seems to ensure that the screen location of an object is preserved during a scene. This allows viewers to shift their gaze between

[145] If the attentional cue is collocated across the cut. If the cue shifted time perception would be distorted by the chronostasis effect.

depicted objects without the object moving from the expected location. The results of the present study indicate that even if such a gaze shift occurs in response to a cut there will be no difference in how attention shifts between locations or how time is perceived across the shift. Evidence of synchrony of gaze shifts has already been presented (section 3.4.2.2) but this is the first evidence that perception is also synchronised and unaffected by the cut.

What is not currently known is how perceived temporal distortions combine with spatial distortions to create spatiotemporal expectations across attention shifts. Yarrow et al (Yarrow et al., 2005) have produced some evidence that space perception distorts to accommodate the temporal distortion occurring during a saccade. It is not clear from their results how the combined distortions of space and time should be accommodated to create the perception of spatiotemporal continuity although there is some indication that an overlap of action, such as that suggested by Dmytryk may be required. There is also uncertainty about whether precise matching of spatiotemporal expectations is required for the perception of spatiotemporal continuity. Occlusion studies have shown that viewers are quite tolerant of temporal discontinuities[146] (Spelke et al., 1995). Future studies need to investigate the perceived continuity of moving objects across cuts.

In summary, this study has shown that editing creates a range of different attentional conditions which affect time perception in specific ways. There is some evidence that editors are aware of the variability of perceived temporal continuity but without analysing temporal continuity across existing films there is no way of knowing if editors actually match time by accommodating the temporal distortions reported by this study. The two key techniques that can be used to create attentional and perceptual synchrony between viewers is cuing of the cut and relocation of the focal-object. Both minimise the perceived disruption caused by the cut, allowing editors to accommodate the distortions caused by the attentional shifts and create perceived

[146] Similar spatial discontinuities have not been investigated although it can be predicted that similar tolerance would be observed.

temporal continuity. Whether this temporal continuity results in perception of existence constancy across the cut will be left for the general discussion (chapter 7).

Chapter 7: General Discussion

This chapter will first summarise the motivation behind this thesis before discussing the three main questions. The theory and empirical findings related to each question will be discussed and possible extensions and future work proposed.

7.1 Summary of Motivation

The motivation for this thesis came from the two assumed benefits of continuity editing: to make a cut "invisible" (Bordwell & Thompson, 2001; Dmytryk, 1986; Reisz & Millar, 1953) and to ensure that "the spectator's illusion of seeing a continuous piece of action is not interrupted" (Reisz & Millar, 1953). The first of these benefits requires viewers to have no awareness of the editing. Awareness of a sensory event occurs when sufficient attention is focussed to the processing of that event (Simons, 2000). If attention is captured by the event (is automatically drawn towards it), awareness should emerge automatically. The sudden change in visual information (referred to as "visual transients") associated with a cut has the potential to capture attention (see 2.3).

Evidence exists indicating that attention capture occurs when a cut violates the continuity editing rules and results in the viewer's increased awareness of the editing (d'Ydewalle & Vanderbeeken, 1990; Schröder, 1990). However, cuts composed according to the continuity editing rules show fewer signs of attentional capture (see 2.3). In terms of the level of visual change across a cut, both continuity and discontinuity cuts should be identical. The absence of attention capture for continuity cuts seems to indicate that the continuity editing rules specify editing conditions under which the potential for attentional capture by the cut is limited. The cut occurs

without the viewer automatically becoming aware of it, i.e. the cut is "invisible". However, the mechanisms by which continuity editing controls attention capture are not currently known.

As well as hiding a cut, the other assumed benefit of continuity editing, is that it does not disrupt the viewer's perception of continuous motion (Reisz & Millar, 1953). The instantaneous change in viewpoint associated with a cut may alter the spatial and temporal location of the depicted action, both on the screen and in the depicted 3D world. This spatiotemporal discontinuity would seem to rule out the perception of "continuity" across a cut. However, the concept of "continuity" as referred to by film editors is not fully understood.

The goal of this thesis was to identify how continuity editing rules control attention to influence awareness of cuts and create and maintain the perception of "continuity". These issues where addressed in stages according to three main questions:

How does continuity editing:

1. minimise awareness of a cut,
2. create the perception of "continuity" across a cut, and
3. ensure that "continuity" is not violated as a consequence of the cut?

The evidence and theories developed in answer to each of these questions will be presented. Conclusions will be made for each question and further work proposed.

7.2 Question 1: How does continuity editing minimise awareness of a cut?

The first question was investigated in chapter 3 by surveying existing attention literature and developing theoretical attention-capture avoidance mechanisms. Four techniques were proposed that seem to be used by continuity editing to limit the

attention capture caused by a cut: *focus on a constant, expect a visual change, direct attention internally,* or *suppress attention.*

If viewers focus their attention on a part of the visual scene that is unaffected by the cut (i.e. *focus on a constant*), it was proposed that their attention should be resistant to capture by the cut. This technique was based on the phenomenon of *inattentional blindness* (Mack & Rock, 1998). In film, *inattentional blindness* could be created by graphically matching the focal-object across a cut. The focussing of attention towards the focal-object combined with the limiting of visual transients due to maintaining the focal-object across a cut may limit the viewer's awareness of the cut. However, it was concluded that the level of precision required for a graphical match to successfully cause inattentional blindness is unrealistic for most cuts. Instead a similar effect can be created by gradually dissolving between shots (removing the transients that would capture attention) or digitally compositing a shot. Both of which may limit the viewer's awareness of the "join" between two shots but may have unforeseen side-effects on perceived continuity.

The second method for minimising attention capture across a cut was identified as the use of attentional cues. If a viewer expects a certain type of change, such as the appearance of a new object, the extraneous visual transients caused by the cut will be less likely to capture attention. (Most et al., 2005). The expectation may also form a conceptual bridge across the cut that allows the viewer to create a cohesive mental representation of the relationship between shots. The use of editing to match the viewer's expectations and answer *perceptual enquiries* (Hochberg & Brooks, 1978a) is believed to be the "*fundamental psychological justification for editing*" (Lindgren, 1948; page 54).

Viewers can be encouraged to expect a change through the use of attention cues: sensory events that *pull*, p*oint*, or *push* attention. In film, the use of *pull* cues, such as sudden onsets of motion constitutes one of the most common forms of continuity cut, match-action, and are thought by some to be critical to continuity editing (Pepperman, 2004). The types of attention cues differ in how they create perceptual

expectations across the cut. *Pull* cues seem to create limited perceptual expectations (e.g. "What was that motion?") where as *point* and *push* cues create the expectation that the target of the cue is spatially related to the cue. A *push* cue, the only known form of which is a gaze-shift also attributes an intentional relationship between the cue and its target. The use of gaze-shifts to cue and connect shots is a readily acknowledged part of the continuity editing style (Messaris, 1994). Whilst there appears to be evidence that attentional cues are used to cue cuts there is no understanding of how such cues contribute to the perception of "continuity". Investigating the perceptual consequences of attentional cues and expectation across cuts was the intention of the second empirical study of this thesis (see next question. 7.3).

The third technique continuity editing may use to minimise attention capture by a cut is to occupy the viewer with the cognitive processing of a previous sensory event. This creates a 500ms period during which new stimuli will not reach the level of awareness (known as the *attentional blink;* (Raymond et al., 1992). Tentative evidence of attentional blinks was found during event perception (Smith et al., in press). When viewers observe visual events they perceptually segment the events into units of activity (Zacks et al., 2001). The *breakpoint* between events is significant to our successful comprehension of the event (Newtson & Enqguist, 1976) and is associated with a rise in cognitive load (Smith et al., in press). Immediately after a *breakpoint* cognitive load decreases and frequency of saccadic eye movements increases (Smith et al., in press). This increase in eye movements may be due to the onset of motion at the beginning of the new event attracting attention. If a cut is made at this point, the visual transients of the cut should be hidden by the onset of focal-motion (Levin & Varakin, 2004) and attention should be focussed towards the new event. Therefore, it was proposed that locating a cut immediately after the onset of a new event should result in the lowest awareness of the cut. This suggestion matches the convention of match-action editing and is believed by editors to be the point a cut should be made to create the greatest impression of continuous motion (Reisz & Millar, 1953). This compatibility between

the conclusions of event perception and the editing conventions suggests that editors are identifying *breakpoints* when deciding where to cut.

The final technique suggested as a way to minimise attention capture combined attentional cues with periods of attentional suppression. When a person blinks or performs a saccade they are unable to perceive visual information due to suppression mechanisms (Burr et al., 1994; Burr, 2005). These periods of suppression provide perceptual "holes" in which cuts could be hidden. The possibility of using blinks to hide saccades has previously been suggested (Murch, 2001). However, a perceptual "hole" has to be predictable if it is to be used by an editor to hide a cut. Murch suggests that a viewer's blinks may synchronise with an actors blinks or with cognitive activity, providing a degree of predictability. However, no evidence exists that people synchronise blinks and the relationship between cognitive activity and blinks is unreliable (Fukuda et al., 2005). It has also been shown that there is no correlation between blinks and event *breakpoints* which suggests that they also do not correlate with edit points (Whitwell, 2005),

Whilst the period of suppression provided by blinks is unreliable, saccadic suppression is frequent and controllable through the use of attentional cues. Eye tracking evidence was presented indicating that viewers synchronised their saccadic eye movements across cuts and agree on the centre of interest in most shots (May et al., 2003; Stelmach et al., 1991; Tosi et al., 1997; Treuting, 2004). Tentative evidence was also presented indicating that viewers may perform saccades across cuts, suppressing the visual transients (May et al., 2003). This use of saccades as an attentional and perceptual "bridge" across cuts was investigated in the first study of this thesis (chapter 4).

In conclusion, the Hiding a Cut chapter (3) proposed mechanisms by which awareness of the cut could be minimised by controlling attention and perception. The key component in all of these mechanisms seems to be the use of attentional cues. Attentional cues provide editors with the ability to attract, direct, and synchronise the viewer's attention across cuts and during shots. Cues, such as gaze-shifts, are also

thought to create perceptual enquiries that may provide a basis for "continuity". However, the mechanisms proposed in the Hiding a Cut chapter (3) were hypothetical in that they were not derived from dedicated empirical studies. Finding direct empirical evidence for these mechanisms was the goal of the following chapter (4)

The first experiment of this thesis investigated the use of attentional cues to influence attention and provide perceptual "holes" in which a cut could be hidden (chapter 4). The attentional cues discussed in the previous section were all real-world cues that viewers should be primed to respond to by everyday life. For example, the reflexive attentional *push* caused when a person we are observing shifts their gaze is believed to result from a dedicated neurological system (Klin et al., 2002a) and plays a crucial role in our social ability (Langton et al., 2000). Such real-world attentional cues are supplementary to the attentional cue primitive to the film medium: occlusion of the focal-object by the screen edge. Editors believe that as a focal-object moves off-screen, the viewer's attention is *pushed* across the screen in expectation of the focal-object's reappearance at the opposite screen edge (Dmytryk, 1986). A cut composed to satisfy these expectations was identified and named: matched-exit/entrance cut. This cut was manipulated in an empirical study and a reaction time task used to probe attention across the cut.

The results of this study indicated that viewers expect a cut to occur when the focal-object is fully occluded by the screen edge, not before. This expectation is manifested as a saccade coincident with the cut. This saccade provides a period of saccadic suppression in which the visual transients of the cut can be hidden (see 3.4.2). If a cut occurs before the focal-object is fully occluded, viewers show no signs of anticipatory attention withdrawal and do not adapt to the cut over repeated presentations. However, when the focal-object is half occluded by the screen edge prior to the cut, saccades performed in response to the cut only last 125ms. This is insufficient time for a voluntary saccade to be performed (typically lasting 150-200ms) indicating that the saccade must be reflexive (i.e. controlled by sensory event, not the viewer). This interpretation is supported by evidence that saccades are

shorter when the focal-object appears fully on the screen immediately after the cut. This sudden appearance seems to capture attention, speeding the saccade across the screen.

Apparently, reflexive saccades are also seen when the focal-object is fully occluded prior to the cut. As the focal-object moves behind the screen edge attention is withdrawn from the focal-object. When the cut then occurs the saccade to the opposite screen edge takes less than 83ms. The important difference between cuts that fully occlude the focal-object before the cut and those that do not is that with full occlusion the reflexive saccade does not appear to require an object to capture attention. The rapid reflexive saccade occurs even when the focal-object relocates to a position behind the opposite screen edge. This occlusion removes the visual transients required for attention capture. Therefore, the reflexive saccade seems to be due to attention being *pushed* across the screen by the occluding screen edge.

This identification of the outgoing screen edge as an attentional *push* cue could explain the belief that matched-exit/entrance cuts create the impression of continuous action across a cut (Reisz & Millar, 1953). When an object gradually moves out of sight behind an occluder it is perceived as continuing to exist (referred to as *existence constancy;* (Gibson et al., 1969; Michotte, 1955). However, for existence constancy to continue the object must travel along a spatiotemporally continuous path and re-emerge from the occluder when expected (Hirsch, 1982; Spelke et al., 1995). When the occluder is a screen edge occlusion is followed by a reflexive shift in attention back across the screen. A similar *push* of attention associated with an actor's gaze-shift is interpreted as forming a perceptual enquiry (e.g. "What are they looking at?"), which forms a "bridge" between the cue and its target (Messaris, 1994). If the *push* associated with the screen edge also produces a perceptual enquiry, this may allow existence constancy to be maintained across the cut by shifting spatiotemporal expectations to the opposite screen edge. This study does not provide any evidence of existence constancy across matched-exit/entrance cuts but the constituents seem to be in place: occlusion, expectation, and saccadic suppression of the cut. Investigating

the perceptual consequences of such cuts was the objective of the next two chapters (5 and 6). What is Continuity and Accommodating Expectations.

7.2.1 Question 1: Conclusions and Further Work

The two chapters (3 and 4) summarised here provided the beginnings of an answer to the first question of this thesis: "How does continuity editing minimise awareness of a cut?" Continuity editing appears to limit awareness of editing by ensuring that insufficient attention is available for the visual discontinuity associated with the cut to be processed to the level of awareness. Continuity editing either influences attention through the use of attentional cues or identifies points during visual events when attention is not available (e.g. occupied with the processing of previous events or suppressed). Editors believe that attentional cues are a key component of the continuity style of filmmaking (Dmytryk, 1986; Murch, 2001; Pepperman, 2004) but empirical evidence that attention is exogenously controlled by film does not exist. Tentative evidence from a cross-section of research areas was presented in this section but this evidence does not provide a direct test of the effect of continuity editing on attention.

The first experiment of this thesis provided direct evidence that attention was influenced by a focal-object's motion relative to the screen edge. This influence seemed to be reflexive but this conclusion is based on an indirect measure of overt attentional shifts (e.g. correct response rates to a speeded identification task). A better measure would be eye tracking recordings of eye movements across the same matched-exit/entrance cuts. The study was also limited in the extent of conclusions that can be made due to the repetition of experimental conditions. Repetition was required to map attention across the focal-object's path but it may have resulted in faster and more direct signs of attentional shift than would be displayed under normal viewing conditions. If, in extension to this study an eye tracking experiment was performed the frequency of cuts and location to which the focal-object relocated would be varied to examine whether the attentional shift actually is reflexive and

whether it is directed towards a specific entry location or distributed across the screen.

The results of the present study indicated that the presence of a focal-object fully on-screen immediately after the cut seemed to create continuity of attention across the cut even when the cut had not been cued (i.e. when the focal-object was only half or not occluded before the cut). The possibility of attention capture across cuts was also discussed in chapter 2 and tentative eye tracking evidence that attention across subjects is synchronised was presented (May & Bannard, 1995). This possibility must be explored further both through controlled manipulations of film analogues (such as the animations used in this study) and observational studies of existing films.

If attentional cues and natural fluctuations of attention are used by editors to hide cuts, as suggested by the evidence of these two chapters (2 and 3), it would suggest that editors are especially sensitive to attentional cues. The "rules" of continuity editing may exist only as rules-of-thumb or guidelines because the real intention of the "rules" is to train editors to become sensitive to attentional cues. Once an editor has developed this sensitivity, they can then move away from the "rules" adapting each cut to the attentional dynamics of their viewer. This sensitivity to attentional cues should provide editors with an abnormal ability for predicting where people will tend to look in a dynamic visual scene at any particular point in time. Empirically establishing this ability would provide interesting support for an attentional theory of continuity editing.

In terms of applications of this attentional theory of continuity editing, it may be possible to automate the task of making cuts "invisible". Some cuts, such as match-action cuts could be automated by creating a computational system that identifies the potential for attention capture within a shot. Such systems already exist (Boccignone, Marcelli, & Somma. G., 2002; Böhme, Krause, Barth, & Martinetz, 2004; Dorr, Böhme, Drewes, Gegenfurtner, & Barth, 2005). These systems calculate *conspicuity maps* for regions of a visual scene based on such factors as changes in brightness,

colour, and motion (Boccignone et al., 2002). The most conspicuous part of a visual scene is seen as the part most likely to capture attention. This allows such a system to identify time points during a shot when attentional capture occurs. This information could then be used to cut to another shot, possibly depicting the same action from a different viewpoint, to create a match-action cut. Considerations of temporal overlap or ellipsis could be incorporated in the system based on the findings of the second study in this thesis.

A computational system for creating "invisible" cuts could provide automated assistance to an editor. Such a system would be unable to choose which shots to cut to[147], how to create causal, logical, or symbolic connections between shots, or how to modify each cut to the perceptual expectations and preferences of the viewer. However, it would be able to either identify potential edit-points which the editor could then accept or reject or "smooth" the joins between sequenced shots. Both of these tasks are very time consuming and not always seen as the best use of an editor's time[148]. Given the current state of computational vision, the creation of a fully automated editing system is unrealistic and will probably remain so for the foreseeable future.

7.3 Question 2: How does continuity editing create the perception of "continuity" across a cut?

The Hiding a Cut and Cuing a Cut chapters (2 and 3) established the potential for attentional cues to "minimise awareness of a cut" (question 1 of main questions) and the possibility that viewers maintain some form of perceptual representation (existence constancy) of the focal-object may be across a cut. This representation

[147] This would require the ability to automatically recognise and track objects within shots. This is currently computationally difficult.

[148] Which is why this task often falls to a junior or assistant editor.

may provide the perception of "continuous action" across a cut assumed by editors to be the product of continuity editing (Reisz & Millar, 1953). However, "continuity", as referred to by film editors does not seem compatible with the psychological concept of spatiotemporal continuity. An exploration of "continuity" was performed in the What is Continuity chapter (5) to attempt to identify the constituents of this concept and suggest ways by which it could be created and maintained by continuity editing.

As a first indicator of how editors use the concept "continuity", instances were identified where "continuity" is said to be absent: continuity errors. These errors refer to unexpected changes in a film, usually between shots such as a person's or object's sudden relocation within a scene. A taxonomy of continuity errors was derived and the three dimensions of "continuity"[149] identified: object, spatial, and temporal. For a viewer to perceive a cut as creating a continuity error their expectations about the actions depicted must be violated by the editing. It was suggested that viewers expect "continuity" along these three dimensions by default and it is the job of editing to change the viewer's expectations so that a change in object or omission of space or time can be accepted by the viewer. The three dimensions of "continuity" were used to identify the perceptual expectations present across eight categories of cut.

Supporting evidence for these three dimensions of "continuity" was sought in the psychological literature. A theory compatible with these three levels of "continuity" was found in Kahneman and Treisman's theory of *object files* (Kahneman & Treisman, 1984). Object files are collections of abstract conceptual information about an object's visual features, properties, and identity (Kahneman et al., 1992). Change blindness experiments have shown that viewers are quite insensitive to changes to these *object files* when the visual transients accompanying the change are obscured (Levin & Simons, 2000). The preservation of an object's *object file* is associated with the continued perception of that object, i.e. existence constancy.

[149] Referring to concept identified by editors not psychologists (n.b. "continuity" is used to refer to editor' concept whilst continuity with out quotation marks is use to refer to the psychological concept)

Developmental studies have shown that infants are insensitive to object discontinuities (e.g. when an object spontaneously changes identity) but will cease to perceive existence constancy if an occluded object fails to satisfy spatiotemporal expectations (Xu, 1999). Adults also experience existence constancy based solely on spatiotemporal continuity when an object is followed covertly (Scholl, Pylyshyn, & Franconeri, 1999) or when they are unaware of the possibility of an *object*'s discontinuity (Simons & Levin, 1997).

This dominance of spatiotemporal continuity over object continuity has been explained as due to the *visual index* (Pylyshyn, 1989) upon which the *object file* may be constructed (Kahneman et al., 1992). *Visual indices* (or FINSTs) pre-attentively segment the visual space into regions of interest (Pylyshyn, 1989). These regions are defined by their spatiotemporal continuity i.e. they move together (Wolfe & Bennett, 1997). When focal attention is first allocated to this index the visual information collocated with the index is "bound" together into an *object file* (Treisman & Gelade, 1980). This *object file* "sticks" via the *visual index* to the object as it moves about the visual scene (Leslie et al., 1998). If sight of the object is momentarily lost, such as due to occlusion, existence constancy will be perceived if the visual index can be immediately reassigned to the object after the object has returned to view (Scholl & Pylyshyn, 1999). If the object preserves spatiotemporal continuity across the visual disruption the visual index will be maintained and existence constancy will be assumed (Leslie et al., 1998). Viewers will only check if the object-level properties of the current object match those stored in the *object file* (*correspondence;* Ullman, 1979), if their attention is captured by an object discontinuity or they adopt a change detection task. Such *correspondence* does not appear to occur during film viewing (Levin & Simons, 1997).

Spatiotemporal expectations appear to be incompatible with film viewing. Film does not represent a spatially or temporally continuous world and the limited space of the screen means that motion cannot continue as it would in the real-world. However, our spatial expectations seem to be imprecise and based on relative rather than absolute locations (Deubel, 2004). Only if the saccade target cannot be found within

30ms of a saccade and within a few degrees of its expected location will information about the observer's eye movement be used to reconstruct their impression of the 3D space (Deubel & Schneider, 1994; McConkie & Currie, 1996; Hollingworth & Henderson, 2002a). This imprecision of spatial expectations combined with assumed object-level continuity provides a possible explanation of how spatial continuity can be perceived in film. The 180° Rule ensures that objects remain in roughly the same position across all cuts (Block, 2001). Attentional cues such as gaze-shifts and head turns can be used to direct attention to an object (see chapter 3) which, when a cut occurs during the attentional shift, should be perceived as located in the same position. Spatial expectations can also be satisfied by matched-exit/entrance cuts but only if the incoming screen edge is identified as the saccade target or spatial expectations are shifted by the reflexive attentional shift caused by the focal-object's occlusion by the screen edge. Evidence for such a shift of spatial expectation can be found in studies where *object files* are seen to be preserved across matched-exit/entrance cuts (Levin & Simons, 1997; Williams & Simons, 2000). If the cut violated spatiotemporal expectations the *visual index* would be lost and the *object file* erased.

However, satisfying spatial expectations across a visual disruption is not sufficient for existence constancy to be perceived: temporal expectations also have to be satisfied. Ensuring that time is perceived as continuous across a visual disruption is not just an issue of matching absolute time as our ability to perceive time is subject to distortion by attention (Zakay & Block, 1996). Perceived durations are extended by saccades and concentration (Block et al., 1980; Yarrow et al., 2001b) and shortened by distraction and divided attention (Zakay, 1992; Zakay et al., 1983b). As continuity editing has already been shown to minimise awareness of the cut by manipulating attention (chapter 3) these changes in attention should also cause distortions of perceived time. If temporal continuity is to be perceived across a cut these distortions must be accommodated by continuity editing. Evidence for such accommodation and the mechanisms by which it functions were investigated in the following chapter, Accomodating Expectations (6).

7.3.1 Question 2: Conclusions and Further Work

It was the conclusion of the 'What is Continuity?' chapter (5) the concept referred to by editors as "continuity" is equated to the perceived *existence constancy* of the focal-object across a cut. This conclusion was based on the observation that viewers seem to be sensitive to similar categories of visual information when they are watching film (identified as a result of *continuity errors*) or an object moving through real space: object, spatial, and temporal. Viewers are insensitive to object-level changes (Levin & Simons, 1997) but notice violations of spatial and temporal expectations more readily. These categories of information have been formalised in the theory of *object files* (Kahneman et al., 1992).

The preservation of an *object file* requires the associated object to move through space with spatiotemporal continuity. This chapter (5) identified the mechanism by which the *object file* is tracked through space as being a *visual index*: a pre-attentive object marker that speeds attention allocation to the associated object (Pylyshyn, 1989). This connection has previously been made by other researchers including Kahneman, Treisman, and Gibbs (Kahneman et al., 1992; Leslie et al., 1998; Scholl & Leslie, 1999). However, research into visual indices and object files are traditionally distinct so this connection is not currently established. Further investigation into both visual indices and object files is required to understand both their connection and their properties.

It is the assumption of most *existence constancy* research that the spatiotemporal continuity required for the maintenance of *existence constancy* is quite inflexible and based on real-world physics (e.g. related to the principle of inertia; Hirsch, 1982). However, the remainder of the 'What is Continuity?' chapter outlined the distortions and tolerances associated with the perception of space and time. These distortions are due to limitations of the human perceptual system as well as side effects associated with the redistribution of attention. When presented with a spatiotemporally continuous visual scene there is no guarantee that a viewer perceives continuity. As

such, to avoid the constant erasure of *object files*, viewers must be tolerant to violations of spatiotemporal expectations.

Evidence for such tolerance comes from recent evidence that explicit change blindness may actually be accompanied by signs of implicit change detection (Brockmole & Henderson, in press; Henderson & Hollingworth, 2003; Hollingworth & Henderson, 2002b; Williams et al., in press). Implicit change detection is observed when viewers show signs of being affected by a change in the visual scene but are not explicitly aware of the change (Simons, 2000). Viewers are believed to update their *object files* to accommodate the changed visual information (Hollingworth & Henderson, 2002b). A similar process of implicit change detection followed by *object file* update could explain how *existence constancy* is perceived across the discontinuities of a cut.

The continuity editing rules, such as the 180° Rule appear to ensure that an object's location on the screen is maintained across a cut and that there is usually only one centre-of-attention (see Bordwell & Thompson, 2001; pages 156-192). Eye-tracking evidence shows that this ensures both continuity and synchrony of attention across subjects (see 3.4.2.2). This may indicate that continuity editing, whilst not eradicating all signs of spatiotemporal or object discontinuity across cuts, is ensuring that any discontinuities are small and attention is still focussed directly towards the object. A *visual index* is principally a mechanism of attention that allows attention to be shifted to objects with minimal planning (Scholl & Pylyshyn, 1999). The rapid attentional shifts across cuts reported in section 3.4.2.2) and measured in chapter 4 may indicate that the visual index associated with the focal-object before the cut remains associated with the focal-object after the cut even though it has experienced a spatiotemporal discontinuity. For a Reverse-Angle cut constructed according to the 180° System the reallocation of the visual index to the focal-object in the new shot may only require a slight spatial shift (see Figure 5-4). However, in matched-exit/entrance cuts this shift covers the entire screen width.

Shifting a *visual index* across a matched-exit/entrance cut could either occur automatically or through active accommodation of the editing convention. This question of whether a film convention is learnt or the result of innate or developed cognitive ability (Hochberg, 1986) was previously seen in the background chapter 2 (see section 2.2). If the shift was due to the learning of an arbitrary convention it could be predicted that viewers would be able to adapt to an alternate convention. No sign of adaptation was observed in the first experiment even over thousands of presentation of a cut with the same exit conditions. This would seem to suggest that the matched-exit/entrance cut has an innate or developed origin and is not just a learnt convention.

In the conclusion to the first study (4.5), the rapid saccades observed across matched-exit/entrance cuts were attributed to the occlusion of the focal-object by the screen edge *pushing* attention across the cut. The subject of this *push* could actually be interpreted as the *visual index*. If the *visual index* were relocated to the opposite screen edge saccades to the opposite screen edge would be expected to be faster and the viewer should perceive *existence constancy* of the focal-object across the cut. Rapid saccades were observed in the first experiment and studies showing viewers ability to detect object-level changes across matched-exit/entrance cuts suggest that the *object file* is being preserved (Levin & Simons, 1997; Williams & Simons, 2000).

The implications of a shifting *visual index* need to be explored in future studies. In general, the assumptions of continuity editing need to be acknowledged by the field of visual cognition as they raise interesting questions about the rigidity of object and spatial perception.

7.4 Question 3: How does continuity editing ensure that "continuity" is not violated as a consequence of the cut?

The final empirical study of this thesis (chapter 6) used a duration estimation task to investigate temporal expectations across a variety of cuts: Jump Cuts, Stop-motion Cuts, and Reverse-Angle Cuts. Evidence was found in the editing literature suggesting that editors would overlap 2-5 frames of an action across a match-action cut to create perceived temporal continuity (Anderson, 1996; Dmytryk, 1986). This temporal overlap is compatible with the effects of distraction on duration perception but not saccadic eye movements (i.e. the chronostasis effect). It was hypothesised that this incompatibility could be due to a lack of evidence of the combined effect of a distracting background change and a focal-object relocation (requiring a saccadic eye movement to follow) on duration perception. This combination was investigated by estimating durations perceived by viewers as being equal to 1000ms (referred to as "matched estimates") under a variety of viewing conditions.

It was found that distraction by a sudden unexpected background change (Jump Cut) shortened perceived duration by 58.44ms on average (~1 frame at 24fps). Distraction also created a high degree of variance between the durations perceived by viewers as 1000ms. This variance may explain the editing convention of overlapping 2-5 frames: a longer period of overlap may be used to ensure that all viewers perceive all actions even if they suffer from a longer period of distraction. If viewers expected the background change before the cut no effect on perceived duration was seen. This can be attributed to attention being focussed to the focal-object allowing the viewer to ignore the background change. The same effect could be created through the use of attentional cues: a sudden change in the focal-object would focus attention and decrease the likelihood that a background change captures attention. The use of such attentional cues has been established for match-action and point-of-view shots (see chapter 3). The absence of any effect of an expected background change on

perceived duration suggests that perceived temporal continuity could be created across a collocated[150] match-action cut by presenting temporal continuity.

When an unexpected relocation of the focal-object occurred across the cut, the accompanying saccade was seen to extend perceived duration by 43.63ms (~1 frame at 24fps). The same effect occurs with or without an accompanying background change. These results indicate that relocation of the focal-object captures attention ensuring that attention cannot be captured by other visual transients associated with the cut. The visual experience of this condition can be seen as analogous with a Reverse-Angle cut created according to the 180° Rule. The use of attentional cues to shift attention between objects on the screen whilst maintaining their position across a cut has previously been established (see 3.2 and 5.3). The results of this study indicate that even if the shift of attention occurs in response to the cut, there will be no difference in how time is perceived across the cut. The temporal distortion created by the saccade is reliable so can be easily accommodated (omit 1 frame) by an editor to create the perception of temporal continuity across the cut.

The results of this study failed to validate the editing convention of overlapping 3-5 frames of an action across a Reverse-Angle cut (Dmytryk, 1986). It was hypothesised that this incompatibility between the empirical results and editing practice could be due to the absence of motion in the experimental stimuli used. The misperception of a moving object's onset position is widely established (Fröhlich, 1923). Recent evidence has suggested that when a saccade is performed to a moving object, its perceived spatial location is distorted to accommodate the perceived temporal discontinuity (Yarrow et al., 2005). The impression is of the object jumping forwards both in space and time. To compensate for this jump the object can be relocated back along its path so that the subsequent perceptual distortion shifts the object to the right location. The size of the required relocation is not as large as Dmytryk's 3-5 frames but the size of the effect could vary depending on such factors as level of attention, speed of motion, or predictability of saccade.

[150] The focal-object would have to be located in the same position across the cut so that no distortion due to the saccade is experienced.

This study also highlighted the flexibility of temporal expectations. Confidence intervals were generated for each matched estimate that indicated the range of durations perceived by subjects as being equal to 1000ms. The size of this confidence interval varied across presentation conditions. An unexpected relocation of the focal-object accompanied by a background change (i.e. a Reverse-Angle cut) produced the smallest confidence interval and unexpected background change during fixation (i.e. a Jump Cut) created the largest interval. This can be interpreted as indicating that Reverse-Angle cuts create reliable temporal distortions that are consistent across subjects where as Jump Cuts effect different subjects to different degrees. This indication of the perceptual control exerted by attentional capture (Reverse-Angle cut) compared with distraction (Jump Cut) confirms continuity editing's aversion to Jump Cuts.

7.4.1 Question 3: Conclusions and Further Work

The empirical study presented in chapter 6 investigated the question "How does continuity editing ensure that 'continuity' is not violated as a consequence of the cut?" Selecting just the temporal component of continuity perception, a duration estimation task was used to detect distortions to perceived temporal continuity across a variety of "continuity" and "discontinuity" cuts. The results indicate that expectation, distraction, and saccadic eye movements all influence the perception of temporal continuity across cuts. Expectation minimises the disruptive effects of a cut. This supports the editing convention of cutting on action. If a cut is not expected and the focal-object does not relocate across the cut, viewers will be distracted by the cut resulting in a large degree of variability in perceived duration. This makes it difficult for an editor to compensate for perceptual distortions. To avoid distractions the focal-object should change location across a cut. This unexpected relocation captures attention and leads to predictable distortions of perceived duration. This supports the editing convention of reverse-angle editing.

The most surprising finding of this study was the absence of chronostasis effect under predictable viewing conditions. Yarrow et al. have produced consistent chronostasis effects across a variety of saccade types (Yarrow et al., 2001b; Yarrow et al., 2004). This study should have also produced chronostasis when the focal-object relocation was predictable but failed to do so due to apparent variance between subject's matched estimates. This can be attributed to the experimental design which could not compensate for saccades moved in anticipation of the cut. If eye tracking was used to monitor saccades relative to the onset of the target duration, any anticipation could be accommodated by the algorithm used to derive matched estimates. A clear chronostasis effect would be expected to emerge.

However, the variability of matched estimates shown when the focal-object relocation was predictable indicates the difficulty of accommodating perceptual expectations when viewers are free to perform attention shifts whenever they choose. If, for example, a viewer chose to shift their gaze away from the focal object before a cut occurred the editor would be unable to predict what effect the cut would have on their perception of continuity. Only when all subjects move their attention in unison, such as in an unexpected Reverse-Angle cut can the editor create a single cut that accommodates the resulting perceptual distortions.

In the discussion of the 'What is continuity?' section (7.3.1), the possibility of shifting *visual indices* whilst maintaining *object files* was discussed. It was suggested that if an object underwent either a slight and insignificant spatial discontinuity, as in a Reverse-Angle cut or a large and expected spatial discontinuity, such as across a matched-exit/entrance cut the *visual index* may shift with the focal-object and *existence constancy* would be perceived. In this study large unexpected spatial discontinuities occurred with no apparent effect on perceived continuity. However, the only dimension of continuity measured in this study was temporal. There is no evidence within this study to indicate that existence constancy was perceived across the unexpected Reverse-Angle cuts.

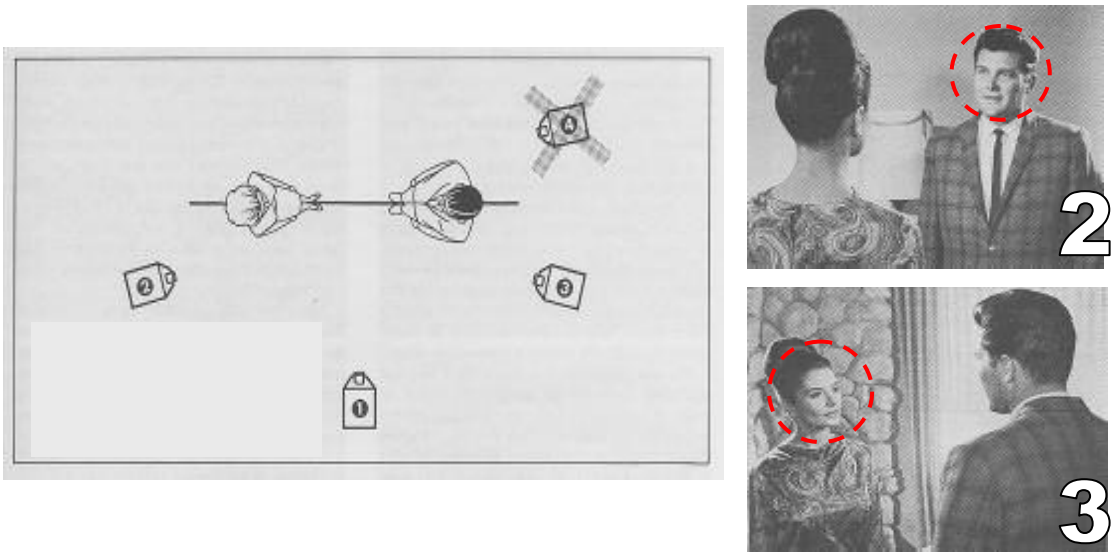


Figure 7-1: The 180° Rule (left) and its application for creating a Reverse-Angle cut (right). The cut from shot 2 to shot 3 presents the woman's face as the most salient part of the visual scene. This sudden change in visual salience will probably capture viewer's attention. The dashed circle represents the focal-object in each shot.

The Reverse-Angle cuts used in this study were intended as analogues for the type of cut depicted in Figure 7-1: Two characters, both present on screen either side of the cut with the viewer's attention shifting between characters after the cut. However, the presence of only one object on the screen in the actual stimuli may have created a different effect: apparent motion. Instead of two separate objects, the viewer may have perceived the stimuli as containing one object that moved through apparent motion. It is not known if an object's *visual index* shifts with it when it undergoes apparent motion. If the *index* did shift, combined with the perceived temporal continuity the result may be perceived *existence constancy*.

One way of testing if *existence constancy* was preserved across apparent motion would be to change the onset of the focal-object relocation relative to the beginning of the estimated duration. If the target relocation occurred at an unpredictable point during the target duration, the spatial discontinuity associated with the relocation may erase the current *object file* and create a new *object file* for the relocated object. The viewer would then be expected to perceive the target duration as starting after the new *object file* was created, not before. Yarrow et al. (2001) showed a similar result when the saccade target shifted during a saccade. Viewers performed no

perceptual "filling in" of the saccade duration; instead they perceived the shifted object as beginning to exist after the saccade. Performing the modified study proposed above would provide an interesting extension of the current study whilst also providing evidence of the tolerances of the spatiotemporal prerequisites for *existence constancy.*

This methodology could also be extended to the situation where there are two objects on the screen across the Reverse-Angle cut (Figure 7-1). The spatial continuity created by having attention shift between objects in this version of the cut would provide a control for the single-object study. This double-object version would allow the effect on *existence constancy* of object discontinuities, such as the apparent rotations caused by the cut to be investigated.

Many more possible extensions of this study exist and many questions about the perception of *existence constancy* and object, spatial, and temporal continuity remain to be answered. By laying out the methodology used in this study as a valid method for investigating such concepts it is hoped that further studies will increase our knowledge of how continuity editing creates and accommodates perceived continuity.

Chapter 8: Conclusion

Continuity editing functions by:

1. providing changes in viewpoint analogous with real-world shifts of attention (as previously thought);
2. influencing attention through the use of cues and identifying points during visual events when insufficient attention is available for the visual disruption of the cut to be processed to the level of awareness; and
3. accommodating the perceptual distortions created by these fluctuations in attention to create the perception of continuity across cuts.

This thesis presented a theoretical and empirical investigation of the cognitive foundations of continuity editing. The result of this thesis is an attentional theory of continuity editing. Attention is functional at every level of control exhibited by continuity editing: composing shots to ensure all viewers look at the same object, shifting viewers' attention about a shot, creating expectations across a cut, minimising the disruptive effects of a cut, and matching-action to ensure "continuity" after the cut. This thesis has presented existing and new empirical evidence that attention is both affected by editing and affects the perception of continuity across cuts. However, this evidence is limited in scope due to the lack of previous experiments investigating film perception. It is the intention of this thesis to motivate future empirical studies by outlining a valid theoretical and methodological foundation.

The two main assumptions of continuity editing are that its application makes a cut "invisible" (Bordwell & Thompson, 2001; Dmytryk, 1986; Reisz & Millar, 1953) and ensures that "the spectator's illusion of seeing a continuous piece of action is not interrupted" (Reisz & Millar, 1953). This thesis has equated the first function to

limiting of awareness of the cut. Continuity editing rules have been shown to have the potential to manipulate awareness of a cut through the use of attentional cues and by identifying time points relative to visual event boundaries when attention is occupied. Editing theory suggests that editors are aware of this potential and actively use it. The first study presented in this thesis provids direct evidence that attention is influenced by editing: when a cut occurs in terms of the focal-object's position relative to the screen edge. However, it is not known whether editors actually capitalise on this potential for attentional control in film. Dedicated eye-tracking studies are required to identify whether the control of attention suggested by continuity editing is actually employed by editors.

The second assumed benefit of continuity editing is that it ensures the perception of "continuity" across a cut. This is interpreted as equivalent to *existence constancy*: "the experience that objects persist through space and time despite the fact that their presence in the visual field may be discontinuous" (Butterworth, 1991). This thesis outlines ways in which continuity might minimise the spatiotemporal discontinuities that are usually thought to be incompatible with the perception of *existence constancy.* The second empirical study identifies possible compatibilities between the conventions employed by editors for matching action and the distorted perceptual expectations of spatiotemporal continuity caused by attentional shifts. However, this study was unable to establish *existence constancy* across cuts.

The current theories concerning *existence constancy* seem to be incompatible with the spatiotemporal discontinuities created by editing. However, there is evidence that *object files,* the representation maintained under *existence constancy*, are maintained across cuts (Levin & Simons, 1997; Williams & Simons, 2000). This incompatibility between the assumed benefit of continuity editing and the current theories of *existence constancy* suggest that the spatiotemporal requirements for *existence constancy* may be more flexible than previously thought. This possibility needs to be empirically investigated both to increase understanding of film perception and to clarify our experience of reality as a dynamic continuous space.

This theoretical and empirical investigation of the “rules” of continuity editing has lead to the conclusion that the creation of continuity is not about the application of rigid rules. The continuity editing rules are “rules of thumb” that provide a baseline from which adjustments can be made for the attentional and perceptual consequences of each cut. To create more precise rules every possible combinatorial effect of changes in the visual scene, changes in eye position, shifts of attention both internal and external, intention, interest, and perceptual enquiry would need to be considered. Such complexity highlights the skills employed by editors in successfully appreciating these factors. It also highlights the tolerance of the human perceptual system to accept deviations from expectation.

Not all cuts achieve the “smooth” continuity prescribed by the continuity editing rules. This should not be seen as a failing of these cuts, just an alternate intention of effect. The intentional use of degrees of discontinuity can provide the editor with control over the ease or difficulty with which their viewer processes the action presented across cuts. Such control may allow the editor to position their viewer along a continuum of attentional and perceptual activity from “reactive” to “active” search and reconstruction. The classic continuity editing rules prescribe a viewer located mid-way along such a continuum: their attention is shifted both reflexively and intentionally around shots and across cuts and their perception of continuity is based both on active perceptual enquiry and passive assumption of continuity.

Modern modifications of the continuity style, particularly the collocational editing style employed in television and recently evident in cinema as well, limits the attentional activity of the viewer by maintaining the centre of attention in one position. The perceptual consequences of such a limited composition are not known but this thesis has shown that overt attentional shifts are essential for minimising awareness of the cut and ensuring consensual perception of continuity. Without these attentional shifts the theory presented in this thesis cannot explain how continuity will be perceived.

However, the predominance of this compositional style suggests that it does not obstruct the viewer's comprehension of the presented action. This incompatibility between the prerequisites for continuity as *existence constancy* presented in this study and the collocational style used in television may suggest that other levels of continuity exist. One level may be the continuity of narrative previously described as the reason why cuts are perceptually tolerable (Bordwell & Thompson, 2001; Messaris, 1994). Viewers do not lose the ability to perceive objects after a discontinuity cut; they just have to re-perceive the objects because the *object files* have been erased by the discontinuity. However, *object files* are not the only level at which object information is stored. *Object files* are only the initial temporary representation of object information (Kahneman et al., 1992). Eventually this object information is passed into a longer term memory store that is less susceptible to disruption. If a cut fails to maintain an *object file* the characters, events, and narrative of the subsequent shot can still be comprehended based on these memories. The shot will lack a direct relationship of *existence constancy* with the previous shot, which may have consequences on the film's overall representation (as shown by Frith & Robson, 1975; Kraft, 1987), but the viewer should still be able to follow the narrative. These variable levels of continuity and their possible inter-relatedness require future investigation.

In conclusion, the attentional theory of continuity editing developed in this thesis allows an appreciation of film perception from a cognitive science perspective. The questions raised and incompatibilities between the apparent perceptual experience of film and current theories of *existence constancy* should indicate that film is a topic desperately in need of empirical investigation. Previously, technical limitations have made the empirical investigation and manipulation of film difficult but recent advances in eye tracking, computer graphics, and low-cost digital film production have eradicated most of these difficulties. It is the intention of this thesis to highlight the potential benefits, both for cognitive scientists and film makers of the empirical investigation of film perception. Maybe, as originally envisaged by Hugo Münsterberg in 1916, film will "become more than any other art the domain of the psychologist" (Münsterberg , 1970/1916: page 181).

Chapter 9: Appendices

9.1 Appendix A

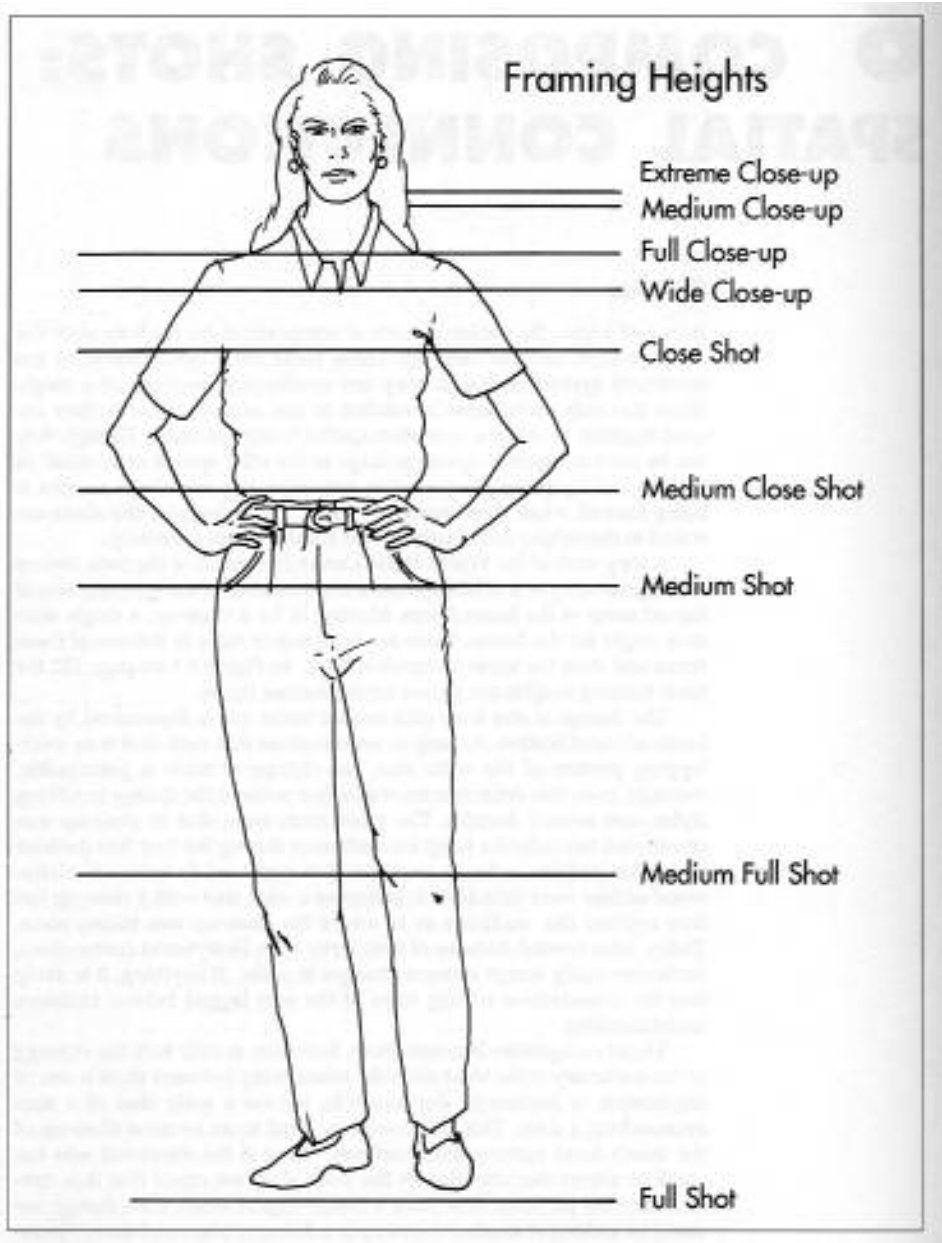


Figure 9-1: Camera-Subject Distance. Lines represent the bottom of a shot framed at a certain distance to a person. The top of all shots will be just above the person's head. (Taken from Katz, 1991)

9.2 Appendix B

Questionnaire for experiment 1

Subjects responded to the questions by selecting one of the possible responses. These are represented below as text enclosed in brackets e.g. “(1)” or “(no opinion)”. If an answer, e.g. “(Yes)*”, is followed by an asterisk the sub-questions beneath it were only revealed if the subject ticked that answer. The questions consisted of:

1. Frame-by-frame, how smooth do you think the motion of the focal-object was:
    a. in the centre of the screen?
       very jerky (1) - (2) - (3) - (4) - (5) perfectly smooth (no opinion)
    b. on the right edge of the screen?
       very jerky (1) - (2) - (3) - (4) - (5) perfectly smooth (no opinion)
    c. on the left edge of the screen?
       very jerky (1) - (2) - (3) - (4) - (5) perfectly smooth (no opinion)
2. After a cut, did you ever expect the focal-object to appear on the screen when it didn't? (Yes)* / (No)
    a. At which time point during the whole animation did this happen most? (early) - (middle) - (late) - (throughout)
    b. How often did this happen?
       (Rarely) -(Occasionally) -(Half) -(Most) –(All of the time) (Can't say)
3. Did the focal-object ever seem to jump unexpectedly from one side of the screen to the other? (Yes)* / (No)
    a. At which time point during the whole animation did this happen most? (early) - (middle) - (late) - (throughout)
    b. How often did this happen?
       (Rarely) -(Occasionally) -(Half) -(Most) –(All of the time) (Can't say)
4. In relation to the previous animations, how easy was it to respond to all the cues in this animation?
       Very difficult (1) - (2) - (3) - (4) - (5) Very easy (No opinion)

9.3 Appendix C

The set of ‘bins’ used in Experiment 2:

```
100-299, 300-499, 500-589, 590-609, 610-629, 630-649,
650-669, 670-689, 690-709, 710-729, 730-749, 750-769,
770-789, 790-809, 810-819, 820-829, 830-839, 840-849,
850-859, 860-869, 870-879, 880-889, 890-899, 900-909,
910-919, 920-929, 930-939, 940-949, 950-959, 960-969,
970-979, 980-989, 990-999, 1000-1009, 1010-1019, 1020-
1029, 1030-1039, 1040-1049, 1050-1059, 1060-1069, 1070-
1079, 1080-1089, 1090-1099, 1100-1109, 1110-1119, 1120-
1129, 1130-1139, 1140-1149, 1150-1159, 1160-1169, 1170-
1179, 1180-1189, 1190-1209, 1210-1229, 1230-1249, 1250-
1269, 1270-1289, 1290-1309, 1310-1329, 1330-1349, 1350-
1369, 1370-1389, 1390-1409, 1410-1499, 1500-1699, 1700-
1799.
```

9.4 Appendix D

		95% Confidence Interval				
Condition	Presentation Group	Lower Boundary (ms)	Mean (ms)	Upper Boundary (ms)	Interval (ms)	Rank
all	all	875.30	905.73	936.16	60.85	1
	random 1st	912.47	971.52	1030.58	118.10	14
	blocked 1st	807.31	883.64	959.96	152.66	35
	random 2nd	818.27	881.11	943.95	125.69	19
	blocked 2nd	827.33	886.64	945.96	118.63	16
1	all	894.28	927.14	959.99	65.71	4
	random 1st	911.56	990.98	1070.41	158.85	38
	blocked 1st	822.77	897.60	972.43	149.66	32
	random 2nd	854.10	914.25	974.40	120.30	17
	blocked 2nd	837.10	905.72	974.34	137.24	26
2	all	836.68	870.03	903.38	66.71	6
	random 1st	867.30	938.05	1008.80	141.49	28
	blocked 1st	759.87	830.97	902.06	142.18	29
	random 2nd	761.37	837.10	912.83	151.46	33
	blocked 2nd	809.18	874.00	938.82	129.63	21
3	all	874.33	906.07	937.80	63.47	2
	random 1st	898.72	959.90	1021.08	122.36	18
	blocked 1st	803.07	883.65	964.23	161.15	41
	random 2nd	815.66	893.67	971.68	156.02	36
	blocked 2nd	829.72	887.05	944.38	114.65	13
4	all	893.16	931.17	969.19	76.02	9
	random 1st	928.36	1007.70	1087.04	158.67	37
	blocked 1st	825.83	919.05	1012.27	186.44	50
	random 2nd	854.57	920.40	986.23	131.66	23
	blocked 2nd	794.52	877.55	960.58	166.06	42
5	all	893.79	934.81	975.82	82.03	11

	random 1st	969.35	1049.43	1129.52	160.17	39
	blocked 1st	808.47	897.97	987.46	178.99	48
	random 2nd	836.36	901.33	966.31	129.95	22
	blocked 2nd	802.02	890.50	978.98	176.97	47
6	all	859.50	892.00	924.49	64.98	3
	random 1st	888.59	942.40	996.21	107.63	12
	blocked 1st	790.73	877.80	964.87	174.15	45
	random 2nd	798.97	867.27	935.57	136.60	25
	blocked 2nd	804.60	880.52	956.44	151.84	34
7	all	850.39	885.25	920.10	69.71	8
	random 1st	866.51	938.58	1010.65	144.14	31
	blocked 1st	805.83	896.53	987.23	181.40	49
	random 2nd	768.83	840.05	911.27	142.44	30
	blocked 2nd	796.69	865.82	934.94	138.24	27
8	all	881.82	914.99	948.16	66.34	5
	random 1st	901.38	960.53	1019.68	118.30	15
	blocked 1st	798.95	885.30	971.65	172.71	44
	random 2nd	833.29	901.32	969.35	136.06	24
	blocked 2nd	832.30	912.82	993.34	161.04	40
9	all	856.49	890.11	923.73	67.24	7
	random 1st	915.46	956.13	996.81	81.34	10
	blocked 1st	777.84	863.87	949.90	172.06	43
	random 2nd	766.77	854.60	942.43	175.67	46
	blocked 2nd	821.51	885.83	950.16	128.65	20

9.5 Appendix E

The rest of the Appendices are conference abstracts:

Presentation at the Cognitive Studies of the Moving Images Conference (CCSMI), July 23rd, 2004

Author: Tim J. Smith
Title: EDITING ATTENTION: the perceptual foundations of continuity editing.

Abstract: Continuity editing uses established rules-of-thumb that are believed to take advantage of "assumptions" intrinsic to the human perceptual system. For example, it is assumed that the visual world is a constant, chronological space in which changes do not occur instantaneously. Therefore, an artificial visual world (e.g. cinema) will be compatible with these assumptions as long as the perceptual system is not alerted to a violation of expectation (e.g. a sudden change). To investigate the validity of this claim an experiment was conducted to show how perception of an edit can be masked by the absence of attention occurring over an eye movement hence creating the illusion of continuity.

Subjects were shown a series of animations depicting an object leaving the right of the screen and re-entering on the left, as if its continuous motion had been filmed by two cameras side-by-side. Exactly when, where, and how the cut occurs were controlled as independent variables. The hypotheses were based upon previous findings in the areas of visual occlusion, eye movement, and attention. It was found that the factors that created the highest level of continuous attention were those that accommodated the eye movements necessary to track the moving object across the edit. An edit should occur as soon as the object has completely left the screen and the action should be overlapped by at least two frames to allow the reorienting of attention. These findings are in direct accordance with match-action continuity rules.

9.6 Appendix F

Poster presented at the **Cognitive Science Conference**, Chicago, August 5-7th 2004

*Attached to print version. Abstract can be viewed at:

http://www.cogsci.northwestern.edu/cogsci2004/ma/ma287.pdf

9.7 Appendix G

Presentation at Society for Cinema and Media Studies Conference
London, UK. March 31-April 3, 2005

Author: Tim J. Smith

Title: Editing Time: an empirical investigation of time perception across match-action cuts.

Abstract When editing together two shots of an action filmed from different viewpoints it is common practice to overlap the action by a couple of frames to create the illusion of temporal continuity. It was previously believed that this overlap was necessary due to the "masking" of the last few frames of action by the visual information of the new shot (Anderson, 1996). The intention behind this study was to empirically investigate this claim and gain a greater understanding of the perception of time across a match-action edit.

Previous investigations of time perception have shown that subjects under-estimate the duration of a visual event when their attention is distracted during the event (Block, George & Reed, 1980) and over-estimate the duration when they perform an eye movement during the event (Yarrow et al, 2001). However, the effects of distraction and eye movements have previously only been investigated in isolation using simple stimuli that is not analogous with the rich visual experience of cinema. Therefore, it is the aim of this study to show that it is the interaction of these effects in visually rich scenes that permits the perception of temporal continuity from discontinuous visual information.

The empirical investigation of film perception is methodologically difficult. Applying a sound empirical method requires the precise manipulation of film content and editing whilst controlling any confabulating factors such as sound and lighting. For this study a new experimental paradigm was devised which could automatically

generate visually complex animations in real-time and manipulate the editing according to viewer response.

In this study it was found that viewers perceive visual events as being longer when the event is presented immediately following an edit in which the centre of attention shifts across the screen (necessitating a saccadic eye movement to re-fixate it). When the camera rotates around the centre of attention during a cut without an accompanying translation (creating a 'Jump Cut'), the viewer perceives the visual event as being shorter than when the same camera rotation occurs without distracting background information. These results allow us to increase the precision of current continuity editing rules: when editing together two shots of an action filmed from different viewpoints, the action can be made to appear temporally continuous by overlapping one frame (41.7ms) of the action when the screen location of the principle object doesn’t change, or omitting one frame when it does.

9.8 Appendix H

Poster presented at European Conference on Visual Perception
A Coruna, Spain 22-26th August 2005

Author: Tim J. Smith

Title: The screen edge as an occluder: expectation of existence constancy during spatiotemporally discontinuous motion.

Abstract: When an object gradually moves behind another object in the visual scene (occlusion) the occluded object is perceived as continuing to exist (existence constancy) and to continue moving in the same direction and speed (spatiotemporal continuity). However, when the occluder is the edge of a television screen the object is not expected to reappear outside of the screen. Instead conventional film technique would relocate the object to the opposite screen edge as a "cut" is made to an adjacent camera, violating the normal expectation of spatiotemporal continuity during occlusion. This study investigated whether viewers expect the object to continue to exist after leaving the screen by showing preference for gradual occlusion over sudden disappearance. This would suggest a dissociation of existence constancy and spatiotemporal continuity.

Subjects were required to respond to a binary-decision reaction time (RT) cue overlaid on an object as it moved across the screen, relocating from one edge to the other on exit. The RT task was used to estimate visual attention. When the object relocation is predictable a saccadic eye movement will occur that coincides with and mirrors the object relocation. Preparation for the saccade can be observed as a ~100ms withdrawal of attention prior to the saccade.

In this study it was found that subjects only withdrew attention in preparation for a relocation that occurred after the object was fully occluded by the screen edge. If relocation occurred when the object was touching the edge or only half occluded no

preparation was observed and recovery of full attention following the relocation took significantly longer.

These results indicate that viewers expect an object to satisfy the requirements for existence constancy (gradual occlusion) even though the occluder, the screen edge, violates assumptions of spatiotemporal continuity. This suggests a dissociation of existence constancy and spatiotemporal continuity.

9.9 Appendix I

Poster to be presented at the **Eye Tracking Research & Applications (ETRA)** conference, March 27-29, 2006. San Diego, CA.

*Attached to print version. Can be viewed at:
http://homepages.inf.ed.ac.uk/s9732397/stuff/smith_whitwell_lee.pdf

Chapter 10:Bibliography

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. Psychological Science, 14(427-432).

Allport, A. (1989). Visual attention. M. I. Posner (Ed.), Foundations of Cognitive Science (pp. 631-682). Cambridge, MA: MIT Press .

Anderson, D., & Levin S. (1976). Young childrens attention to Sesame Street. Child Development, 47, 806-811.

Anderson, J. (1996). The Reality of Illusion: An Ecological Approach to Cognitive Film Theory. Southern. Illinois University Press.

Anderson, J., & Anderson, B. (1985). Motion Perception in Motion Pictures. T. De Lauretis, & S. Heath (Eds.), The Cinematic Apparatus (pp. 76-95). Palgrave Macmillan.

Andreassi, J. L. (1995). Psychophysiology: Human Behaviour and Physiological Response. Hillsdale, NJ: Lawrence Erlbaum Associates.

Angelone, B. L., Levin, D. T., & Simons, D. J. (2003). The relationship between change detection and recognition of centrally attended objects in motion pictures. Perception, (32), 947-962.

Arnheim, R. (1957). Film as Art. Berkeley, California: University of California Press.

Assad, J., & Maunsell, J. H. R. (1995). Neuronal correlates of inferred motion in primate posterior cortex. Nature, 373, 518-521.

Bachmann, G. (1965). Interview with John Huston. Film Quarterly, 19(1).

Baird, J. A., & Baldwin, D. A. (2001). Making sense of human behavior: action parsing and intentional inference. B. F. Malle, Moses L. J. , & D. A. Baldwin (Ed.), Understanding Intentions and Intentionality: Foundations of Social Cognition . Cambridge, MA: MIT Press.

Baker, C. I., Keysers, C., Jellema, T., & Wicker, B. (2001). Neuronal representation of disappearing and hidden objects in the temporal cortex of the macaque. Experimental Brain Research, 140, 375-381.

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, A. (2001). Infants parse dynamic action. Child Development, (*72*), 655-948.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. Behavioral and Brain Sciences, 20, 723-767.

Baron-Cohen, S. (1995). Mindblindness: An essay on Autism and Theory of Mind. Cambridge, New England: MIT Press.

Becklen, R., & Cervone, D. (1983). Selective looking and the noticing of unexpected events. Memory & Cognition, 11, 601–608.

Bentivoglio, A. R., Bressman, S. B., Cassetta, E., Carretta, D., Tonali, P., & Albanese, A. (1997). Analysis of blink rate patterns in normal subjects. Movement Disorder, 12(6), 1028-1034.

Berlyne, D. E. (1971). Aesthetics and Psychobiology. New York: Meredith Corporation.

Block, B. (2001). The Visual Story: Seeing structure of film, TV, and New Media. MA, USA: Focal Press.

Block, R. A. (1978). Remembered duration: Effects of event and sequence complexity. Memory and Cognition, 6, 320-26.

Block, R. A. (1982). Temporal judgements and contextual change. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 530-44.

Block, R. A. (1990). Models of Psychological Time. R. A. Block (editor), Cognitive models of psychological time (pp. 1-35). Hillsdale, NJ: Erlbaum.

Block, R. A., George, E. J., & Reed, M. A. (1980). A Watched pot sometimes boils: a study of duration experience. Acta Psychologica, 46, 81-94.

Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for contextual-change hypothesis. Journal of Experimental Psychology: Learning, Memory, and Cognition, 4, 656-65.

Boccignone, G., Marcelli, A., & Somma. G. (2002). Analysis of dynamic scenes based on visual attention. Proceedings of AIIA 2002 .

Bolls, P., Hibbs, H., & Lang, A. (1995). A message is a message is a message....Structure predicts memory for random television messages. Mass Communication Interest Group of the Western Speech Communication Association .

Bolls, P., Potter, R., & Lang, A. (1996). Television arousal and memory: The effects of production pacing and arousing content on encoding, storage, and retrieval of television messages. Online proceedings of the 1996 Midwest Artifical Intelligence and Cognitive Science Conference .

Boot, W. R., Brockmole, J. R., & Simons, D. J. (in press). Attention capture is modulated in dual-task situations. Psychonomic Bulletin & Review.

Bordwell, D., Staiger, J., & Thompson, K. (1985). The Classical Hollywood Cinema: Flm Style & Mode of Production to 1960. London, UK: Routledge.

Bordwell, D. (1997). On the History of Film Style. Cambridge, Massachusetts, USA: Harvard University Press.

Bordwell, D., & Thompson, K. (2001). Film Art: An introduction 6th ed.,). New York, USA: Mc Graw Hill.

Bottomore, S. (1990). Shots in the Dark - The Real Origins of Film Editing. T. Elsaesser (Editor), Early cinema: space - frame - narrative . London, UK: British Film Institute.

Bower, T. G. R. (1967). The development of object-permanence: Some studies of existence constancy. Perception & Psychophysics, 2(9), 411-418.

Bower, T. G. R. (1974). Development in infancy. San Francisco, CA: W. H. Freeman.

Branigan, E. (1984). Point of View in the Cinema: A Theory of Narration and Subjectivity in Classical Film. Berlin: Mouton.

Breitmeyer, B. G. (1984). Visual Masking: An integrative approach. New York, USA: Oxford University Press.

Bristow, D., Haynes, J.-D., Sylvester, R., Frith, C. D., & Rees, G. (2005). Blinking Suppresses the Neural Response to Unchanging Retinal Stimulation. Current Biology, 15, 1296-1300.

Brockmole, J. R., & Henderson, J. M. (2005). Attentional prioritization of new objects in real-world scenes: Evidence from eye movements. Journal of Experimental Psychology: Human Perception and Performance, 31, 857-868.

Brockmole, J. R., & Henderson, J. M. (in press). Object appearance, disappearance, and attention prioritization in real-world scenes. Psychonomic Bulletin & Review.

Brown, S. W. (1995). Time, change and motion: The effects of stimulus movement on temporal perception . Perception and Psychophysics, 57(1), 105-116.

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. Perception & Psychophysics, 59(7), 1118-1140.

Bruce, V., & Young, A. (1998). In the Eye of the Beholder: the science of face perception. New York: Oxford University Press.

Burch, N. (1990). Life to those shadows. London, UK: British Film Institute.

Burr, D. C., Holt, J., Johnstone, J. R., & Ross, J. (1982). Selective depression of motion sensitivity during saccades. Journal of Physiology, 333, 1-15.

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular pathway during saccadic eye movements. Nature, (371), 511-513.

Burr, D. (2004). Eye Movements: Keeping Vision Stable. Current Biology, 14, 195-197.

Burr, D. (2005). Vision: in the Blink of an Eye. Current Biology, 14(15).

Butterworth, G. (1991). Phenomenal Permanence. G. Thines, A. Costall, & G. Butterworth (Eds.), Michotte's experimental phenomenology of perception (pp. 117-122). Hillsdale, NJ: Lawrence Erlbaum.

Böhme, M., Krause, C., Barth, E., & Martinetz, T. (2004). Eye Movement Predictions Enhanced by Saccade Detection. Brain Inspired Cognitive Systems 2004 .

Campbell, F. W., & Wurtz, R. H. (1978). Saccadic ommission: why we do not see a greyout during a saccadic eye movement. Vision Research, 18, 1297-1303.

Campbell, S. S., & Birnbaum, J. M. (1994). Time flies when you're cool: Relationship between bodz temperature and estimated interval duration. Sleep Research, 23, 483.

Carpenter, D. L., & Wojtaszcyk, J. A. (2002). The Effects Of Enjoyment On Perceived Duration: Limits To The Temporal Processing Model. 73rd Annual Meeting of the Eastern Psychological Association .

Carroll, J. M., & Bever, T. G. (1976). Segmentation in Cinema Perception. Science, 191(4231), 1053-1055.

Churchland, M. M., Chou, I.-H., & Lisberger, S. G. (2003). Evidence for Object Permanence in the Smooth-Pursuit Eye Movements of Monkeys. Journal of Neurophysiology, 90, 2205-2218.

Comuntzis, G. M. (1987). Children's Comprehension of Changing Viewpoints in Visual Presentations. Visual Communication Conference .

Comuntzis-Page, G. (1991). Perspective-taking theory: shifting views from Sesame Street. Visual Communication Conference .

Coren, S., Ward, L. M., & Enns, J. T. (1993). Sensation and Perception (4th ed.). New York: Harcourt Brace College.

Coull, J. T., Vidal, F., Nazarian, B., & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. Science, 303, 1506-1508.

Croy, H. (1918). How Motion Pictures Are Made.

Currie, C. B., McConkie, G. W., Carlson-Radvansky, L. A., & Irwin, D. E. (2000). The role of the saccade target object in the perception of a visually stable world. Perception & Psychophysics, 62, 673-683.

Cutting, J. (2004). Perceiving Scene in Film and in the World. J. D. Anderson, & B. F. Anderson (Eds.), Moving image theory: ecological considerations .

d'Ydewalle, G., Desmet, G., & Van Rensbergen, J. (1998). Film Perception: The processing of film cuts. G. Underwood (Ed.), Eye guidance in reading and scene perception (pp. 357-367). Oxford: Elsevier.

d'Ydewalle, G., & Vanderbeeken, M. (1990). Perceptual and cognitive processing of editing rules in film. R. Groner, G. d'Ydewalle, & R. Parnham (Eds.), From eye to mind: Information acquisition in perception, search, and reading. (pp. 129-139). Amsterdam: Elsevier (North Holland).

Dennett, D. C. (1991). Consciousness Explained. Boston: Little, Brown.

Deubel, H., & Schneider, W. X. (1994). Can man bridge a gap? Behavioural and Brain Sciences, 17, 259-260.

Deubel, H. (2004). Localisation of targets across saccades: Role of landmark objects. Visual Cognition, 11(2-3), 173-202.

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal Control of Saccadic Suppression. The Journal of Neuroscience, 20(9), 3449-3455.

Dmytryk, E. (1986). On Filmmaking. London, UK: Focal Press.

Dodge, R. (1900). Visual perception during eye movements. Psychological Review, 7, 454-465.

Dorr, M., Böhme, M., Drewes, J., Gegenfurtner, K. R., & Barth, E. (2005). Variability of eye movements on high-resolution natural videos. TWK 2005 Proceedings of the 8th Tübinger Perception Conference .

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Mawell, E., & Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. Visual Cognition, 6 (5), 509-540.

Duhamel, J.-R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. Science, 255(90-92).

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion Integration and Postdiction in Visual Awareness. Science, 287, 2036-2038.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. Annual Review of Psychology, 48, 269-297.

Enns, J. T., Austen, E. L., Di Lollo, V., Rauschenberger, R., & Yantis, S. (2001). New objects dominate luminance transients in setting attentional priority. Journal of Experimental Psychology: Human Perception and Performance , 27, 1287-1302.

Eriksen, C. W., & Hoffman, J. E. (1972). Temporal and spatial characteristsics of seöective encoding from visual displays. Perception & Psychophysics, 12, 201-204.

Eriksen, C. W., & St. James, J. D. (1986). Visual attenton within and around the field of focal attention: a zoom lens model. Perception & Psychophysics, 40, 225-240.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns rather than novel ones. Science, 146, 668-670.

Faraday, P., & Sutcliffe, A. (1997). Multimedia: design for the "moment" . Proceedings of the fifth ACM international conference on Multimedia (pp. 183-192). ACM Press.

Flombaum, J. I., & Scholl, B. J. (in press). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. Journal of Experimental Psychology: Human Perception and Performance.

Fogarty, C., & Stern, J. A. (1989). Eye movements and blinks: their relationship to higher cognitive processes. International Journal of Psychophysiology, 8(1), 35-42.

Folk, C. L., & Gibson, B. S. (2001). Attraction, distraction, and action: Multiple perspectives on attentional capture. Amsterdam: Elsevier.

Folk, C. L., & Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture. Journal of Experimental Psychology: Human Perception and Performance , 24, 847-858.

Folk, C. L., Remington, R. W., & Johnstone, J. C. (1992). Involuntary covert orienting on attentional control settings. Journal of Experimental Psychology: Human Perception and Performance , 18, 1030-1044.

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onsets, and colour. Journal of Experimental Psychology: Human Perception and Performance , 20, 317-329.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary Covert Orienting Is Contingent on Attentional Control Settings. Journal of Experimental Psychology: Human Perception and Performance, 18(4), 1030-1044.

Forster, K. I. (1979). Visual perception of rapidly presented word sequences of varying complexity. Perception & Psychophysics, 8, 215-221.

Fraisse, P. (1963). The Psychology of Time. New York: Harper and Row.

Franconeri, S. L., Hollingworth, A., & Simons, D. J. (2005). Do new objects capture attention? Psychological Science, 16(4), 275-281.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. Perception & Psychophysics, 65, 999-1010.

Franconeri, S. L., Simons, D. J., & Junge, J. A. (in press). Searching for stimulus driven shifts of attention. Psychonomic Bulletin & Review.

Freyd, J. J., & Finke, R. A. (1984). Representational Momentum. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10(1), 126-132.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it!: reflexive orienting is triggered by non-predictive gaze. Psychonomic Bulletin and Review, 5, 490-495.

Frith, U., & Robson, J. E. (1975). Perceiving the language of films. Perception, 4(1), 97-103.

Fröhlich, F. W. (1923). Über die Messung der Empfindungszeit. Zeitschrift Für Sinnesphysiologie, 54, 58-78.

Fukuda, K., Stern, J. A., Brown, T. B., & Russon, M. B. (2005). Cognition, blinks, eye-movements, and pupillary movements during performance of a running memory task. Aviation, Space and Environmental Medical, 76(7 Suppl.), C75-85.

Geiger, S., & Reeves, B. (1993). The effects of scene changes and semantic relatedness on attention to television. Communication Research, 20(155-175).

Gibson, B. S., & Kelsey, E. M. (1998). Stimulus-driven attentional capture is contingent on attentional set displaywide visual features. Journal of Experimental Psychology: Human Perception and Performance , 24(699-706).

Gibson, J. J. (1979). The ecological approach to visual perception. Boston, USA: Houghton Mifflin.

Gibson, J. J., Kaplan, G. A., Reynolds, H. N. Jr., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. Perception & Psychophysics, 5(2), 113-115.

Goldberg, M. E., Eggers, H. M., & Gouras, P. (1991). Ch 43: The Ocular Motor System. E. R. Kandel, J. H. Schwartz, & T. H. Jessel (Eds), Principles of neural science (3rd ed.,). Norwalk, CT: Appleton & Lange.

Gregory, J. R. (1961). Some psychological aspects of motion picture montage.

Grimes, J. (1996). On the failure to detect changes to scenes across saccades. K. Akins (editor), Perception (Vanvouver studies in Cognitive Science) volume 2. (pp. 89-110). New York: Oxford University Press .

Grüsser, O. J., Krizic, A., & Weis, L. R. (1987). Afterimage movement during saccades in the dark. Vision Research, 27, 215-226.

Gullberg, M., & Holmqvist, K. (1999). An Eye on Gesture. Pragmatics & Cognition , 7(1), 35-63.

Haimson, C., & Behrmann, M. (2001). Cued visual attention does not distinguish between occluded and occluding objects. Psychonomic Bulletin & Review, 8(3), 496-503.

Helmholtz, H. v. (1867-1925). Treatise on physiological optics from 3rd German edition, Trans) ed.). New York: Dover Publications.

Henderson, J. M. (1994). Two representational systems in dynamic visual identification. Journal of Experimental Psychology: General, 123, 410-426.

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. Psychological Science, 8, 51-55.

Henderson, J. M., & Anes, M. D. (1994). Effects of object-file review and type priming on visual identification within and across eye fixations. Journal of Experimental Psychology: Human Perception and Performance, 20, 826-839.

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. Psychological Science, 10(5), 438-443.

Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: detecting changes to saccade targets in scenes. Perception & Psychophysics, 65(1), 58-71.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. Science, 143, 1190-1192.

Hirsch, E. (1982). The Concept of Identity. Cambridge, MA: Harvard University Press.

Hirsch, I. J., & Sherrick, C. E. (1961). Perceived order in different sense modalities. Journal of Experimental Psychology, 62, 423-432.

Hitchcock, A. (1995). Hitchcock on HitchcockFaber and Faber.

Hobbs, R., & Frost, R. (1988). How first-time viewers comprehend editing conventions. Journal of Communication, 38(4), 50-60.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of Perception and Human Performance (Vol. I: Sensory Processes and Perception). New York: Wiley.

Hochberg, J., & Brooks, V. (1978a). Film Cutting and Visual Momentum. J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), Eye Movements and the Higher Psychological Functions (pp. 293-317). Hillsdale, NJ: Lawrence Erlbaum.

Hochberg, J., & Brooks, V. (1978b). The Perception of Motion Pictures. E. C. Carterette, & M. P. Friedman (Eds.), Handbook of Perception, Vol. X . New York: Academic Press.

Holland, M. K., & Tarlow, G. (1975). Blinking and Thinking. Perceptual Motor Skills , 41(2), 503-506.

Hollingworth, A., & Henderson, J. M. (2002a). Accurate visual memory for previously attended objects in natural scenes. Journal of Experimental Psychology: Human Perception and Performance, 28, 113-136.

Hollingworth, A., & Henderson, J. M. (2002b). Sustained change blindness to incremental scene rotation: a dissociation between explicit change detection and visual memory. Michigan State University Eye Movement Laboratory Technical Report

Hood, B. M. (1998). An eye direction detector triggers shifts of visual attention in human infants. Psychological Science, 9, 131-134.

Irwin, D. E. (1991). Information integration across saccadic eye movements. Cognitive Psychology, 23, 420-456.

Irwin, D. E. (1992). Visual memory within and across fixations. K. Rayner (Ed.), Eye movements and visual cognition: Scene perception and reading (pp. 146-165). New York : Springer-Verlag.

Ivry, R. B., & Hazeltine, R. E. (1992). Models of timing-with-a-timer. F. Macar, V. Pouthas, & W. J. Friedman (editors), Time, action and cognition: Towards bridging the gap (pp. 183-189). Dordrecht, Netherlands: Kluer Academy.

Jesionowski, J. E. (1982). Thinking in Pictures: Dramatic Structure in D. W. Griffith's Biograph films. Berkeley and Los Angeles: University of California Press.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organisation of visual short-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 683-702.

Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. Proceedings of the National Academy of Sciences of the United States of America, 100(18), 10568-10573.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and response to time. Psychological Review, 96(3), 459-91.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. T. Field, & N. Fox (Eds.), Attention and Performance Vol. IX (pp. 187-203). 1981: Erlbaum.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. D. A. Davies (Ed.), Varieties of attention . New York: Academic Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. Cognitive Psychology, 24(2), 175-219.

Kaplan, G. A. (1969). A psychophysical study of occlusion. Perception & Psychophysics.

Katz, S. D. (1991). Film Directing Shot by Shot: visualizing from concept to screen. CA, US: Michael Wiese Productions.

Kellman, P. J., & Shipley, T. F. (1991). A theory of object interpolation in object perception. Cognitive Psychology, 23(2), 141-221.

Kleinke, C. L. (1986). Gaze and eye contact: a research review. Psychological Review, 100, 78-100.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002a). Defining and quantifying the social phenotype in autism. American Journal of Psychiatry , 159 (6), 895-908.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002b). Visual Fixation Patterns During Viewing of Naturalistic Social Situations in Individuals with Autism. Arch Gen Psychiatry, 59, 809-816.

Koffka, K. (1935). Principles of Gestalt psychology. New York: Harcourt Brace.

Korte, A. (1915). Kinematoskopische Untersuchungen. Zeitschrift Fur Psychologie, 72, 193-206.

Kraft, R. N. (1987). Rules and strategies of visual narratives. Perceptual and Motor Skills, 64, 3-14.

Kuhn G., & Tatler, B. W. (2005). Magic and fixation: Now you don't see it, now you do. Perception, 34, 1155-1161.

Lamme, V. A. F. (2003). Why Visual Attention and Awareness are different. Trends in Cognitive Science, 7(1), 12-18.

Lamont, P., & Wiseman, R. (1999). Magic in Theory. Hartfield: Hermetic Press.

Lang, A., Zhou, S., Schwartz, N., Bolls, P. D., & Potter, R. F. (2000). The Effects of Edits on Arousal, Attention, and Memory for television Messages: When an Edit is an Edit Can An Edit Be Too Much? Journal of Broadcasting and Electronic Media, 44, 94-109.

Lang, A. (1990). Involuntary attention and physiological arousal evoked by structural features and mild emotion in TV commercials. Communication Research, 17, 275-299.

Lang, A. (1991). Emotion, formal features, and memory for televised political advertisements. F. Biocca (Ed.), Television and political advertising, Vol. I: Psychologicyl processes. Hillsdale, NJ: Erlbaum.

Lang, A. (2000). The Limited Capacity model of Mediated Message Processing. Journal of Communication, 50, 46-70.

Lang, A., & Basil, M. D. (1998). Attention, resource allocation, and communication research: What do secondary task reaction times measure anyway? M. Roloff (Ed.), Communication Yearbook 21 . Beverley Hills, CA: Sage.

Lang, A., Geiger, S., Strickwerda, M., & Sumner, J. (1993). The effects of related and unrelated cuts on viewers memory for television: A limited capacity theory of television viewing. Communication Research, 20, 4-29.

Lang, P. J., Simons, R. F., & Balaban, M. (1997). Attention and Orienting: Sensory and motivational processes. Hillsdale, NJ: Erlbaum.

Langton, S. R. H., & Bruce, V. (2000). You *must* see the point: automatic processing of cues to the direction of social attention. Journal of Experimental Psychology: Human Perception and Performance, 26(2), 747-57.

Langton, S. R. H., O'Malley, C., & Bruce, V. (1996). Actions speak louder than words: symmetrical crossmodal interference effects in the processing of verbal gestural information. Journal of Experimental Psychology: Human Perception and Performance, 22, 1357-1375.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. Trends in Cognitive Sciences, 4(2), 50-59.

Lappe, M., Awater, H., & Krekelberg, B. (2000). Postsaccadic visual references generate presaccadic compressions of space. Nature, 403, 892-895.

Lawrence, D. H. (1971). Two studies of visual search for word targets with controlled rates of presentation. Perception & Psychophysics, 10, 85-89.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: developing 'what' and 'where' systems. Trends in Cognitive Sciences, 2(1), 10-18.

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. Perception, 13, 287-305.

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. Psychonomic Bulletin and Review, 4, 501-506.

Levin, D. T., & Simons, D. J. (2000). Perceiving Stability in a Changing World: Combining Shots and Integrating Views in Motion Pictures and the Real World. Mediapsychology, 2, 357-380.

Levin, D. T., & Varakin, D. A. (2004). No pause for a brief disruption: Failures of visual awareness during ongoing events. Consciousness and Cognition, 13(2), 363-372.

Lindgren, E. (1948). The Art of the Film. London, UK: George Allen & Unwin Limited.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not recognised during the attentional blink. Nature, 383, 616-618.

Lynn, R. (1966). Attention, Arousal, and the Orienting Reaction. London: Pergamon Press.

Macar, F., Lejeune, H. F., Bonnet, M., Ferrara, A., Pouthas, V., & Maquet, P. (2002). Acivation of the supplementary motor area and of attentional networks during temporal processing. Experimental Brain Research, 142, 539-550.

Mack, A., & Rock, I. (1998). Inattentional Blindness. Cambridge, MA: MIT Press.

Macromedia Inc. (2004). Macromedia Director MX 2004 (Version 10.1) : Macromedia Inc.

Matin, E. (1974). Saccadic suppression: A review and an analysis. Psychological Bulletin, 81(12), 899-917.

May, J., & Bannard, P. (1995). Cinematography and Interface Design. K. Nordby, P. H. Helmersen, D. J. Gilmore, & S. A. Arnesen (Eds), Human-Computer Interaction: Interact'95 (pp. 26-31).

May, J., Dean, M. P., & Barnard, P. J. (2003). Using Film Cutting Techniques in Interface Design. Human-Computer Interactions, 18, 325-372.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. Journal of Experimental Psychology: General, 114, 159-188.

McConkie, G. W., & Currie, C. B. (1996). Visual-stability across saccades while viewing complex pictures. Journal of Experimental Psychology: Human Perception and Performance , 22, 563-581.

Messaris, P. (1994). Visual Literacy: Image, Mind, and Reality. Boulder: Westview Press.

Michon, J. A. (1985). The compleat time experiencer. J. A. Michon, & J. L. Jackson (editors), Time, mind, and behavior (pp. 20-52). Berlin: Springer-Verlag.

Michotte, A. (1955). Perception and cognition. Acta Psychologica, 11, 69-91.

Michotte, A. (1963). The Perception of Causality. Andover, hants.: Methuen.

Michotte, A. (1991). On phenomenal permanence: facts and theories. G. Thines, A. Costall, & A. Butterworth (Editor), Michotte's experimental phenomenology of perception (pp. 122-139). Hillsdale, NJ, US: Erlbaum. Notes: (Original work published 1950)

Michotte, A., Thines, G., & Crabbe, G. (1991). Amodal completion of perceptual structures. G. Thines, A. Costall, & A. Butterworth (Editors), Michotte's experimental phenomenology of perception (pp. 140-167). Hillsdale, NJ, US: Erlbaum. Notes: (Original work published 1964)

Miller, G. (2005). Reflecting on Another's Mind. Science, 308(5724), 945-947.

Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The relationship between object files and conscious perception. Cognition, 96, 67-92.

Mitroff, S. R., Scholl, B. J., & Wynn, K. (under review). One plus one equals one: The effects of merging on object files. Manuscript Submitted for Publication.

Mitroff, S. R., Simons, D. J., & Levin, D. T. (2004). Nothing compares two views: Change blindness can occur despite preserved access to the changed information. Perception & Psychophysics, 66, 1268-1281.

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. Psychological Science, 9, 104-110.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. Nature Neuroscience, 8(7), 950-954.

Most, S. B., Scholl, B. J., Simons, D. J., & Clifford, E. R. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. Psychological Review, 112(1), 217-242.

Murch, W. (2001). In The Blink Of An Eye: a perspective on film editing. Los Angeles, USA: Silman-James Press.

Münsterberg, H. (1970). The film: a psychological study; The silent photoplay in 1916 . New York, USA: D. Appleton and Company.

Müsseler, J., & Aschersleben, G. (1998). Localising the first position of a moving stimulus: the Fröhlich effect and an attention-shifting explanation. Perception & Psychophysics, 60, 683-695.

Neisser, U. (1979). The control of information pickup in selective looking. A. D. Pick (Ed.), Perception and its Development: A Tribute to Eleanor J. Gibson (p. 201–219). Erlbaum .

Neisser, U., & Becklen, R. (1975). Selective looking: attending to visually specified events. Cognitive Psychology, 7, 480–494.

Nelmes, J. (2003). An Introduction to Film Studies (3rd ed.). London, UK: Routledge.

Nesthus, T. E., & Stern, J. A. (2002). Ocular correlates of fatigue. Aviation, Space and Environmental Medical, 73(4), 414.

Nevarez, H. G., & Scholl, B. J. (in preparation). Why so slow?: The role of speed discontinuities in maintaining object persistence through occlusion. Manuscript in Preparation.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. Journal of Personality and Social Psychology, 28, 28-38.

Newtson, D., Engquist, G., & Bois, J. (1976). The reliability of a measure of behaviour perception. JSAS Catalog of Selected Documents in Psychology, 6(5), 1173.

Newtson, D., Engquist, G., & Bois, J. (1977). The Objective Basis of Behavior Units. Journal of Personality and Social Psychology, 35(12), 847-862.

Newtson, D., & Enqguist, G. (1976). The perceptual organization of ongoing behaviour. Journal of Experimental Social Psychology, 12, 436-450.

Nichols, B., & Lederman, S. J. (1985). Flicker and Motion in Film. T. De Lauretis, & S. Heath (Eds.), The Cinematic Apparatus (pp. 96-105). Palgrave Macmillan.

Noles, N. S., Scholl, B. J., & Mitroff, S. R. (2005). The persistence of object file representations. Perception & Psychophysics, 67(2), 324-334.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: the world as an outside memory. Canadian Journal of Psychology, 46, 461-488.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture Changes During Blinks: Looking Without Seeing And Seeing Without Looking. Visual Cognition, 7(1-3), 191-211.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences, 24, 939-1011.

Olson, I. R., Gatenby, J. C., Leung, H.-C., Skudlarski, P., & Gore, J. C. (2004). Neuronal representation of occluded objects in the human brain. Neuropsychologica, 42, 95-104.

Ornstein, R. E. (1969). On the experience of time. Hammondsworth,UK: Penguin.

Palmer, S. E. (1999). Vision Science: Photons to Phenomenology. Cambridge, Mass: MIT Press.

Pashler, H. E. (1998). The psychology of attention. Cambridge, MA: MIT Press.

Pavlov, I. P. (1927). Conditioned Reflexes: An investigation of the physiological activity of the cerebral cortex. London: Wexford University Press.

Penton-Voak, I. S., Edwards, H., Percival, A., & Wearden, J. H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. Journal of Experimental Psychology: Animal Behavior Processes, 22, 307-320.

Pepperman, R. D. (2004). The Eye is Quicker: film editing; making a good film better. Los Angeles, USA: Michael Wiese Productions.

Peterson, L. B., & Peterson, M. J. (1959). Short-term retention of individual items. Journal of Experimental Psychology, 58 , 193-198.

Piaget, J. (1929). The child's conception of the world. New York: Harcourt Brace.

Piaget, J. (1954). The Construction of reality in the Child. New York: Basic Books.

Posner, M., Snyder, C., & Davidson, B. (1980). Attention and the detection of signals. Journal of Experimental Psychology: General, 109, 160-174.

Posner, M. I. (1980). Orienting of attention. Quarterly Journal of Experimental Psychology, 32, 3-25.

Posner, M. I. (1994). Attention: the mechanisms of consciousness. Proceedings of the National Academy of Sciences., 91, 7398-7403.

Poynter, W. D. (1989). Judging the duration of time intervals: A process of remembering segments of experience. I. Levin, & D. Zakay (editors), Time and human cognition: A life-span perspective (pp. 305-31). Amsterdam: North Holland.

Prunes, M., Raine, M., & Litch, M. (2002) Yale Film Analysis Web Site [Web Page]. URL http://classes.yale.edu/film-analysis/index.htm [1929, December 5].

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. Cognition, 32, 65-97.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. Cognition, 80, 127-158.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parrallel tracking mechanism. Spatial Vision, 3, 179-197.

Pöppel, E. (1970). Excitability cycles in central intermittency. Psychological Forschung, 34, 1-9.

Pöppel, E. (1997). A hierarchical model of temporal perception. Trends in Cognitive Sciences, 1(2), 56-61.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? Journal of Experimental Psychology. Human Perception and Performance, 18 (3), 849-60.

Reeves, B., Thorson, E., Rothschild, M., McDonald, D., Hirsch, J., & Goldstein, R. (1985). Attention to television: Intra stimulus effects of movement and scene changes on alpha variation over time. International Journal of Neuroscience, 25, 241-255.

Reisz, K., & Millar, G. (1953). Technique of Film Editing. London, UK: Focal Press.

Rensink, R. A. (2000). The dynamic representation of scenes. Visual Cognition, 7, 17-42.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: the need for attention to perceive changes in scenes. Psychological Science, 8, 368-373.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (2000). On the failure to detect changes to scenes across brief disruptions. Visual Cognition, 7, 127-145.

Rock, I. (1985). PerceptionScientific American Library.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 8(3), 382-439.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Caro, D. M. (2002). A Longitudinal Study of Visual Expectation and reaction Time in the First Year of Life. Child Development, 73(1), 47-61.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. Nature, 386, 598-601.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. Trends in Neuroscience, 24(2), 113-120.

Ruz, M., & Lupiánez, J. (2002). A review of attentional capture: On its automaticity and sensitivity to endogenous control. Psicologica, 23, 283-309.

Salt, B. (1983). Film Style and Technology: History and Analysis . London: Starword.

Sandys, J. (2005). Movie Mistakes: Take 4. UK : Virgin Books.

Saylor, M., & Baldwin, D. (2005). Action Parsing in Infancy . D. T. Levin (Ed.), Thinking and Seeing: Visual Metacognition in Adults and Children . Cambridge, MA: MIT Press.

Scholl, B. J., & Leslie, A. M. (1999). Explaining the infant's object concept: Beyond the Perception/Cognition dichotomy. E. Lepore, & Z. W. Pylyshyn (Eds), What is Cognitive Science? Oxford: Blackwell.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: clues to visual objecthood. Cognitive Psychology, 38, 259-290.

Scholl, B. J., Pylyshyn, Z. W., & Franconeri, S. L. (1999). When are featural and spatiotemporal properties encoded as a result of attentional allocation? Investigative Opthamology and Vision Science, 39, S797.

Schröder, J. (1990). Die psychologische Realität von Prinzipien des Continuity Cinema. G. Schumm, & H. G. Wulff (Eds.), Film und Psychologie I. Kognition-Rezeption-Perzeption (pp. 109-142). Münster: Maks Publikationen.

Schwan, S., & Garsoffky, B. (2004). The Cognitive Representation of Filmic Event Summaries. Applied Cognitive Psychology, 18, 37-55.

Schwan, S., Garsoffky, B., & Hesse, F. W. (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? Memory & Cognition, 28(2), 214-223.

Seddon, M. D. (2003). Investigating the effect on attention of action continuity. Notes: Unpublished MSc. Thesis.

Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. Trends in Cognitive Sciences, 1(8), 291-296.

Shapiro, K. L., Driver, J., Ward, R., & Sorensen, R. (1996). Priming from the attentional blink: A failure to extract viual tokens but not visual types. Psychological Science, 8, 95-100.

Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1995). Spatial attention and eye movements. Exp Brain Res, 105, 261-275.

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. Psychological Review, 91(4), 441-447.

Simons, D. J., & Levin, D. T. (1997). Change Blindness. Trends in Cognitive Science, 1, 261-267.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during real-world interaction. Psychonomic Bulletin and Review, 5, 644-649.

Simons, D. J. (2000). Attentional capture and inattentional blindness. Trends in Cognitive Sciences, 4(4), 147-155.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. Perception, 2, 1059–1074.

Simons, D. J., & Rensink, R. A. (2005). Change Blindness: past, present, and future. Trends in Cognitive Science, 9(1).

Singer, J. (1980). The power and limitations of television: A cognitive-affective analysis. P. Tannenbaum (Ed.), The entertainment functions of television. (pp. 36-42). Hillsdale, NJ: Erlbaum.

Smith, T. J., Whitwell, M., & Lee, J. (in press). Eye Movements and Pupil Dilation During Event Perception. Proceedings of the Eye Tracking Research and Applications conference 2006 . Notes: Poster Abstract

Speer, N. K., Swallow K. M., & Zacks, J. M. (2003). On the role of human areas MT+ and FEF in event perception. Cognitive, Affective and Behavioral Neuroscience, 3, 335-345.

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. British Journal of Developmental Psychology, 13, 113-142.

Sperling, G. (1960). The information available in brief visual presentations. Psychological Monographs, 74 (11).

War of the Worlds. (2005). S. Spielberg (Producer), & S. Spielberg (Director).

Stelmach, L., Tam, W. J., & Hearty, P. (1991). Static and dynamic spatial resolution in image coding: An investigation of eye movements. *Human Vision, Visual Processing and Digital Display II*, 1453, 147-152.

Sweeney, L. (1973). Interview with John Huston. Christian Science Monitor.

Szelag, E., von Steinbüchel, N., Reiser, M., de Langen, E. G., & Pöppel, E. (1996). Temporal constraints in processing nonverbal rhythmic patterns. Acta Neurobiol Exp, 56, 215–225.

Tan, E. S.-H. (1995). Constraint and Convention in Psychological Film Aesthetics: On the Psychological Basis of Cinematic Conventions. J. Müller (Ed.), Towards a pragmatics of the audiovisual. (pp. 67-86). Münster: Nodus.

Theeuwes, J. (1994). Stimulus-driven capture and attentional set: Selective search for colour and visual abrupt onsets. Journal of Experimental Psychology: Human Perception and Performance , 20, 799-806.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. Psychological Science, 9(379-385).

Tosi, V., Mecacci, L., & Pasquali, E. (1997). Scanning eye movements made when viewing film: preliminary observations. International Journal of Neuroscience, 92, 47-52.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12(1), 97-136.

Treisman, M. (1963). Temporal discrimination and indifference interval: Implications for a model of the "internal clock". Psychologial Monographs, 77(13, whole No. 576), 1-13.

Treuting, J. (2004). Eye Tracking and the Cinema: A study of Film Theory and Visual Perception.

Ullman, S. (1979). The interpretation of visual motion. Cambridge, MA: MIT Press.

Van der Heijden, A. H. C. (1992). Selective Attention in Vision. London: Routledge.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information Processing in the Primate Visual System: An Integrated Systems Perspective. Science, 255(5043), 419-423.

Van Essen, D. C., & Maunsell, J. H. R. (1983). Hierarchical organisation and functional streams in the visual cortex. Trends in Neuroscience, 6, 370-375.

Volkmann, F. C., Riggs, L. A., & Moore, R. K. (1980). Eyeblinks and visual suppression. Science, 207, 900-902.

Vollrath, M., Kazenwadel, J., & Krüger, H. P. (1992). A universal constant in temporal segmentation of human speech. Naturwissenschaften, 79, 479-480.

von Steinbüchel, N., Wittmann, M., & Pöppel, E. (1996). Timing in perceptual and motor tasks after disturbances of the brain. (Pastor, M. A. Artieda, J.), Time,internal clocks and movement (pp. p. 281-304.). Amsterdam: Elsevier.

Walker, R., Walker, D. G., Husain, M., & Kennard, C. (2000). Control of voluntary and reflexive saccades. Experimental Brain Research, 130(4), 540-544.

Whitwell, M. (2005). Using Eye-Tracking Techniques to Investigate the Segmentation of Moving Images. Notes: Unpublished MSc. Thesis

Wikipedia. (Continuity (fiction) [Web Page]. URL http://www.wikipedia.org [2005, October 30].

Williams, C. C., Henderson, J. H., & Zacks, R. T. (in press). Incidental Visual Memory for Targets and Distractors in Visual Search. Perception and Psychophysics.

Williams, P., & Simons, D. J. (2000). Detecting Changes in Novel, Complex, Three-dimensional Objects. Visual Cognition, 7(1/2/3), 297-322.

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive Object Files: Shapeless Bundles of Basic Features. Vision Research, 37(1), 25-43.

Woodworth, W. S. (1906). Vision and localisation during eye movements. Psychological Bulletin, 3, 68-70.

Worth, S., & Adair, J. (1972). Through Navajo Eyes: An Exploration in Film Communication and Anthropology. Bloomington: Indiana University Press.

Wynn, K. (1996). Infants' individuation and enumeration of actions. Psychological Science, 7, 164-169.

Xu, F. (1999). Object individuation and object identity in infancy: The role of spatiotemporal information, object property information and language. Acta Psychologica, 102, 113-136.

Xu, F., & Carey, S. (1996). Infants' metaphysics: the case of numerical identity. Cognitive Psychology, 30, 111-153.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. Journal of Experimental Psychology: Human Perception and Performance , 10, 601-621.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. Journal of Experimental Psychology: Human Perception and Performance, 16, 121-134.

Yarbus, A. (1967). Eye Movements and Vision. New York: Plenum Press.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001a). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. Nature, 414.

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001b). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. Nature, 414.

Yarrow, K., Haggard, P., & Rothwell, J. C. E. (in press). Action, arousal, and subjective time. Consciousness and Cognition.

Yarrow, K., Johnson, H., Haggard, P., & Rothwell, J. C. E. (2004). Consistent Chronostasis effects across saccade categories imply a subcortical efferent trigger. Journal of Cognitive Neuroscience, 16(5), 839-847.

Yarrow, K., & Rothwell, J. C. E. (2003). Manual chronostasis: Tactile perception precedes physical contact. Current Biology, 13, 1334-1339.

Yarrow, K., Whiteley, L., Rothwell, J. C., & Haggard, P. (2005). Spatial consequences of bridging the saccadic gap. Vision Research, 46, 545-555.

Yi, G.-J., Kim, M.-S., & Chun, M. M. (2003). Inhibition of return to occluded objects. Perception & Psychophysics, 65(8), 1222-1230.

Zacks, J., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. Nature Neuroscience, 4(6), 651-655.

Zacks, J. M. (2004). Using movement and intentions to understand simple events. Cognitive Science, 28, 979-1008.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and cognition. Psychological Bulletin, 127, 3-21.

Zakay, D. (1989). Subjective time and attentional resource allocation: An integrated model of time perception. I. Levin, & D. Zakay (editors), Time and human cognition: A life-span perspective (pp. 365-397). Amsterdam: North Holland.

Zakay, D. (1992). The role of attention in children's time perception. Journal of Experimental Child Psychology, 54, 355-371.

Zakay, D. (1993). Time estimation methods. Do they influence prospective duration estimates? Perception, 22, 91-101.

Zakay, D., & Block, R. (1996). Role of Attention in Time Estimation Processes. Time, Internal Clocks, and Movement . Elsevier Science.

Zakay, D., Nitzan, D., & Glicksohn, J. (1983a). The influence of task difficulty and external tempo on subjective time estimation. Perceptual Psychophysics, 34(5), 451-56.

Zakay, D., Nitzan, D., & Glicksohn, J. (1983b). The influence of task difficulty and external tempo on subjective time estimation. Perceptual Psychophysics, 34(5), 451-56.

Zakay, D., Tsal, Y., Moses, M., & Shahar, I. (1994). The role of segmentation in prospective and retrospective time estimation processes. Memory and Cognition , 22(3), 344-51.