Accepted for publication in *Media Psychology* on 02/02/16

Match-Action:

The role of motion and audio in creating global change blindness in film.

Tim J. Smith* & Janet Yvonne Martin-Portugues Santacreu

Birkbeck, University of London

*Corresponding Authors contact details

Department of Psychological Sciences

Birkbeck, University of London

Malet Street

LONDON

UK

WC1E 7HX

Tel: +44 207 631 6359

e-mail: tj.smith@bbk.ac.uk

ABSTRACT

An everyday example of change blindness is our difficulty to detect cuts in an edited moving-image. *Edit Blindness* (Smith & Henderson, 2008) is created by adhering to the continuity editing conventions of Hollywood, e.g. coinciding a cut with a sudden onset of motion (*Match-Action*). In this study we isolated the roles motion and audio play in limiting awareness of match-action cuts by removing motion before and/or after cuts in existing Hollywood film clips and presenting the clips with or without the original soundtrack whilst participants tried to detect cuts. Removing post-cut motion significantly decreased cut detection time and the probability of missing the cut. By comparison, removing pre-cut motion had no effect suggesting, contrary to the editing literature, that the onset of motion before a cut may not be as critical for creating edit blindness as the motion after a cut. Analysis of eye movements indicated that viewers reoriented less to new content across intact match-action cuts than shots with motion removed. Audio played a surprisingly large part in creating edit blindness with edit blindness mostly disappearing without audio. These results extend film editor intuitions and are discussed in the context of the *Attentional Theory of Cinematic Continuity* (Smith, 2012a). [198 words]

INTRODUCTION

In real-world audiovisual scenes, whole-field visual changes only occur when our viewpoint of a scene changes gradually via ego-motion or rapidly through the volitional movement of the eye. For both types of change we are aware of the change due to our being complicit in its creation. However, modern sensory life is no longer constrained by physical laws as we spend a significant portion of our waking lives engaged in experiences mediated by computer, smart phone, TV or cinema screens (one estimate is as much as a fifth of an adult's waking day: Men=19.92%, Women=17.28%; Bureau of Labor Statistics, 2014). The audiovisual information presented by screens can change instantaneously without any contingency on the visual scene within which it is embedded. Such changes should therefore be highly salient, attract attention and require cognitive effort to comprehend how the new information relates to the old. However, filmmakers believe they have the ability to limit the negative impact of cuts and effectively make cuts invisible by adhering to a suite of filming and compositional techniques known as the *Continuity Editing Rules* (or Hollywood Style; Bordwell & Thompson, 2001).

A typical ninety minute Hollywood film contains between one thousand and two thousand edits, a change in shot every 2.7 to 5.4 seconds (Bordwell & Thompson, 2001), yet film editors assume that for the majority of these edits the "*spectator's illusion of seeing a continuous piece of action is not interrupted*" (Reisz & Millar, 1953; pg 216). The Continuity Editing (CE) Rules make strong predictions about how viewers will perceive an edited sequence (Bordwell & Thompson, 2001), however, up until recently they received very little empirical psychological investigation (see Lang, 2000 for review). Cognitive scientists have recently begun to turn their attention to such questions under the guise of *neurocinematics* (Hasson, Landesman, Knappmeyer, Valines, Rubin, & Heeger, 2008), *film cognition* (Smith, Levin, & Cutting, 2012), *psychocinematics* (Shimamura, 2013) or *cognitive media theory*

(Nannicelli & Taberham, 2014). Given that these techniques permeate all forms of film and TV programming and filmmakers religiously stick to them it is surprising that our understanding of their cognitive and perceptual foundations are not better understood. For example, one of the most powerful CE techniques is known as *Match-Action* editing (or *Match-On-Action*). Match-Action editing attempts to create the impression of a continuous action by coinciding a cut with the onset of the action and showing the continuation of that action in the second shot (Anderson, 1996; Reisz & Millar, 1953). Match-Action editing was one of the earliest techniques used to edit a scene (films as early as 1901 use this technique; Salt, 2009) and matching content and action across a cut is believed to be a fundamental principle of the continuity style (Smith, 2012a). For example, in a clip from *Six Days Seven Nights* (Figure 1), Anne Heche's character unexpectedly drops out of a plane in shot 1 and is shown crashing face first into the sand in shot 2. The continuation of her motion is believed by editors to make the cut invisible (Dmytryk, 1986; Murch, 2001; Pepperman, 2004). This belief has been confirmed by a study in which existing match-action cuts in feature films were missed significantly more often during a cut detection task (32.4%) than cuts between unrelated scenes (9.4%) (Smith & Henderson, 2008). Failure to detect a cut constitutes global *change blindness* (Rensink, ORegan, & Clark, 1997) as the entire visual scene changes (assuming the film screen fills most of the viewer's visual field). Every cut involves an instantaneous and total replacement of one image with a different image, and should therefore be highly salient but something about the composition of Match-Action cuts limits awareness of the large and unnatural sensory event.

What is it about match-action cuts that make their detection so difficult? By definition, all cuts involve a change in the visual content of the cinematic image either through a change in subject, perspective, shot composition, camera location, elision of time or a combination of these factors. These sudden changes create a range of cognitive and

physiological impacts on the viewer including slowed secondary task reaction time (Geiger & Reeves, 1993), focussing of attention (Reeves et al., 1985), increased eye movements (d'Ydewalle, Desmet, & Van Rensbergen, 1998; Hochberg & Brooks, 1978; Smith & Henderson, 2008), improved recognition memory (Frith & Robson, 1975; Lang, 1991) and faster recognition for information presented after a cut (Carroll & Bever, 1976) as well as heart rate deceleration (Lang, 1990) and an increase in self-reported arousal (Lang, Zhou, Schwartz, Bolls, & Potter, 2000). As the degree of change increases so does viewer awareness of the cut (Smith & Henderson, 2008) and the cognitive load experienced by the viewer (Lang, Kurita, Gao, & Rubenking, 2013). The most common type of cuts, *Shot-Reverse shots* between two characters within a scene (Bordwell, Staiger, & Thompson, 1985) preserve temporal continuity of the scene whilst changing the subject and perspective, resulting in a 25.1% miss rate (Smith & Henderson, 2008). Such cuts typically also retain audio continuity across the cut with dialogue, environmental sounds and musical score creating a "perceptual scaffold" between the two shots (Smith, 2012a).

In addition to the features preserved across a within scene cut, match-action cuts also preserve the subject of the shot (e.g. a hand or head) and the action depicted (e.g. the hand reaches out and grabs an object across the cut or a head turns). A large movement such as the sudden whip-pan of the camera will obviously make a cut hard to detect as due to the low frame rate of most movie cameras (e.g. 24fps) any high velocity movement blurs the visual content of the image making the timing of the change to a new shot hard to detect. However, film editors (Dmytryk, 1986; Murch, 2001; Pepperman, 2004) also believe that subtle movements within a stable frame such as a shift in actor gaze, changes in facial expressions and blinks can be used to hide cuts. Such small changes will create high motion contrast relative to the stable background. Motion contrast is known to be one of the most reliable visual features to capture attention in classic cuing or search paradigms (Abrams & Christ,

2003; Wolfe & Horowitz, 2004) and to predict gaze location during dynamic scene free-viewing (Carmi & Itti, 2006; Mital, Smith, Hill, & Henderson, 2011; Smith & Mital, 2013). The onsets of motion may elicit an overt attentional shift (i.e. saccadic eye movement) towards the source of the motion, creating a period of insensitivity to visual information (due to saccadic suppression) or covert attentional shift that may make detection of the subsequent cut difficult (due to an attentional blink; Levin & Saylor, 2008). Such brief distractions may operate in a similar manner to the flickers (Rensink et al., 1997) or mudsplashes (O'Regan, Rensink, & Clark, 1999) used in classic change blindness paradigms. However, Smith & Henderson (2008) found no evidence that missed cuts coincided with saccadic eye movements any more than detected cuts. There was also no evidence that missed cuts coincided with viewer eye blinks, another mechanism hypothesised by editors for hiding cuts (Murch, 2001).

Previous studies investigating edit blindness have confirmed the phenomena's existence but they have not provided explanation of why it occurs. Which factors are critical for the creation of edit blindness across a Match-Action cut? The most obvious factor is action. *"…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). Previous eye tracking studies have revealed a peak in saccadic activity immediately following all cuts (Hochberg & Brooks, 1978; Mital et al., 2011; Smith, 2013) which increases with continuity errors (d'Ydewalle et al., 1998; Germeys & d'Ydewalle, 2007). Smith & Henderson (2008) showed that the size of this saccadic peak varied with the type of cut: match-action cuts were followed by a low peak whereas cuts

within a scene without such motion cues resulted in a larger peak. This increase in saccadic activity has been interpreted as indicative of viewer awareness of the onset of a new scene content and reorientation to it (Hochberg & Brooks, 1978). Although caution should be taken when attributing awareness to overt attention as fixating an object does not necessitate awareness of the object and involuntary saccades, like those occurring across a cut, may not result in awareness for the target of the saccade (Smith, Lamont, & Henderson, 2012, 2013).

Previous studies (d'Ydewalle et al., 1998; Germeys & d'Ydewalle, 2007; Smith & Henderson, 2008) have compared saccadic activity across different shots and, as such, the influence of shot content cannot be separated from the match-action techniques itself. The present experiment will overcome these content issues by comparing saccadic activity across a match-action cut to the same shot pair with or without the motion before or after the cut. By comparing four versions of the same cut, *Intact* (both pre and post-cut motion), *-Pre* (pre-cut motion removed), *-Post* (post-cut motion removed) and *–Both* (without any motion) the independent contributions of the two periods of motion on the edit blindness effect will be able to be identified. It is hypothesised that if motion prior to the cut is capturing attention and occupying it during the cut, awareness of cuts and saccade frequencies will be lower immediately following *Intact* cuts and *–Post* cuts (i.e. cuts with only pre-cut motion) compared to cuts with motion before the cut removed (i.e. *–Pre*) or with no motion (i.e. *–Both*). This is the strong hypothesis made by the *Attentional Theory of Cinematic Continuity* (AToCC; Smith, 2012a). AToCC is a cognitive theory of how continuity is perceived across an edited film sequence that is derived from contemporary theories of real-world scene perception and active vision. Within AToCC, the cognitive explanation of Match-Action states "*Sudden onsets of movement within a shot draw attention to the screen location of the movement and its future trajectory…. Expectations about the visual scene and object features are abandoned as we focus on the spatiotemporal details of the action. If a match-action cut*

is timed to coincide with the onset of the action and presents a new viewpoint of the same action at the same screen location as pre-cut the viewers' expectations will be satisfied and a priori continuity will be perceived." (Smith, 2012a; pg. 15). This posits pre-cut motion as a critical component of a match-action cut.

Indirect evidence in support for AToCC comes from studies investigating how first-time adult film viewers perceive visual content across cuts. Schwan and Ildirar (2010) presented short film sequences to participants from a remote community in Turkey who had never previously been exposed to TV or film. When asked to describe what they had seen, first-time viewers were only able to accurately describe edited sequences in which a familiar line of action bridged the cut. In a follow-up study (Ildirar & Schwan, 2015), the authors compared cuts combining shots with different types of relations, such as pictorial (e.g. the same person depicted from two viewpoints), causal (e.g. passing a bucket across a Match-Action cut) or conceptual (e.g. a person looking down at their shoes depicted in a close-up or intending to throw a stone at a troublesome dog). First-time viewers were able to perceive all types of relationships but struggled when the activity depicted was not clearly established before the cut or when a character's intention had to be inferred. Given their absence of prior knowledge of editing conventions, these results suggest that first-time viewers were using familiarity with the everyday actions depicted in the first shot to guide comprehension across the cut. Whether the novice viewers experienced more or less edit blindness than experienced viewers is not known.

An alternate mechanism for edit blindness may be *motion silencing* (Suchow & Alvarez, 2011). Motion silencing is the inability to detect changes to object hue, luminance, size or shape when they or the scene background move relative to a viewer's stable fixation (Suchow & Alvarez, 2011). If the viewer pursues the moving object with their eyes, silencing disappears but if they hold fixation and attempt to pursue with covert attention silencing

reappears. The phenomenon is believed to demonstrate the difficulty in updating multiple object representations simultaneously when they all move on the retina. Whilst smooth pursuit of a moving target allows changes to the features of that object to be perceived, similar changes in the periphery, away from the pursued target may be missed due to blurring of these peripheral features on the retina (Tse, 2009). Both phenomena could account for edit blindness if the motion onset shifted critical features of the shot relative to the viewer's stable gaze (motion silencing) or the viewer pursued a point of low velocity motion, decreasing their peripheral sensitivity to the non-moving elements of the shot. Such effects would be visible in gaze mobility across a match-action cut. Low velocity eye motion (e.g. smooth pursuit or drift) due to tracking moving objects before or after the cut will create blurring of the image on the retina making cut detection difficult. If gaze mobility is a causal factor in the creation of edit blindness, removing the pre-cut motion should stop gaze mobility across the cut and eradicate edit blindness. By comparison, removing only post-cut motion may actually increase awareness of the cut as the sudden cessation of the visual motion tracked prior to the cut stops gaze mobility.

Attentional capture, peripheral motion silencing during fixation or smooth pursuit may all provide sensory or cognitive mechanisms for creating edit blindness (independently or in combination) but all would require a motion onset prior to the Match-Action cut. Most specifications of match-action editing require motion onset prior to the cut (even if the motion is brief; only a few frames; Dmytryk, 1986). However, they also state that this onset must be matched by similar motion in the next shot. Intriguingly, the subsequent shot does not need to have any semantic relationship to the previous shot. A cut can be made to a completely unrelated scene and continuity will still be perceived if the motion before and after the cut match (this technique is often used in TV advertisements and music videos;

Smith, 2012a)[1]. Film editors would not consider a cut without continuation of motion after the cut to be a Match-action and would predict awareness of the cut to be greater than if the post-cut motion was present (Anderson, 1996). If this is the case the post-cut motion may be as important for the creation of edit blindness as the pre-cut motion. The sudden onset of motion following a cut may create backward masking (Breitmeyer & Öğmen, 2000) for the visual information immediately preceding the cut, thereby making perceptual comparison between the new and old visual information difficult and limiting cut awareness. However, whilst Enns and Di Lollo (2009) have proposed that visual backward masking may be relevant to change blindness, masking typically requires a very specific timecourse of visual stimulation which is unlikely to be matched by film cuts. Also, in order for the post-cut motion to mask the visual transients created by the cut they would need to be of at least the same magnitude as the cut transients (and, as a result, indistinguishable from the cut). Otherwise, it would be the cut that draws attention and instead masks the pre-cut content. If the post-cut motion is responsible for edit blindness due to backward masking, removing the post-cut motion should eradicate edit blindness. Although, it should be noted that the importance of post-cut motion in creating edit blindness is typically downplayed in the film literature (e.g. Dmytryk, 1986) and AToCC (Smith, 2012a).

As well as manipulating pre/post-cut motion, the role audio plays in the creation of edit blindness will also be investigated. Audio is the only property of a film that can remain continuous across a cut and as such it may provide a perceptual "scaffold" upon which audiovisual continuity can be perceived. Recent evidence from a secondary task reaction time study suggested attention is attenuated immediately following cuts on dialogue but only when the film is presented with the original audio (Shimamura, 2013). Also, several recent studies

[1] An extreme example of match-action editing that attempts to minimises awareness of cuts joining shots from multiple movies is available here: http://vimeo.com/35167016

have suggested that viewing activity changes when watching videos without sound when the sound references an object within the image, accents key events or is produced by an object, e.g. speech (Coutrot, Guyader, Ionescu, & Caplier, 2012; Hirvenkari et al., 2013; Mera, & Stumpf, 2014; Song, Pellerin, & Granjon, 2013; Võ, Smith, Mital, & Henderson, 2012) whereas gaze in scenes with only off-screen sounds (Song et al., 2013), audio narration (Ross & Kowler, 2013) or scenes accompanied by a musical score (Smith, 2014) differ less when the audio is removed. Audio may also help first-time adult film viewers perceive the connection between shots which otherwise are perceived as two unrelated images (Ildirar, Levin, Schwan, & Smith, 2014). Audio may operate in a similar way in match-action cuts although no clear mention of this effect can be found in the editing literature which typically focuses on the visual aspects of the technique (Anderson, 1996; Bordwell & Thompson, 2001; Dmytryk, 1986; Reisz & Millar, 1953). If the audio continuity across a match-action cut is important for obscuring the cut and creating edit blindness removal of the audio should increase cut detection. How audio will interact with pre and post-cut motion is unknown.

The present experiment investigated the role pre/post-cut motion and accompanying soundtrack plays in creating edit blindness across match-action cuts. The presence of pre- and post-cut motion (Cut Type) and soundtrack (Audio) in short feature film clips were manipulated within participants whilst they performed a cut detection task in which they were instructed to press the spacebar every time they saw a cut. Their eye movements were also recorded during the task.

METHODOLOGY

Participants

Thirty-two participants completed the experiment: 21 females and 11 males, 19 to 35 years of age (M = 24.61, SD = 4.41). Participants were excluded if their critical cut reaction time or miss rate was greater than three standard deviations above the sample mean (N=1) or

if the eyetracker identified them as spending greater than 50% of the experiment in an eyeblink (N=4; this is a sign of poor tracking or fatigue). Twenty-seven participants were entered into the final analysis reported below. Participants were asked to rate their "Knowledge of film theory" on a five point scale ("None, Passing, Amateur, Undergraduate, Postgraduate/professional level"). All participants rated their familiarity as "None" or "Passing" and can therefore be considered non-experts in the formal cinematic features manipulated in this study.

Participants were either paid £8 for participation or received course credit. The experiment received ethical approval by the College Ethics Committee.

Design

A within subjects design with two independent variables (Audio and Cut Type) was used. There were four types of match-action cut according to the level of manipulation (see Figure 1): 1) *Intact*: there was no manipulation to the movement in the clip; 2) *–Pre*: the movement before the cut was removed but the movement after the cut was left intact; 3) *–Post*: the movement after the cut was removed but the movement before the cut was left intact; 4) *–Bot*h: both pre-cut and post-cut motion were removed. The audio variable had two levels: 1) the original audio was present throughout the clip (*Audio* condition); 2) all audio was removed from the clip (*Silent* condition). See Video 1 for a demonstration of match-action editing and the manipulations used in this study[2].

The behavioural task was to detect every cut in the movie clips by pressing the spacebar as soon as they saw a cut. Although participants detected all cuts we were only interested in their perception of match-action cuts. As such, the behavioural dependent

[2] Demonstrations of the cut type and audio manipulations made to one clip used in this study, *Six Days Seven Nights* (Ivan Reitman, 1998). Miss probabilities are included in the video as well as a slow-motion break-down of the motion manipulated across cut conditions. http://youtu.be/RJfOPHicLKY

variables derived from this task were miss rate and reaction times for match-action cuts. Cut miss rate was the probability of missing a match-action cut. Reaction time was the time elapsed from the critical cut to when the participant pressed the spacebar.

The eye movement dependent variables were saccade frequency in 100ms time bins relative to the critical cut and gaze mobility. Gaze mobility was calculated as the average velocity of viewer gaze during the frame across the critical cut. Gaze was parsed of blinks and saccades and then the Euclidean displacement of raw gaze during the fixation crossing the critical cut was calculated. Given that smooth pursuit movements were not identified in this dataset (see Smith & Mital, 2013 for discussion of why this is acceptable) all low velocity movements would be classified as fixations unless the velocity exceeded the saccade threshold (>30 °/s and >8000 °/s^2). Any tracking of image motion by gaze would therefore be manifest in a small but significant increase in gaze velocity during the critical fixation.

<<Insert Figure 1 here>>

Stimuli

Eighty film clips (~20s in duration) were extracted from feature films. Each clip contained one match-action cut. The selection of the excerpts conformed to the following criteria: a) each excerpt should contain only one match-action cut but it could also contain other types of cut (mean number of cuts per clip was 4.52, *SD*=2.22); b) the movement in the match-action cut should be short (<2s) and involve a person; for example, a head or body turn, an arm lift, or a short walk; c) the content of the excerpt should have a coherent narrative; d) the target duration of an excerpt was 20 seconds; however, the exact duration of the excerpt varied to preserve the narrative; e) there should be no speech across the cut; f) the excerpt should not contain offensive material or violence. All films were classified 15 or under by the British Board of Film Classification (equivalent to either a US PG-13 or R rating, depending on the film). However, a small number of excerpts used in the experiment

slightly contravened these criteria, e.g. a non-critical sentence of dialogue across the cut.

The 80 clips chosen as stimuli for this study were taken from a range of feature films (see Appendix Table 1 for full details) and involved a variety of actions: 5% of clips cut on an upper torso movement (e.g. *Victor Victoria* (1982) cut as a character leans down to kiss Julie Andrews' hand), 17.5% depicted a head movement (e.g. *Groundhog Day* (1993) cut as Bill Murray's character looked down at himself), 28.75% of clips depicted an action involving the hands (e.g. *The American President* (1995) cut on glasses being chinked together during a toast), and 48.75% cut on a whole body movement (e.g. *Contact* (1997) cut as Jodie Foster's character sits down). Across the cut 38.75% of clips involved a cut in to a closer shot (e.g. Long Shot to Close-Up), 30% involved a cut out to a longer shot and 31.25% involved a cut between two similar shot sizes (typically involving a change in camera angle on the subject). A range of changes in camera angle were also depicted (modal angle change was 45%; see Appendix Table 1 for further details).

Each of the original 80 excerpts was manipulated to generate eight versions that corresponded to the four cut conditions crossed with the two audio conditions. The critical motion that was manipulated was identified as the action that crossed the cut. Other movement occurring before or after this cross-cut action was left intact. For example, in a clip from *Six Days Seven Nights* (Figure 1), Anne Heche's drop out of the plane was removed in all –Pre conditions (Figure 1; 2nd and 4th row) and the continuation of this movement was removed in all –Post conditions (Figure 1; 3rd and 4th row). All of Anne Heche's actions up to this fall or after were kept intact. Any accompanying audio was also removed (see Video 1).

In order to ensure our manipulations did not introduce weird or highly salient discontinuities (such as a noticeable disappearance of an object or audiovisual artefact) an on-line control study was performed using the Birkbeck Psychological Sciences web experiment platform. Eight participants were shown each of the eighty clips in one of the eight conditions

at home via their web browser. Participants were instructed to wear headphones, minimise distractions in their surroundings and maximise the browser window so that it filled their entire screen. Their task was to detect if each clip had been "*re-edited to remove some frames from the original clip*". The mean probability that an edited clip (-Pre, -Post, -Both) was detected was 0.30 (*SD*=.204) and the mean probability that an unedited clip (Intact) was falsely reported as containing a re-edit was 0.19 (*SD*=.136). These detection rates did not differ significantly ($t(7)=-1.54$, p=.168, n.s.). This confirms that our manipulations did not increase the salience of the edited clips beyond that of the original unedited clips. Any subsequent behavioural results can therefore be attributed to the removal of the critical motion or audio and not the introduction of artefacts.

Procedure and Apparatus

Film clips were presented using Experiment Builder (SR Research) on a 21 inch CRT monitor (720x576 pixel resolution DVD quality on an 800x600 screen; viewing angle = 36.44 degrees) at a distance of 60cm with participant head stabilised on a chinrest. Stereo audio was presented on Sennheiser stereo headphones. Eye movements were recorded monocularly at 1000Hz using an Eyelink 1000 (SR Research). Dominant eye was tracked. Reponses were made via a Microsoft Sidewinder gaming controller.

The experimental procedure began with onscreen instructions and then eye tracker calibration. Nine point calibration was used and any points that exhibited an offset of >1 degree were recalibrated. Participants were then shown a short introductory video explaining what a generic film cut was and how they should respond to cuts. Participants completed a practice session, consisting of four videos, each with a different number of cuts. Feedback was given at the end of the practice session in terms of number of cuts per video. All clips were presented full-screen (800 x 600 resolution) but aspect ratio varied between clips with the top and bottom of the screen filled with black bars.

In the main experiment, each participant was presented with a total of 80 different videos. There were 5 trial blocks, each containing 16 videos. The order of videos was randomised. Each trial began with a 3 second central fixation cross followed by the test video. All button presses were recorded during playback. No feedback on accuracy was provided to participants during the main experiment. The whole experiment took around 45 minutes to complete including eyetracker setup and calibration.

RESULTS

Behavioural results (Cut detection RTs and Miss Rates) will be explored before eye movement measures (saccade frequency and gaze mobility) are analysed.

Reaction Time

The time taken to identify the critical cut across all four cut types (Intact, -Pre, -Post, -Both) and two audio conditions (Silent vs. Audio) were analysed with a within-subjects ANOVA. Any reaction times over 2000ms but before the next cut occurred were reclassified as misses. Differences in cut characteristics (Appendix Table 1) did not have any significant effect on reaction times (Size Change In vs. Out, $t(44)=-.596$, $p=.554$; Angle change, $F(4,66)=.851$, $p=.498$) so all subsequent analysis is collapsed across clip.

The ANOVA revealed a main effect of Audio, $F(1,26)=7.098$, $p=.013$ $p\eta^2=.214$, a main effect of Cut type, $F(3,78)=6.710$, $p<.001$, $p\eta^2=.205$, but no interaction between Audio and Cut Type, $F(3,78)=.155$, $p=.926$, n.s., $p\eta^2=.006$. The audio effect can clearly be seen in Figure 2. Cuts accompanied by audio took longer to detect than the same cuts presented without audio (Silent=618.4ms, *SD*=122.1; Audio=650.01ms, *SD*=146.8) and this difference was preserved across all cut types.

<<Insert Figure 2 here>>

To investigate the main effect of Cut Type, reaction times were averaged across audio conditions and paired comparisons performed between the cut conditions. The main effect of

Cut Type was mostly due to Intact cuts taking significantly longer to detect (M=680.6ms, *SD*=153.5) than –Post (M=597.23ms, *SD*=117.4; p=.005) and –Both (M=624.91ms, *SD*=125.02; p=.01) but not –Pre (M=634.07ms, *SD*=134.1; p=.111). All post-hoc t-tests are corrected for multiple comparisons using Bonferroni correction (here and throughout this manuscript). Comparison between the altered cut conditions (-Pre, -Post, -Both) reveals no significant differences.

Miss Rate

Differences in cut characteristics (Appendix Table 1) did not have any significant effect on miss rates (Size Change In vs. Out, $t(53)=1.618$, p=.112; Angle change, $F(4,79)=.741$, p=.567) so all subsequent analysis is collapsed across clip.

A within-subjects ANOVA with factors Cut Type (Intact, -Pre, -Post, -Both) and Audio (Silent vs. Audio) revealed a marginally significant main effect of Audio, $F(1,26)=3.319$, p=.080 $p\eta^2=.113$, a marginal interaction, $F(3,78)=2.187$, p=.096, $p\eta^2=.078$, and no effect of Cut Type, $F(3,78)=2.095$, p=.108, $p\eta^2=.075$, n.s.. Looking at Figure 3, the interaction can be clearly seen in the different effect of Cut Type between the two Audio conditions. Within the Audio condition there was a significant main effect of Cut Type, $F(3,78)=3.465$, p=.020, $p\eta^2=.118$, with the Intact cuts being missed significantly more often (M=.141, *SD*=.097) than the –Post (M=.078, *SD*=.131; $t(26)=2.565$, p=.016) and –Both cuts (M=.067, *SD*=.130; $t(26)=2.768$, p=.01) but not –Pre (M=.115, *SD*=.129; $t(26)=1.022$, p=.316, n.s.). Critically, -Pre cuts were also missed significantly more than –Both cuts ($t(26)=2.164$, p=.040) and there was also a trend for them to be missed more than –Post cuts ($t(26)=1.474$, p=.152, n.s.). This suggests that the presence of post-cut motion (as present in the –Pre and Intact conditions) may be more critical for the creation of edit blindness than pre-cut motion when presented with accompanying audio. When cuts were presented without audio, the main effect of Cut Type disappeared, $F(3,78)=.180$, p=.910, $p\eta^2=.007$, n.s..

<<Insert Figure 3 here>>

Saccade Frequency

Now that we have established the prominent role post-cut motion and audio appear to have on edit blindness how does this influence manifest in viewer eye movements? As predicted earlier, is there a trade-off of high-velocity eye movements (i.e. saccades) for more low velocity eye movements (i.e. fixational eye movements) immediately after the cut creating image blurring? Does this result in fewer saccadic eye movements in the conditions with the greatest cut miss rates, e.g. Intact and –Pre?

<<Insert Figure 4 here>>

Saccade frequency was calculated by identifying all saccades occurring between -200ms and +700ms around the critical cut and calculating the percentage of saccades within each 100ms time bin within this time window. Only cuts that were identified by participants (i.e. hits) were entered into the analysis as this ensures the response actions are matched across all cut types even though they differ in terms of miss rates. Analysis of saccade frequencies during misses was not possible due to the low number of misses within some conditions. As such, the saccade frequencies presented here should be considered a conservative estimate of what may be observed when the same cut types are missed.

Figure 4 shows the saccade frequency relative to the critical cut for the different cut and audio conditions. The stereotypical peak in saccadic activity 250-350ms after the cut can be seen. Prior to the cut and after this initial peak there are no noticeable differences between cut types, confirming previous evidence that there are no saccade differences across cuts of different types (Smith & Henderson, 2008)**.** Only during the peak do we see a separation in the saccadic activity between conditions. A repeated-measures ANOVA of saccade frequency revealed a main effect of cut type ($F(3,75)=3.204$, $p=.028$, $p\eta^2=.114$) and time

(F(9,225)=22.728, p=.000, $p\eta^2$=.476), no effect of Audio (F<1), and a significant three-way interaction between audio, cut type and time (F(27,675)=55.799, p=.015, $p\eta^2$=.064). The effect of time can clearly be seen in

Figure 4. Three-hundred milliseconds after the cut the frequency of saccades increases from ~8% to ~16%. The amplitude of this peak varied across cut types and across audio conditions. Whilst there was no overall main effect of audio, the differing impact of audio on each cut type within this peak can be clearly seen. Within the Audio condition during the +300ms time bin (i.e. between +250 and +349ms) there is a main effect of cut type (F(3,75)=3.340, p=.024, $p\eta^2$=.118) with the saccade frequency in the Intact edit condition being significantly lower (mean=15.49%, *SD*=6.57) than the –Both condition (mean=20.80%, *SD*=7.56, t(25)=-3.123, p=.004) and numerically, but not significantly lower than the –Pre (mean=17.54%, *SD*=7.69, t(25)=-1.199, p=.242, n.s.) and –Post (mean=17.73%, *SD*=6.62, t(25)=-1.52, p=.141, n.s.). The saccade probability in the -Both cut type during the same time bin is marginally larger than –Pre (t(25)=2.023, p=.054) and -Post (t(25)=1.713, p=.099) suggesting a potential benefit in having either motion before or after the cut compared to no motion. Although any benefit of having one source of motion over the other was not clear as –Pre and –Post do not differ significantly from Intact.

Within the Silent condition the overall pattern of saccade frequencies over time remained the same but the differences between cut types within this critical bin (+300ms) was less clear. The main effect of cut type remained during this critical +300ms bin (F(3,75)=4.096, p=.01) but the effect was driven by the -Both condition which had a significantly higher saccade frequency (mean=20.76%, *SD*=9.83) than -Pre (mean=13.97%, *SD*=7.75; t(25)=2.854, p=.009), -Post (mean=15.32%, *SD*=7.38; t(25)=-2.909, p=.008) and marginally higher than Intact cuts (mean=16.65%, *SD*=8.75; t(25)=-1.938, p=.064). No other

differences were significant confirming –as in the Audio condition- that the critical impact of either motion before or after the cut on edit blindness cannot be identified from the saccadic frequencies. The lack of differences between Intact and –Post in both audio conditions is a little surprising considering the general reaction time and miss rate differences found in this experiment. This suggests that participants are exhibiting behavioural differences (RT and Miss Rates) in response to subtler degrees of motion manipulation than are necessary to elicit differences in saccadic activity. However, the fact that only hit trials have been included in this analysis has to be considered as misses may have shown a slightly different pattern.

Gaze mobility

Given that match-action cuts are defined by the presence of visual motion before and after the cut there are two oculomotor hypotheses for how this motion may hinder awareness of the cut 1) the eyes may be stationary on the screen and the visual scene (or elements of it) move relative to the gaze creating a blurring of the visual scene on the retina, or 2) the eyes may track objects in the scene, moving relative to the background. Both conditions would blur the background image projected on to the retina, making frame-by-frame comparison across shots difficult. To tease apart these two hypotheses the average velocity of viewer gaze during the frame when the critical cut occurred was calculated for all cut and audio conditions. Gaze was parsed of blinks and saccades and then the Euclidean displacement of raw gaze during the fixation crossing the critical cut was calculated. Given that smooth pursuit movements were not identified in this dataset (see Smith & Mital, 2013 for discussion of why this is acceptable) all low velocity movements would be classified as fixations unless the velocity exceeded the saccade threshold (>30 °/s and >8000 °/s^2). Any tracking of image motion by gaze would therefore be manifest in a small but significant increase in gaze velocity during the critical fixation.

<<Insert Figure 5 here>>

Ideally, this analysis of gaze mobility would be split by audio condition, cut type and whether the cut was identified or missed. However, due to the low proportion of missed cuts in some conditions, a combined analysis of all three factors was not possible. Instead, separate analyses was performed within the factors and different numbers of participants entered into the analysis according to who returned valid trials in each condition.

Comparison of the gaze velocity between the audio (M=2.82°/s, *SD*=.731) and silent (M=2.77°/s, *SD*=.511) conditions showed no main effect, $t(26)=-.341$, p=.736. Therefore, subsequent analyses were performed collapsing across Audio conditions.

Comparing gaze velocity for cuts that were detected (i.e. hits) across the four cut types (Figure 5, green dotted line) revealed a main effect of cut type ($F(3,78)=3.517$, p=.019) which can attributed to the two cut types that have pre-cut motion (Intact and –Post) having higher gaze velocity than the two conditions without pre-cut motion (-Pre and -Both). Planned comparisons between these two groups of cuts revealed marginally significant higher velocities in Intact cuts (M=3.02°/s, *SD*=1.09) than –Pre (M=2.50°/s, *SD*=.83; $t(26)=1.77$, p=.088) and -Both (M=2.55°/s, *SD*=.82; $t(26)=1.933$, p=.064) and significantly higher velocities in –Post (M=3.17°/s, *SD*=1.06) than –Pre ($t(26)=3.078$, p=.005) and -Both ($t(26)=2.42$, p=.023). These results confirm that pre-cut motion is creating a slight (0.52°/s, i.e. Intact minus -Pre) but significant increase in gaze displacement across the screen which will have the perceptual consequence of slightly blurring the retinal image that may obscure the visual transients created by the cut.

Does higher gaze velocity lead to more missed cuts? If this were the case we would expect higher gaze velocities for misses than hits. Comparing the gaze velocity for hits to misses within each cut type only revealed a significant difference within the –Post condition, $t(17)=2.282$, p=.036. This difference can clearly be seen in Figure 5. The direction of the effect is, however, reversed. Gaze has a higher average velocity during hits (M=3.12°/s,

SD=1.15) than misses (M=2.18/s, *SD*=1.45). This suggests that, rather than forward masking the cut, the pre-cut motion in this condition sets the viewer gaze in motion but the sudden termination of this motion by the cut makes the cut more salient, resulting in less misses. By comparison, in the Intact condition the presence of both pre and post-cut motion allows viewer gaze to continue pursuing the image across the cut irrespective of whether the cut was detected. This suggests that pre-cut motion may contribute to the match-action effect by increasing gaze velocity before the cut, resulting in a slight blurring of the image. However, pre-cut motion appears less important for creating edit blindness than the backward masking effect of post-cut motion as –Pre (i.e. cuts with only post-cut motion) are missed more than –Both (see miss rates results) and the sudden cessation of motion observed in –Post may draw attention to some cuts.

GENERAL DISCUSSION

Filmmakers have been using onsets of motion to create global change blindness of film cuts for over a century. If this match-action technique operates as filmmakers believe it does it constitutes one of the most striking demonstrations of change blindness ever recorded: failure to detect a change of almost the entire visual field. In this study we endeavoured to further validate this match-action effect and investigate which components of the cut are critical for its creation. By manipulating the presence of motion before and after the critical match-action cut and whether it was played with the accompanying audio we were able to influence the degree of edit blindness experienced by participants and how they oriented to the content before and after the cut. We showed that removing post-cut motion (-*Post*) or both pre-cut and post-cut motion (-*Both*) significantly speeded up cut detection time and decreased the probability of missing the cut. By comparison, removing just pre-cut motion (-*Pre*) did not decrease the probability of missing the cut and did not significantly decrease reaction times. In fact, cuts without pre-cut motion (-Pre) were missed significantly more

than cuts without any motion (-Both). This finding is contrary to our first hypothesis which stated that edit blindness requires attention to be captured by pre-cut motion and occupied during the cut. Removing pre-cut motion should have increased awareness of the cut, which was not the case. This unexpected finding suggests that the presence of post-cut motion may be more important than pre-cut motion for the creation of edit blindness through match-action editing. Although, caution should be taken in interpreting this result as *–Pre* only resulted in a greater miss rate than *–Post* when combined with audio suggesting that visual motion alone is insufficient for the creation of a match-action cut.

Our eye movement analyses may provide some initial evidence for why removing only post-cut motion decreases edit blindness more than removing pre-cut motion. The presence of pre-cut motion without the continuing post-cut motion may draw attention to the cut due to the sudden and unexpected termination of motion. Recent evidence from a novel side-by-side film-viewing paradigm provided evidence in support of such motion tracking as participants found it easier to saccade to cuts in which motion was continuous across a cut compared to cuts without continuity of motion (Valuch, Ansorge, Buchinger, Patrone, & Scherzer, 2014). Our analysis of gaze mobility also indicated that *–Post* cuts that were detected had higher gaze velocity than the same cuts that were missed. This suggests that, as hypothesised, gaze was tracking motion prior to the cut and the sudden termination of this motion may have alerted viewers to the existence of the cut, resulting in an increase in saccade frequency in response to the new shot content (see –Both vs. Intact). Only when combined with post-cut motion does the displacement of gaze continue across the cut allow retinal blurring of the image to minimize viewer awareness of the cut. However, caution should be taken when extrapolating cognitive mechanisms for edit blindness during normal movie free-viewing from these findings as it is not known how the primary task of cut detection in the present study affected eye movements. The peak in saccades immediately

following all critical cuts suggests that normal viewing behaviour was being observed but without a free-view control condition the possibility that viewers were adopting a specific "cut detection" viewing strategy cannot be ruled out. Future studies should establish whether the cut detection viewing task alters viewer gaze behaviour and also perform more precise manipulation of pre- and post-cut motion by changing the timing of cuts in purposely shot films.

One of the most surprising findings of this study was the critical role audio played in the creation of edit blindness in match-action cuts. Cut detection time was significantly quicker when clips were presented without audio and the impact of cut type on miss rate disappeared. Technical guidelines for match-action editing (Anderson, 1996; Bordwell & Thompson, 2001; Dmytryk, 1986; Reisz & Millar, 1953) focus almost entirely on the visual aspects, e.g. motion onset before the cut followed by continuation of motion after the cut. Our results suggest that these visual features will only limit viewer awareness of the cut if combined with the continuous perceptual "scaffold" of a soundtrack. Without audio flowing across the cut participants were either less engaged with the visual content (and therefore, not influenced by visual motion) or have more cognitive resources to allocate to the primary task, i.e. cut detection (Lang, 2000). The slower reaction times in the Audio condition support the latter hypothesis: cut detection could be considered a dual task 1) watching and understanding the audiovisual narrative and 2) detecting visual cuts. Even though participants were instructed to only detect cuts, by its nature Hollywood cinema is designed to captivate the viewer and draw them into the narrative (Smith, 2013). From our results it appears that the soundtrack is critical for maintaining this engagement with the visual content. Several recent studies have suggested similar changes in viewing activity when watching videos without sound (Coutrot et al., 2012; Hirvenkari et al., 2013; Smith, 2014; Song et al., 2013; Võ et al., 2012). When viewing videos with their corresponding audio, there were differences

in fixation durations and saccade amplitudes and participant gaze was more coordinated than when the same videos were presented in silence (Coutrot et al., 2012; Song et al., 2013). Gaze is less focussed on faces when watching clips of dialogues without audio, shifts to the speaker are delayed (Hirvenkari et al., 2013) and focus more on the speaker's eyes and less on their mouth (Võ et al., 2012). These differences seem to rely on diegetic (i.e. "on screen") sounds as gaze behaviour during scenes with only off-screen sounds (Song et al., 2013) or scenes accompanied by a musical score differ less when the audio is removed (Smith, 2014). The majority of the scenes used in the present study depicted two or more actors engaged in a dialogue. It may be that participant interest in the dialogue, facilitated by the sound design of the scene within and across the critical cut helped viewer attention transition across the cut and maintain its partial allocation to the audio, limiting the availability of resources for the cut detection task. A similar impact of audio on reaction times immediately following cuts has recently been demonstrated in a probe detection task (Shimamura, 2013).

The use of "found" stimuli in this study (i.e. clips gathered from existing feature films) rather than specifically constructed stimuli allows our findings to be directly related to the experience of watching real film but the limitations of using such naturalistic stimuli are the lack of control. It is possible that the manipulations we made to pre and post-cut motion differ in some way other than the simple removal of motion and that this confound may account for our findings. Whilst our initial control study (see Stimuli section) demonstrated that our manipulations didn't introduce any detectable artefacts the duration of pre and post-cut motion is inherently mismatched and this might explain the cut detection differences. The editing guidelines for creating a match-action cut recommend that the cut occur just after the onset of motion with the following shot depicting the remainder of the motion. Removing the longer period of post-cut motion may have a greater likelihood of creating a visual or semantic incongruity (known in cinema practice as a 'continuity error', such as the

disappearance of an object or change in posture). Examining the average number of frames removed in our stimuli this mismatch in duration is apparent: -Pre clips =14.8 frames removed *(SD*=8.19), -Post = 21.5 frames (*SD*=14.34), -Both=34.9 frames (*SD*=17.67). The number of frames removed for each clip significantly correlated with the mean probability of missing that clip ($r(480)=-0.203$, $p<.001$) but not the reaction time for detection ($r(468)=-.007$, $p=.875$, n.s.). Splitting this correlation by Cut Type it is clear that the effect is driven by clips that involve removing the longer post-cut motion, –Post clips, $r(160)=-0.195$, $p=.013$, and –Both, $r(160)=-0.153$, $p=.053$, but not –Pre ($p=.353$, n.s.). To check that this duration difference could not be accounting for our edit detection and reaction time effects the clips were categorised in terms of the relative number of frames removed before and after the cut. Only clips for which the number of post-cut frames removed was similar to the number of pre-cut frames were entered into the analyses (the number of -Post frames removed <1.5 times –Pre; number of clips remaining = 41). The pattern of misses and reaction times for these matched clips was the same as reported in the main analysis. Therefore, whilst the removal of a large number of post-cut frames in some clips may have increased cut detection it does not appear that these extreme cases can account for our results. It still appears that it is the presence of motion and audio immediately around the cut that is important for the creation of edit blindness.

The *Attentional Theory of Cinematic Continuity* (AToCC; Smith, 2012a), as discussed in the introduction emphasises the role pre-cut motion played in capturing attention and guiding it across the cut. The current findings show that post-cut motion may be more important for creating edit blindness than pre-cut motion. This emphasis on forward prediction and attentional shifts in AToCC has previously been criticised by Levin and Hymel (2012) in their response to AToCC. They presented empirical evidence of failures to detect impossible orderings of actions in an edited sequence (e.g. using a screwdriver before

it was picked up) and proposed that constant prediction was computationally inefficient and a more appropriate mechanism would be postdiction, perceiving the logical sequence of events after the fact. The prediction assumption of AToCC was adapted from event segmentation theory (Zacks & Magliano, 2013) in which constant prediction about the short and long term form of events is used to identify moments when certainty in the future drops below an acceptable threshold and a critical boundary between events is perceived. Such event boundaries are thought to be accompanied by an increase in attention in order to encode new information and formulate new predictions (Levin & Saylor, 2008; Zacks & Magliano, 2013). Evidence for peak attention during event boundaries has been provided by fMRI studies (Zacks et al., 2001) and eye tracking which has shown a peak in attentional synchrony (i.e. coordination of gaze across multiple viewers) at event boundaries primarily directed towards the actor's hands (fine boundaries) or face (coarse boundaries) (Smith, 2012c). Sudden onsets of motion, like those used in the match-action cuts in the present study, may sometimes constitute fine event boundaries and, therefore indicate a momentary breakdown in prediction, contrary to that originally proposed by AToCC. As such, it is the new information presented after the onset of motion that is critical to the perception of the new event, not whether it matches predictions formulated before the onset of motion. The post-cut motion may operate much like the global motion used to induce motion silencing (Suchow & Alvarez, 2011) or backward masking (Breitmeyer & Öğmen, 2000) in more simple displays. An unexpected onset of motion may negate on-going prediction and destabilise the visual array making detection of change (i.e. the cut) hard to detect.

The editing literature generally specifies onsets of motion prior to the cut as essential to the creation of a match-action cut. However, the degree of motion used in examples is often miniscule e.g. a head turn, lifting hand or even actor gaze shift (Dmytryk, 1986; Pepperman, 2004), suggesting that the onset's purpose is simply to attract attention and alert

the viewer to a change. The content of the following shot may be more important for maintaining the *a priori* assumption of continuity as is evidenced by the tendency for editors to overlap the action presented in the following shot by a few frames (Anderson, 1996). In an empirical study in which participants had to choose the degree of overlap between a Match-Action cut depicting a woman drinking from a cup, participants perceived an overlap of three frames (~125ms) as continuous even if the cut occurred to a shot from the same viewpoint separated by a noise mask (Shimamura, Cohn-Sheehy, & Shimamura, 2014; also see Hecht & Kalkofen, 2009). The authors interpreted this effect as being the consequence of recovering from the attentional shift induced by the cut, a similar explanation to that derived from event segmentation theory (Zacks & Magliano, 2013). Therefore, a slight modification to AToCC may be required to accommodate our present findings. Some cuts, such as match-action cuts, may not entail the pre-cut formulation of expectations about the future form of events but may instead either create an absence of expectation (due to a unexpected change in action) or absence of perception due to backward masking of the previous visual information by sudden motion onset. Both situations provide an opportunity for attention to be attracted by post-cut motion and cued to the beginning of a new event sequence. This modification to AToCC enables the theory to accommodate a broader range of cuts (see Smith, 2012b for further extensions).

The findings of the present study and their formulation within AToCC may appear specific to film editing (such as the unexpected importance of audio continuity) and, therefore of marginal relevance to broader cognition. However, we would argue that edit blindness constitutes a *limit case* of audiovisual scene perception. Demonstrations of change blindness in a range of situations and degrees of change have been informative for questioning prior assumptions about the completeness of visual representation (Henderson & Hollingworth, 1999; Rensink, 2000) but most theories of visual perception would assume that a change to

the entire visual scene would be perceived by viewers. Edit blindness is equivalent to a viewer not noticing an instantaneous transportation to a new location while their eyes are open. While demonstration of this effect in the real world is physically impossible film provides a window into this extreme example of audiovisual perception and reveals insights on attention, motion processing, spatial representations and event perception (see Smith, Levin, et al., 2012; for further discussion). As such, we believe the intuitions of film makers formalised in the Hollywood style and Continuity Editing rules provide an incredible wealth of craft knowledge about audiovisual perception that psychologists should exploit in the same way recent scientific investigations have endeavoured to unlock the psychology of music (Levitin, 2006), visual art (Bacci & Melcher, 2013), and magic (Kuhn, Amlani, & Rensink, 2008).

REFERENCES

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science, 14*, 427-432.

Anderson, J. D. (1996). *The Reality of Illusion: An Ecological Approach to Cognitive Film Theory*. Carbondale and Edwardsville, US: Southern. Illinois University Press.

Bacci, F., & Melcher, D. (2013). *Art and the Senses*: Oxford University Press.

Bordwell, D., Staiger, J., & Thompson, K. (1985). *The Classical Hollywood Cinema: Flm Style & Mode of Production to 1960*. London, UK: Routledge.

Bordwell, D., & Thompson, K. (2001). *Film Art: An Introduction* (Vol. 6th). New York, USA: Mc Graw Hill.

Breitmeyer, B., & Öğmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62*(8), 1572-1595.

Bureau of Labor Statistics. (2014). American Time Use Survey.

Carmi, R., & Itti, L. (2006). Visual causes versus correlates of attention selection in dynamic scenes. *Vision Research, 46*(2006), 4333.

Carroll, J. M., & Bever, T. G. (1976). Segmentation in Cinema Perception. *Science, 191*(4231), 1053-1055.

Coutrot, A., Guyader, N., Ionescu, G., & Caplier, A. (2012). Influence of soundtrack on eye movements during video exploration. *Journal of Eye Movement Research, 5*(4:2), 1-10.

d'Ydewalle, G., Desmet, G., & Van Rensbergen, J. (1998). Film Perception: The processing of film cuts. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 357-367). Oxford: Elsevier.

Dmytryk, E. (1986). *On Filmmaking*. London, UK: Focal Press.

Frith, U., & Robson, J. E. (1975). Perceiving the language of films. *Perception, 4*(1), 97-103.

Geiger, S., & Reeves, B. (1993). The effects of scene changes and semantic relatedness on attention to television. *Communication Research, 20*, 155-175.

Germeys, F., & d'Ydewalle, G. (2007). The psychology of film: perceiving beyond the cut. *Psychological Research, 71*(4), 458-466.

Hasson, U., Landesman, O., Knappmeyer, B., Valines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The Neuroscience of Film. Projections: The Journal of Movies and Mind, 2(1), 1-26.

Hecht, H., & Kalkofen, H. (2009) Questioning the rules of continuity editing: An empirical study. *Empirical Studies of the Arts, 27*, 1–23

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243-271.

Hirvenkari, L., Ruusuvuori, J., Saarinen, V.-M., Kivioja, M., Peräkylä, A., & Hari, R. (2013). Influence of Turn-Taking in a Two-Person Conversation on the Gaze of a Viewer. *PLoS One, 8*(8).

Hochberg, J., & Brooks, V. (1978). Film Cutting and Visual Momentum. In J. W. Senders, D. F. Fisher & R. A. Monty (Eds.), *Eye Movements and the Higher Psychological Functions* (pp. 293-317). Hillsdale, NJ: Lawrence Erlbaum.

Ildirar, S., Levin, D. T., Schwan, S., & Smith, T. J. (2014). *The Role of Audio on Perception of Continuity by First-Time Viewers,*. Paper presented at the Society of the Cognitive Study of the Moving Image Lancaster, Pennsylvania, USA.

Ildirar, S., & Schwan, S. (2015). First-time viewers' comprehension of films: Bridging shot transitions. *British Journal of Psychology, 106*, 133-151. doi:10.1111/bjop.12069.

Kuhn, G., Amlani, A. A., & Rensink, R. A. (2008). Towards as science of magic. *Trends in Cognitive Sciences, 12*(9), 349-335.

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. In F. Biocca (Ed.), *Television and political advertising, Vol. I: Psychological processes*. Hillsdale, NJ: Erlbaum.

Lang, A. (2000). The Limited Capacity model of Mediated Message Processing. . *Journal of Communication, 50*, 46-70.

Lang, A., Kurita, S., Gao, Y., & Rubenking, B. (2013). Measuring television message complexity as available processing resources: Dimensions of information and cognitive load. *Media Psychology, 16*(2), 129-153.

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.

Levin, D. T., & Hymel, A. M. (2012). Making the case for nonpredictive continuity perception. *Projections, 6*(1), 61-70.

Levin, D. T., & Saylor, M. M. (2008). Shining spotlights, zooming lenses, grabbing hands, and pecking chickens: The ebb and flow of attention during events. . In T. Shipley & J. M. Zacks (Eds.), *Understanding events: From perception to action* (pp. pp. 522-554). New York: Oxford Univ. Press.

Levitin, D. J. (2006). *This Is Your Brain on Music: The Science of a Human Obsession*: Dutton Adult.

Mera, M., & Stumpf, S. (2014) Eye-Tracking Film Music. *Music and the Moving Image, 7*(3), 3-23.

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. *Cognitive Computation, 3*(1), 5-24.

Murch, W. (2001). *In The Blink Of An Eye: a perspective on film editing*. Los Angeles, USA: Silman-James Press.

Nannicelli, T. & Taberham, P. (2014) *Cognitive Media Theory,* AFI Film Readers. New York, USA: Routledge,

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of “mudsplashes”. . *Nature, 398*(34).

Pepperman, R. D. (2004). *The Eye is Quicker: Film Editing: Making a good Film better*. Studio City, CA, US: Michael Wiese Productions.

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*(1-3), 17-42.

Rensink, R. A., ORegan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8*(5), 368-373.

Ross, N. M., & Kowler, E. (2013). Eye movements while viewing narrated, captioned and silent videos. *Journal of Vision, 13*(4).

Salt, B. (2009). *Film Style and Technology: History and Analysis* (Vol. 3rd). Totton, Hampshire, UK: Starword.

Schwan, S., & Ildirar, S. (2010). Watching film for the first time: How adult viewers interpret perceptual discontinuities in film. *Psychological Science, 21*(7), 970-976.

Shimamura, A. P. (2013), *Psychocinematics: Exploring Cognition at the Movies.* New York: Oxford University Press,

Shimamura, A. P. (2013). *Multimodal Influences on Attention Across Film Edits: An Empirical Analysis.* Paper presented at the Society for Cognitive Studies of The Moving Image, Berlin, Germany.

Shimamura, A. P., Cohn-Sheehy, B. I., & Shimamura, T. A. (2014). Perceiving movement across film edits: A psychocinematic analysis. *Psychology of Aesthetics, Creativity, and the Arts*. Vol 8(1), 77-80.

Smith, T. J. (2012a). The Attentional Theory of Cinematic Continuity. *Projections: The Journal for Movies and the Mind, 6*(1), 1-27.

Smith, T. J. (2012b). Extending AToCC: a reply. *Projections: The Journal for Movies and the Mind, 6*(1), 71-78.

Smith, T. J. (2012c). The relationship between overt attention and event perception during dynamic social scenes. *Journal of Vision, 12*(9), 407.

Smith, T. J. (2013). Watching you watch movies: Using eye tracking to inform cognitive film theory. In A. P. Shimamura (Ed.), *Psychocinematics: Exploring Cognition at the Movies*. New York: Oxford University Press.

Smith, T. J. (2014). Audiovisual correspondences in Sergei Eisenstein's Alexander Nevsky: a case study in viewer attention. In P. Taberham & T. Nannicelli. (Eds.), *Cognitive Media Theory (AFI Film Reader)*.

Smith, T. J., & Henderson, J. M. (2008). Edit Blindness: The Relationship Between Attention and Global Change Blindness in Dynamic Scenes. *Journal of Eye Movement Research, 2*(2), 1-17.

Smith, T. J., Lamont, P., & Henderson, J. M. (2012). The penny drops: Change blindness at fixation. *Perception, 41*(4), 489 - 492.

Smith, T. J., Lamont, P., & Henderson, J. M. (2013). Change blindness in a dynamic scene due to endogenous override of exogenous attentional cues. *Perception, 42*, 884-886.

Smith, T. J., Levin, D. T., & Cutting, J. E. (2012). A Window on Reality: Perceiving Edited Moving Images. *Current Directions in Psychological Science, 21*, 101-106.

Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behaviour in static and dynamic scenes. *Journal of Vision, 13*(8), 16.

Song, G., Pellerin, D., & Granjon, L. (2013). Different types of sounds influence gaze differently in videos. *Journal of Eye Movement Research, 6*(4), 1-13.

Suchow, J. W., & Alvarez, G. A. (2011). Motion Silences Awareness of Visual Change. *Current Biology*, 140-143.

Tse, P. (2009). Smooth-pursuit eye-movements suppress motion processing. *Journal of Vision, 9*(8).

Valuch, C., Ansorge, U., Buchinger, S., Patrone, A. R. & Scherzer, O. (2014). The effect of cinematic cuts on human attention. In *Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video* (TVX '14). ACM, New York, NY, USA, 119-122.

Võ, M. L.-H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the Eyes Really Have it? Dynamic Allocation of Attention when Viewing Moving Faces. *Journal of Vision, 12*(13).

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*, 1-7.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., . . . Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*, 651-655.

Zacks, J. M., & Magliano, J. P. (2013). Film, Narrative, and Cognitive Neuroscience. In F. B. D. P. Melcher (Ed.), *Art and the Senses*. New York: Oxford University Press.

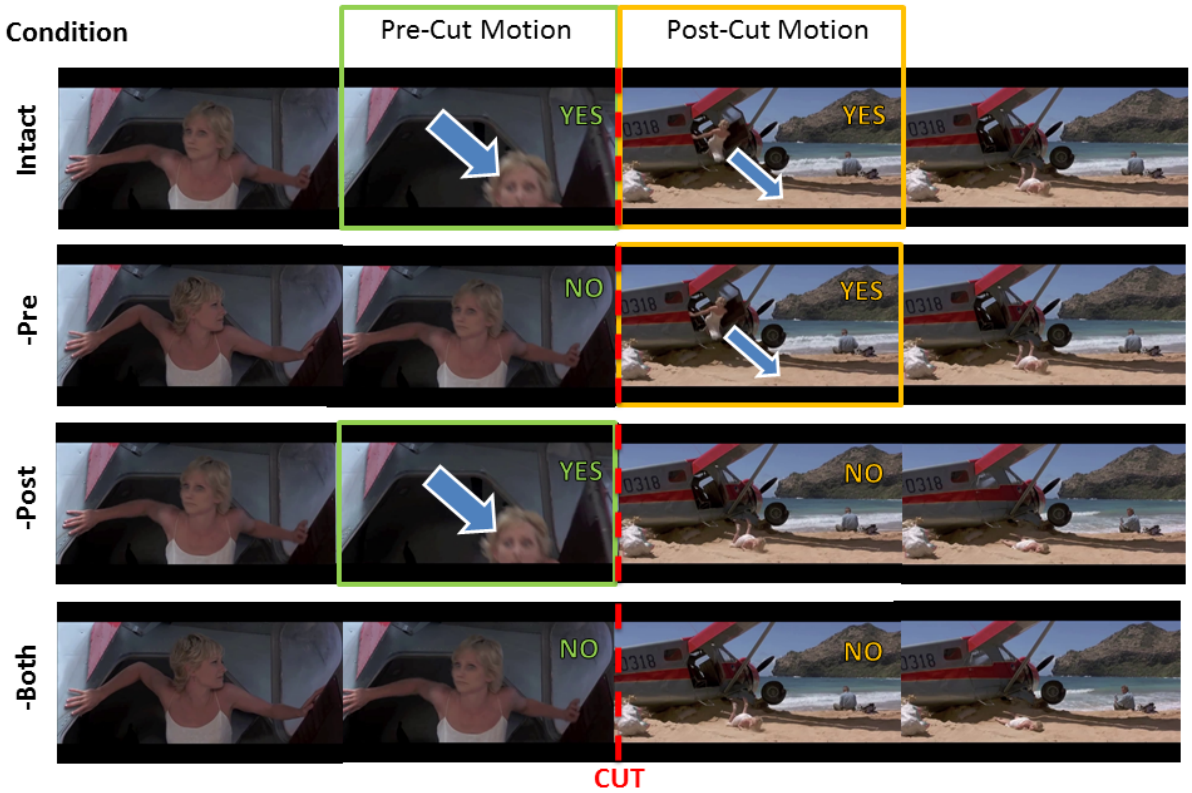

Figure 1: Example of one of the film clips used in this study (taken from Six Days Seven Nights, Ivan Reitman, 1998). Anne Heche's character awakes in a crashed plane and falls out of the door. The clip cuts (via a Match-Action) to her falling on the sand while Harrison Ford's character looks on. Rows depict the categories of cuts used in this study: Top row= full motion (*Intact*); Second row= no pre-cut motion (-*Pre*); Third row= no post-cut motion (-*Post*); Bottom row = neither pre-cut nor post-cut motion (-*Both*). Each cut type was presented with or without original audio. See Video 1 for demonstrations of these manipulations.

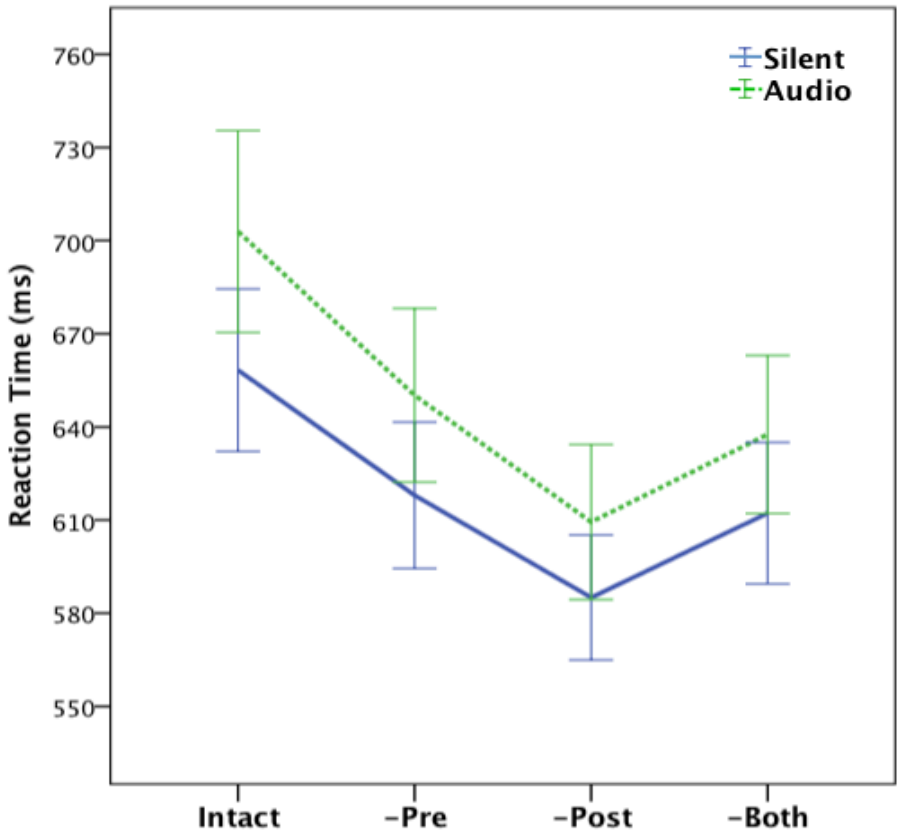

Figure 2: Mean time taken to detect the critical cut (ms) across the four cut types (Intact, -Pre, -Post, -Both) and Audio conditions (Audio vs. Silent).

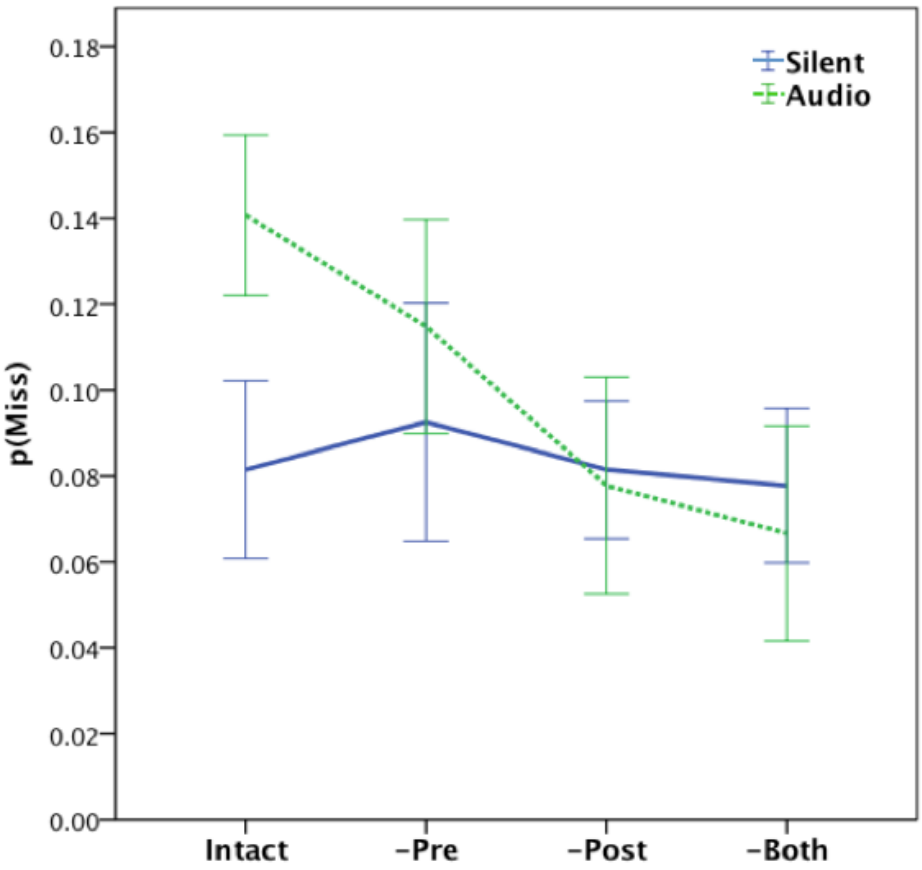

Figure 3: Mean miss rate for the critical cut (probability) across the four cut types (Intact, -Pre, -Post, -Both) and Audio conditions (Audio vs. Silent).

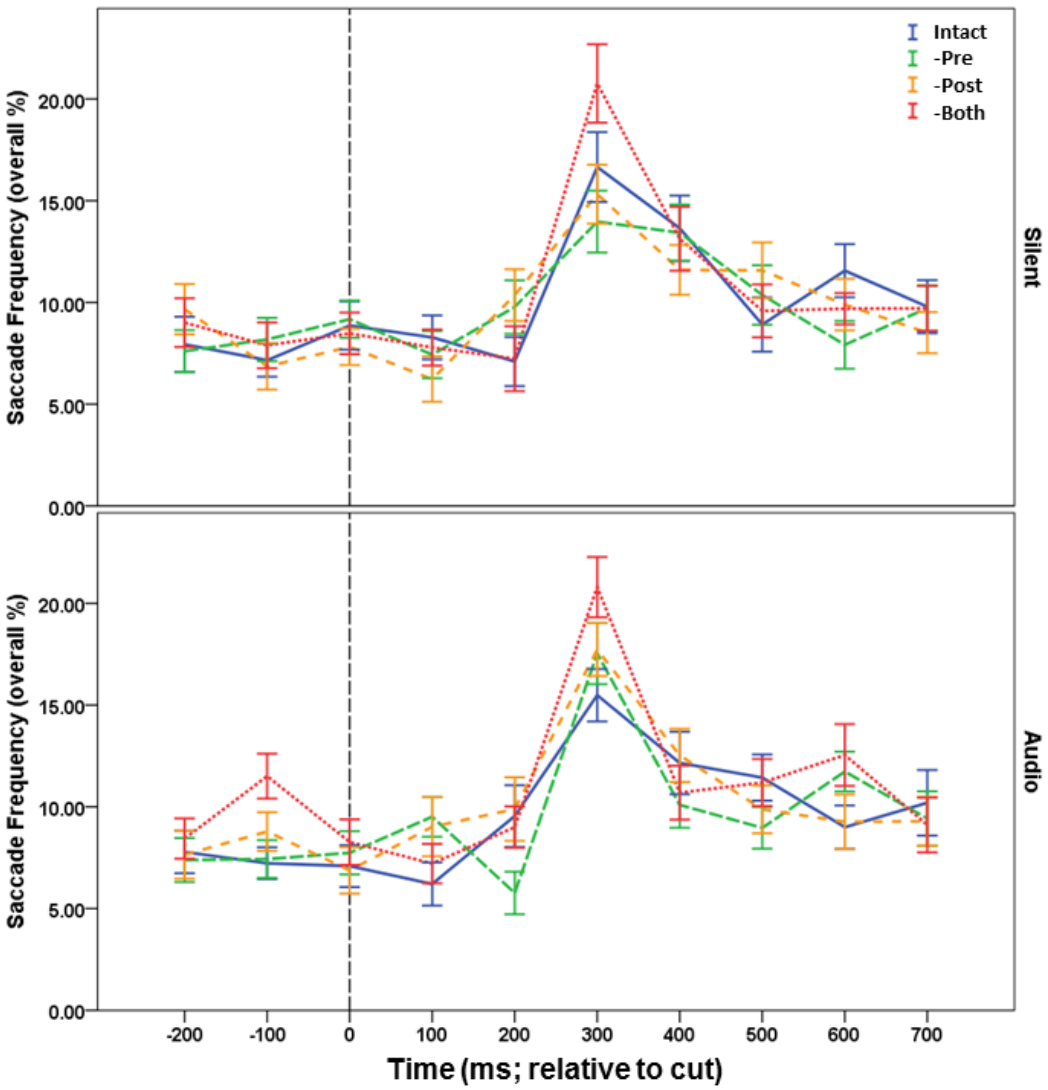

Figure 4: Saccade frequency (% of saccades in each condition) relative to the critical cut (ms; 100ms time bins) across the four edit conditions (Intact=blue solid; -Pre=green dashed; -Post=orange small dash; -Both = red dotted) and the two audio conditions (Silent=Top panel; Audio=Bottom panel).

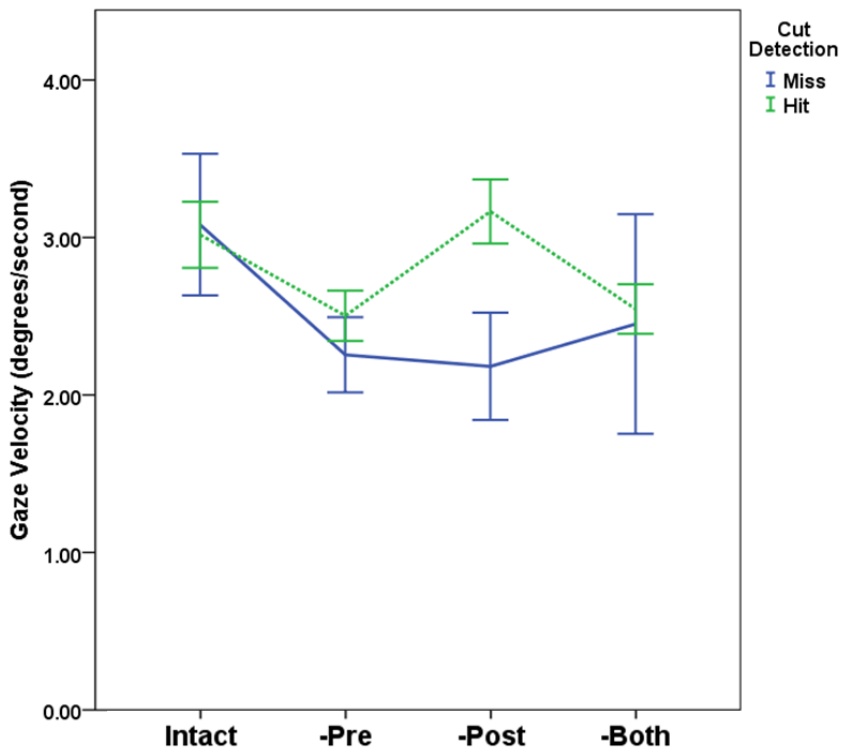

Figure 5: Mean velocity of gaze during the critical cut across the four cut types (Intact, -Pre, -Post, and -Both). Data is split according to whether cut was detected (green dotted line) or missed (blue solid line).

Video Captions

Video 1: Demonstrations of the cut type and audio manipulations made to one clip used in this study, *Six Days Seven Nights* (Ivan Reitman, 1998). Miss probabilities are included in the video as well as a slow-motion break-down of the motion manipulated across cut conditions. http://youtu.be/RJfOPHicLKY

Supplementary Material

Tablet Caption

Table 1: Films from which match-action cuts were sourced for the present study. Each clip is characterised in terms of shot size for the pre and post-cut shots, whether the cut involved a change in shot size, a change in camera angle on the subject, and the body part and action that was used to match the shots across the cut. Key to abbreviations: Shot Size= Close Up (CU), Medium Close-Up (MCU), Medium Shot (MS), Medium Long-Shot (MLS), Long Shot (LS), eXtreme Long Shot (XLS). See Bordwell and Thompson (2001) for details of this coding scheme.

Appendix A

	Film		Date	Studio	Pre-Cut Size	Post-Cut Size	Size Change	Angle Change	Body Part	Action
1	About A Boy	1	2002	Universal Pictures	MS	MS	NONE	45	HANDS	Reach out
2	Along Came Polly	1	2004	Universal Pictures	LS	CU	IN	135	HEAD	Turn
3	American President	1	1995	Universal Pictures	CMS	MLS	OUT	45	HANDS	Glass chink
4	Arachnophobia	1	1990	Hollywood Pictures	MLS	MLS	NONE	90	WHOLE BODY	Walk
5	As Good As It Gets	1	1997	TriStar Pictures	MLS	MLS	NONE	90	WHOLE BODY	Turn
6		2			MS	MLS	OUT	90	WHOLE BODY	Turn
7	Barefoot in the Park	1	1967	Paramount Pictures	MLS	MS	IN	45	HANDS	Reach out
8		2			MLS	MS	IN	0	HANDS	Pick up
9	Bowfinger	1	1999	Universal Pictures	MS	LS	OUT	90	WHOLE BODY	Turn
10		2			MLS	MS	IN	90	HANDS	Touch
11	Bruce Almighty	1	2003	Spyglass Ent.	MS	MS	NONE	90	HEAD	Turn
12	Catch Me If You Can	1	2002	Dreamworks STG	MS	CU	IN	90	HEAD	Turn
13	Chocolat	1	2000	Miramax	MLS	MLS	NONE	90	HANDS	Lift object
14		2			MLS	MLS	NONE	90	HEAD	Turn
15		3			LS	MS	IN	90	WHOLE BODY	Crouch
16	Contact	1	1997	Warner Bros.	MLS	MS	IN	45	WHOLE BODY	Sit
17	The Day After Tomorrow	1	2004	Twentieth Century Fox	LS	MLS	IN	0	WHOLE BODY	Lift object
18	Star Wars: Episode V - The Empire Strikes Back	1	1980	Lucasfilm	CU	CU	NONE	90	WHOLE BODY	Turn
19	E.T. the Extra-Terrestrial	1	1982	Universal Pictures	MLS	MS	IN	0	WHOLE BODY	Turn and crouch
20		2			MS	MS	NONE	90	HANDS	Lift object
21	The Lord of the Rings: The Fellowship of the Ring	1	2001	New Line Cinema	CU	CU	NONE	90	HEAD	Turn
22	Fellowship of the Ring	1			MLS	CU	IN	90	HEAD	Turn
23	Finding Neverland	1	2004	Miramax	MLS	CU	IN	45	HEAD	Lift up
24	The First Wives Club	1	1996	Paramount Pictures	LS	MS	IN	90	WHOLE BODY	Lift up
25	First Wives Club	1			MS	MLS	OUT	90	WHOLE BODY	Throw body forward and down
26	For Love Of The Game	1	1999	Universal Pictures	CU	MLS	OUT	0	HEAD	Turn
27		2			CU	MLS	OUT	45	TORSO	Sit back
28	Ghostbusters	1	1984	Columbia Pictures	MCU	MLS	OUT	45	WHOLE BODY	Turn corner
29	Grosse Pointe Blank	1	1997	Hollywood Pictures	MS	MLS	OUT	0	WHOLE BODY	Stand up
30	Groundhog Day	1	1993	Columbia Pictures	CU	MLS	OUT	45	HEAD	Look down
31		2			MS	MCU	IN	45	HANDS	Eat cake

	Film		Date	Studio	Pre-Cut Size	Post-Cut Size	Size Change	Angle Change	Body Part	Action
32	Hitch	1	2005	Columbia Pictures	CU	LS	OUT	0	WHOLE BODY	Turn (run)
33	Holiday	1	2006	Columbia Pictures	MLS	MS	IN	45	HANDS	Pull out letter
34		2			CU	MLS	OUT	90	HANDS	Lift glass to mouth
35	How To Lose A Guy in 10 Days	1	2003	Paramount	MLS	MCU	IN	90	HANDS	Pick up phone and put to ear
36		2			CU	MS	OUT	0	WHOLE BODY	Stand up
37	Jerry Maguire	1	1996	TriStar Pictures	MCU	LS	OUT	0	WHOLE BODY	Walk forward (step through door)
38		2			MCU	MS	OUT	45	WHOLE BODY	Turn + walk
39	Lawrence of Arabia	1	1962	Horizon Pictures	MLS	CU	IN	0	HANDS	Lift match
40	Legally Blonde	1	2001	Metro-Goldwyn-Mayer	MLS	MLS	NONE	90	HANDS	Lift cake
41	Meet the Fockers	1	2004	Universal Pictures	MCU	MCU	NONE	90	HANDS	Lift to mouth
42		2			MCU	MS	OUT	90	WHOLE BODY	Walk forward
43	Miss Pettigrew Lives For A Day	1	2008	Focus Features	CU	CU	NONE	90	HEAD	Lean in for kiss
44	Moonstruck	1	1987	Metro-Goldwyn-Mayer	LS	MS	IN	45	WHOLE BODY	Sit down
45	Mr Hollands Opus	1	1995	Hollywood Pictures	MS	LS	OUT	180	HANDS	Reach out and in (open window)
46	Music of the Heart	1	1999	Miramax	MLS	MS	IN	45	WHOLE BODY	Turn
47	My Big Fat Greek Wedding	1	2002	Gold Circle Films	MS	MCU	IN	90	HANDS	Pull arm back
48		2			MS	MS	NONE	90	WHOLE BODY	Stand up
49	Nanny McPhee	1	2005	Universal Pictures	CU	MS	OUT	90	HEAD	Slight turn
50	Notting Hill	1	1999	Working Title Films	LS	MS	IN	45	HANDS	Drop
51	One True Thing	1	1998	Universal Pictures	MLS	LS	OUT	90	WHOLE BODY	Reach and turn
52	Out of Africa	1	1985	Mirage Enterprises	CU	MLS	OUT	90	WHOLE BODY	Down/up crank an engine
53		2			MS	MS	NONE	90	WHOLE BODY	Lean back
54	Philadelphia	1	1993	TriStar Pictures	MS	MS	NONE	90	HANDS	Pour tea
55	Pirates of the Caribbean: The Curse of the Black Pearl	1	2003	Walt Disney Pictures	MS	MS	NONE	90	HEAD	Turn
56	Poltergeist	1	1982	Metro-Goldwyn-Mayer	MS	LS	OUT	0	WHOLE BODY	Tilt up (from lean)
57		2			LS	MCU	IN	45	WHOLE BODY	Kneel down
58	Seabiscuit	1	2003	Universal Pictures	LS	CU	IN	90	WHOLE BODY	Sit down + turn

	Film		Date	Studio	Pre-Cut Size	Post-Cut Size	Size Change	Angle Change	Body Part	Action
59	Shadowlands	1	1993	Price Entertainment	MCU	MLS	OUT	90	WHOLE BODY	Lie down + turn
60	Six Days Seven Nights	1	1998	Caravan Pictures	MS	LS	OUT	0	WHOLE BODY	Fall to floor
61		2			MS	MS	NONE	180	HANDS	Reach into bag and pull something out
62	The Sound of Music	1	1965	Robert Wise Productions	XLS	LS	IN	0	WHOLE BODY	Stand up
63	Spider-Man	1	2002	Columbia Pictures	XLS	MS	IN	90	WHOLE BODY	Step up
64	Stranger Than Fiction	1	2006	Columbia Pictures	MLS	MS	IN	0	TORSO	Lean forward + head turn
65	Terminator 2: Judgment Day	1	1991	Carolco Pictures	CU	CU	NONE	45	HEAD	Turn
66	Terminal	1	2004	Dreamworks STG	MS	MLS	OUT	45	HANDS	Close tin
67	The Associate	1	1996	Hollywood Pictures	MS	MS	NONE	90	WHOLE BODY	Walk
68	The Fifth Element	1	1997	Gaumont	MS	MS	NONE	45	HANDS	Lift glass to mouth
69	The Interpreter	1	2005	Universal Pictures	MLS	MLS	NONE	135	WHOLE BODY	Sit down
70	The Mummy	1	1999	Universal Pictures	MS	MCU	IN	0	HANDS	Take glasses off (hands up then down)
71	The Princess Bride	1	1987	Act III Communications	MLS	MS	IN	90	WHOLE BODY	Turn + lean in
72	Thelma and Louise	1	1991	Pathé Entertainment	MS	MS	NONE	90	WHOLE BODY	Walk through doorway
73	Titanic	1	1997	Twentieth Century Fox	MLS	MS	IN	0	HANDS	Take off coat
74	The Towering Inferno	1	1974	Warner Bros.	MS	MS	NONE	90	WHOLE BODY	Turn + walk
75		2			MLS	MS	IN	45	WHOLE BODY	Sit down
76	The Truth About Cats & Dogs	1	1996	Twentieth Century Fox	MCU	MCU	NONE	180	HEAD	Turn
77	Unknown	1	2011	Dark Castle Entertainment	CU	MLS	OUT	90	WHOLE BODY	Walk forward + lean in
78	Victor Victoria	1	1982	Metro-Goldwyn-Mayer	MCU	CU	IN	90	TORSO	Lean in for kiss
79	While You Were Sleeping	1	1995	Hollywood Pictures	MLS	MS	IN	90	TORSO	Lean down
80	You Only Live Twice	1	1967	Eon Productions	MS	MS	NONE	0	WHOLE BODY	Walk through doorway + lean down