Read, Watch, Listen: A commentary on eye tracking and moving images.

Tim J. Smith

Department of Psychological Sciences, Birkbeck, University of London

Keywords: film, eye tracking, experimentation, statistics, psychology, subtitles, digital effects,

Biography:

Dr. Tim J. Smith BSc. Hons, PhD (Edin.). is a senior lecturer in the Department of Psychological, Sciences, Birkbeck, University of London. He applies empirical Cognitive Psychology methods including eye tracking to questions of Film Cognition and has published extensively on the subject both in Psychology and Film journals.

Abstract

Eye tracking is a research tool that has great potential for advancing our understanding of how we watch movies. Questions such as how differences in the movie influences where we look and how individual differences between viewers alters what we see can be operationalised and empirically tested using a variety of eye tracking measures. This special issue collects together an inspiring interdisciplinary range of opinions on what eye tracking can (and cannot) bring to film and television studies and practice. In this article I will reflect on each of these contributions with specific focus on three aspects: how subtitling and digital effects can reinvigorate visual attention, how audio can guide and alter our visual experience of film, and how methodological, theoretical and statistical considerations are paramount when trying to derive conclusions from eyetracking data.

Introduction

I have been obsessed with how people watch movies since I was a child. All you have to do is turn and look at an audience member's face at the movies or at home in front of the TV to see the power the medium holds over them. We sit enraptured, transfixed and immersed in the sensory patterns of light and sound projected back at us from the screen. As our physical activity diminishes our mental activity takes over. We piece together minimal audiovisual cues to perceive rich otherworldly spaces, believable characters and complex narratives that engage us mentally and move us emotionally. As I progressed through my education in Cognitive Science and Psychology I was struck by how little science understood about cinema and the mechanisms filmmakers used to create this powerful experience[i]. Reading the film literature, listening to filmmakers discuss their craft and excavating gems of their craft knowledge I started to realise that film was a medium ripe for psychological investigation. The empirical study of film would further our understanding of how films work and how we experience them but it would also serve as a test bed for investigating complex aspects of real-world cognition that were often considered beyond the realms of experimentation. As I (Smith, Levin & Cutting, 2010) and others (Anderson, 2006) have argued elsewhere, film evolved to "piggy back" normal cognitive development and use basic cognitive tendencies such as attentional preferences, theory of mind, empathy and narrative structuring of memory to make the perception of film as enjoyable and effortless as possible. By investigating film cognition we can, in turn advance our understanding of general cognition. But to do so we need to step outside of traditional

disciplinary boundaries concerning the study of film and approach the topic from an interdisciplinary perspective. This special issue represents a highly commendable attempt to do just that.

By bringing together psychologists, film theorists, philosophers, vision scientists, neuroscientists and screen writers this special issue (and the Melbourne research group that most contributors belong to) provides a unique perspective on film viewing. The authors included in this special issue share my passion for understanding the relationship between viewers and film but this interest manifests in very different ways depending on their perspectives (see Redmond, Sita, & Vincs, this issue; for a similar personal journey into eyetracking as that presented above). By focussing on viewer eye movements the articles in this special issue provide readers from a range of disciplines a way into the eyetracking investigation of film viewing. Eye tracking (as comprehensively introduced and discussed by Dyer & Pink, this issue) is a powerful tool for quantifying a viewer's experience of a film, comparing viewing behaviour across different viewing conditions and groups as well as testing hypotheses about how certain cinematic techniques impact where we look. But, as is rightly highlighted by several of the authors in this special issue eyetracking is not a panacea for all questions about film spectatorship. Like all experimental techniques it can only measure a limited range of psychological states and behaviours and the data it produces does not say anything in and of itself. Data requires interpretation. Interpretation can take many forms[ii] but if conclusions are to be drawn about how the data relates to psychological states of the viewer this interpretation must be based on theories of psychology and ideally confirmed using secondary/supporting measures. For example, the affective experience of a movie is a critical aspect which cognitive approaches to film are often wrongly accused of ignoring. Although, cognitive approaches to film often focus on how we comprehend narratives (Magliano & Zacks, 2011), attend to the image (Smith, 2013) or follow formal patterns within a film (Cutting, DeLong & Nothelfer, 2010) several cognitivists have focussed in depth on emotional aspects (see the work of Carl Plantinga, Torben Grodal or Murray Smith). Eye tracking is the perfect tool for investigating the impact of immediate audiovisual information on visual attention but it is less suitable for measuring viewer affect. Psychophysiological measures such as heart rate and skin conductance, neuroimaging methods such as fMRI or EEG, or even self-report ratings may be better for capturing a viewer's emotional responses to a film as has been demonstrated by several research teams (Suckfull, 2000; Raz et al, 2014). Unless the emotional state of the viewer changed where they looked or how quickly they moved their eyes the eyetracker may not detect any differences between two viewers with different emotional states[iii]. As such, a researcher interested in studying the emotional impact of a film should either choose a different measurement technique or combine eyetracking with another more suitable technique (Dyer & Pink, this issue). This does not mean that eyetracking is unsuitable for studying the cinematic experience. It simply means that you should always choose the right tool for the job and often this means combining multiple tools that are strong in different ways. As Murray Smith (the current President of the Society for the Cognitive Study of the Moving Images; SCSMI) has argued, a fully rounded investigation of the cinematic experience requires "triangulation" through the combination of multiple perspectives including psychological, neuroscientific and phenomenological/philosophical theory and methods (Smith, 2011). An approach taken proudly across this special issue.

For the remainder of my commentary I would like to focus on certain themes that struck me as most personally relevant and interesting when reading the other articles in this special issue. This is by no means an exhaustive list of the themes raised by the other articles or even an assessment of the importance of the particular themes I chose to select. There are many other interesting observations made in the articles I do not focus on below but given my perspective as a cognitive scientist and current interests I decided to focus my commentary on these specific themes rather than make a comprehensive review of the special issues or tackle topics I am unqualified to comment on. Also, I

wanted to take the opportunity to dispel some common misconceptions about eyetracking (see the section 'Listening to the data') and empirical methods in general.

Reading an image

One area of film cognition that has received considerable empirical investigation is subtitling. As Kruger, Szarkowska and Krejtz (this issue) so comprehensively review, they and I believe eyetracking is the perfect tool for investigating how we watch subtitled films. The presentation of subtitles divide the film viewing experience into a dual- task: reading and watching. Given that the media was originally designed to communicate critical information through two channels, the image and soundtrack introducing text as a third channel of communication places extra demands on the viewer's visual system. However, for most competent readers serially shifting attention between these two tasks does not lead to difficulties in comprehension (Kruger, Szarkowska and Krejtz, this issue). Immediately following the presentation of the subtitles gaze will shift to the beginning of the text, saccade across the text and return to the centre of interest within a couple of seconds. Gaze heatmaps comparing the same scenes with and without subtitles (Kruger, Szarkowska and Krejtz, this issue; Fig. 3) show that the areas of the image fixated are very similar (ignoring the area of the screen occupied by the subtitles themselves) and rather than distracting from the visual content the presence of subtitles seems to actually condense the gaze behaviour on the areas of central interest in an image, e.g. faces and the centre of the image. This illustrates the redundancy of a lot of the visual information presented in films and the fact that under non-subtitle conditions viewers rarely explore the periphery of the image (Smith, 2013).

My colleague Anna Vilaró and I recently demonstrated this similarity in an eyetracking study in which the gaze behaviour of viewers was compared across versions of an animated film, Disney's *Bolt* (Howard & Williams, 2008) either in the original English audio condition, a Spanish language version with English subtitles, an English language version with Spanish subtitles and a Spanish language version without subtitles (Vilaró, & Smith, 2011). Given that our participants were English speakers who did not know Spanish these conditions allowed us to investigate both where they looked under the different audio and subtitle conditions but also what they comprehended. Using cued recall tests of memory for verbal and visual content we found no significant differences in recall for either types of content across the viewing conditions except for verbal recall in the Spanish-only condition (not surprisingly given that our English participants couldn't understand the Spanish dialogue). Analysis of the gaze behaviour showed clear evidence of subtitle reading, even in the Spanish subtitle condition (see Figure *1*) but no differences in the degree to which peripheral objects were explored. This indicates that even when participants are watching film sequences without subtitles and know that their memory will be tested for the visual content their gaze still remains focussed on central features of a traditionally composed film. This supports arguments for subtitling movies over dubbing as, whilst placing greater demands on viewer gaze and a heightened cognitive load there is no evidence that subtitling leads to poorer comprehension.

<<Insert Figure 1 about here>>

The high degree of *attentional synchrony* (Smith & Mital, 2013) observed in the above experiment and during most film sequences indicates that all visual features in the image and areas of semantic

significance (e.g. social information and objects relevant to the narrative) tend to point to the same part of the image (Mital, Smith, Hill & Henderson, 2011). Only when areas of the image are placed in conflict through image composition (e.g. depth of field, lighting, colour or motion contrast) or staging (e.g. multiple actors) does attentional synchrony break down and viewer gaze divide between multiple locations. Such shots are relatively rare in mainstream Hollywood cinema or TV (Salt, 2009; Smith, 2013) and when used the depicted action tends to be highly choreographed so attention shifts between the multiple centres of image in a predictable fashion (Smith, 2012). If such choreographing of action is not used the viewer can quickly exhaust the information in the image and start craving either new action or a cut to a new shot. Hochberg and Brooks (1978) referred to this as the *visual momentum* of the image: the pace at which visual information is acquired. This momentum is directly observable in the saccadic behaviour during an images presentation with frequent short duration fixations at the beginning of a scene's presentation interspersed by large amplitude saccades (known as the *ambient* phase of viewing; Velichovsky, Dornhoefer, Pannasch & Unema, 2000) and less frequent, longer duration fixations separated by smaller amplitude saccades as the presentation duration increases (known as the *focal* phase of viewing; Velichovsky et al., 2000). I have recently demonstrated the same pattern of fixations during viewing of dynamic scenes (Smith & Mital, 2013) and shown how this pattern gives rise to more central fixations at shot onset and greater exploration of the image and decreased attentional synchrony as the shot duration increases (Mital, Smith, Hill & Henderson, 2011). Interestingly, the introduction of subtitles to a movie may have the unintended consequence of sustaining visual momentum throughout a shot. The viewer is less likely to exhaust the information in the image because their eyes are busy saccading across the text to acquire the information that would otherwise be presented in parallel to the image via the soundtrack. This increased saccadic activity may increase the cognitive load experienced by viewers of subtitled films and change their affective experience, producing greater arousal and an increased sense of pace.

For some filmmakers and producers of dynamic visual media, increasing the visual momentum of an image sequence may be desirable as it maintains interest and attention on the screen (e.g. Michael Bay's use of rapidly edited extreme Close-Ups and intense camera movements in the *Transformer* movies). In this modern age of multiple screens fighting for our attention when we are consuming moving images (e.g. mobile phones and computer screens in our living rooms and even, sadly increasingly at the cinema) if the designers of this media are to ensure that our visual attention is focussed on their screen over the other competing screens they need to design the visual display in a way that makes comprehension impossible without visual attention. Feature Films and Television dramas often rely heavily on dialogue for narrative communication and the information communicated through the image may be of secondary narrative importance to the dialogue so viewers can generally follow the story just by listening to the film rather than watching it. If producers of dynamic visual media are to draw visual attention back to the screen and away from secondary devices they need to increase the ratio of visual to verbal information. A simple way of accomplishing this is to present the critical audio information through subtitling. The more visually attentive mode of viewing afforded by watching subtitled film and TV may partly explain the growing interest in foreign TV series (at least in the UK) such as the popularity of Nordic Noir series such as *The Bridge* (2011) and *The Killing* (2007).

Another way of drawing attention back to the screen is to constantly "refresh" the visual content of the image by either increasing the editing rate or creatively using digital composition[iv]. The latter technique is wonderfully exploited by *Sherlock* (2010) as discussed brilliantly by Dyer (this issue). *Sherlock* contemporised the detective techniques of Sherlock Holmes and John Watson by incorporating modern technologies such as the Internet and mobile phones and simultaneously

updated the visual narrative techniques used to portray this information by using digital composition to playfully superimpose this information onto the photographic image. In a similar way to how the sudden appearance of traditional subtitles involuntarily captures visual attention and draws our eyes down to the start of the text, the digital inserts used in *Sherlock* overtly capture our eyes and encourage reading within the viewing of the image. If Dyer (this issue) had eyetracked viewers watching these excerpts she would have likely observed this interesting shifting between phases of reading and dynamic scene perception. Given that the appearance of the digital inserts produce sudden visual transients and are highly incongruous with the visual features of the background scene they are likely to involuntarily attract attention (Mital, Smith, Hill & Henderson, 2012). As such, they can be creatively used to reinvigorate the pace of viewing and strategically direct visual attention to parts of the image away from the screen centre. Traditionally, the same content may have been presented either verbally as narration, heavy handed dialogue exposition (e.g. "Oh my! I have just received a text message stating….") or as a slow and laboured cut to close-up of the actual mobile phone so we can read it from the perspective of the character. Neither takes full advantage of the communicative potential of the whole screen space or our ability to rapidly attend to and comprehend visual information and audio information in parallel.

Such intermixing of text, digital inserts and filmed footage is common in advertisements, music videos, and documentaries (see Figure *2*) but is still surprisingly rare in mainstream Western film and TV. Short-form audiovisual messages have recently experienced a massive increase in popularity due to the internet and direct streaming to smartphones and mobile devices. To maximise their communicative potential and increase their likelihood of being "shared" these videos use all audiovisual tricks available to them. Text, animations, digital effects, audio and classic filmed footage all mix together on the screen, packing every frame with as much info as possible (Figure 2), essentially maximising the visual momentum of each video and maintaining interest for as long as possible[v]. Such videos are so effective at grabbing attention and delivering satisfying/entertaining/informative experiences in a short period of time that they often compete directly with TV and film for our attention. Once we click play, the audiovisual bombardment ensures that our attention remains latched on to the *second screen* (i.e. the tablet or smartphone) for its duration and away from the *primary screen,* i.e. the TV set. Whilst distressing for producers of TV and Film who wish our experience of their material to be undistracted, the ease with which we pick up a handheld device and seek other stimulation in parallel to the primary experience may indicate that the primary material does not require our full attention for us to follow what is going on. As attention has a natural ebb-and-flow (Cutting, DeLong, & Nothelfer, 2010) and "There is no such thing as voluntary attention sustained for more than a few seconds at a time" (p. 421; James, 1890) if modern producers of Film and TV want to maintain a high level of audience attention and ensure it is directed to the screen they must either rely on viewer self-discipline to inhibit distraction, reward attention to the screen with rich and nuanced visual information (as fans of "slow cinema" would argue of films like those of Bela Tarr) or utilise the full range of postproduction effects to keep visual interest high and maintained on the image, as *Sherlock* so masterfully demonstrates.

<< Insert Figure 2 about here >>

A number of modern filmmakers are beginning to experiment with the language of visual storytelling by questioning our assumptions of how we perceive moving images. Forefront in this

movement are Ang Lee and Andy and Lana Wachowski. In Ang Lee's *Hulk* (2003), Lee worked very closely with editor Tim Squyers to use non-linear digital editing and after effects to break apart the traditional frame and shot boundaries and create an approximation of a comicbook style within film. This chaotic unpredictable style polarised viewers and was partly blamed for the film's poor reception. However, it cannot be argued that this experiment was wholly unsuccessful. Several sequences within the film used multiple frames, split screens, and digital transformation of images to increase the amount of centres of interest on the screen and, as a consequence increase pace of viewing and the arousal experienced by viewers. In the sequence depicted below (Figure 3) two parallel scenes depicting Hulk's escape from a containment chamber (A1) and this action being watched from a control room by General Ross (B1) were presented simultaneously by presenting elements of both scenes on the screen at the same time. Instead of using a *point of view* (POV) shot to show Ross looking off screen (known as the *glance* shot; Branigan, 1984) followed by a cut to what he was looking at (the *object* shot) both shots were combined into one image (F1 and F2) with the latter shot sliding into from behind Ross' head (E2). These digital inserts float within the frame, often gliding behind objects or suddenly enlarging to fill the screen (A2-B2). Such visual activity and use of shots-within-shots makes viewer gaze highly active (notice how the gaze heatmap is rarely clustered in one place; Figure 3). Note that this method of embedding a POV object shot within a glance shot is similar to *Sherlock's* method of displaying text messages as both the *glance*, i.e. Watson looking at his phone, and the *object*, i.e. the message, are shown in one image. Both uses take full advantage of our ability to rapidly switch from watching action to reading text without having to wait for a cut to give us the information.

<< Insert Figure 3 about here >>

Similar techniques have been used Andy and Lana Wachowski's films including most audaciously in *Speed Racer* (2008). Interestingly, both sets of filmmakers seem to intuitively understand that packing an image with as much visual and textual information as possible can lead to viewer fatigue and so they limit such intense periods to only a few minutes and separate them with more traditionally composed sequences (typically shot/reverse-shot dialogue sequences). These filmmakers have also demonstrated similar respect for viewer attention and the difficulty in actively locating and encoding visual information in a complex visual composition in their more recent 3D movies. Ang Lee's *Life of Pi* (2012) uses the visual volume created by stereoscopic presentation to its full potential. Characters inhabit layers within the volume as foreground and background objects fluidly slide around each other within this space. The lessons Lee and his editor Tim Squyers learned on *Hulk* (2003) clearly informed the decisions they made when tackling their first 3D film and allowed them to avoid some of the issues most 3D films experience such as eye strain, sudden unexpected shifts in depth and an inability to ensure viewers are attending to the part of the image easiest to fuse across the two eye images (Banks, Read, Allison & Watt, 2012).

Watching Audio

I now turn to another topic featured in this special issue, the influence of audio on gaze (Robinson, Stadler & Rassel, this issue). Film and TV are inherently multimodal. Both media have always existed as a combination of visual and audio information. Even early silent film was almost always presented with either live musical accompaniment or a narrator. As such, the relative lack of empirical

investigation into how the combination of audio and visual input influences how we perceive movies and, specifically how we attend to them is surprising. Robinson, Stadler and Rassel (this issue) have attempted to address this omission by comparing eye movements for participants either watching the original version of the Omaha beach sequence from Steven Spielberg's *Saving Private Ryan* (1998) or the same sequence with the sound removed. This film sequence is a great choice for investigating AV influences on viewer experience as the intensity of the action, the hand-held cinematography and the immersive soundscape all work together to create a disorientating embodied experience for the viewer. The authors could have approached this question by simply showing a set of participants the sequence with audio and qualitatively describing the gaze behaviour at interesting AV moments during the sequence. Such description of the data would have served as inspiration for further investigation but in itself can't say anything about the causal contribution of audio to this behaviour as there would be nothing to compare the behaviour to. Thankfully, the authors avoided this problem by choosing to manipulate the audio.

In order to identify the causal contribution of any factor you need to design an experiment in which that factor (known as the *Independent Variable*) is either removed or manipulated and the significant impact of this manipulation on the behaviour of interest (known as the *Dependent Variable*) is tested using appropriate inferential statistics. I commend, Robinson, Stadler and Rassel's experimental design as they present such an manipulation and are therefore able to produce data that will allow them to test their hypotheses about the causal impact of audio on viewer gaze behaviour. Several other papers in this special issue (Redmond, Sita, & Vincs, this issue; Batty, Perkins, & Sita, this issue) discuss gaze data (typically in the form of scanpaths or heatmaps) from one viewing condition without quantifying its difference to another viewing condition. As such, they are only able to describe the gaze data, not use it to test hypotheses. There is always a temptation to attribute too much meaning to a gaze heatmap (I too am guilty of this; Smith, 2013) due to their seeming intuitive nature (i.e. they looked here and not there) but, as in all psychological measures they are only as good as the experimental design within which there are employed[vi].

Qualitative interpretation of individual fixation locations, scanpaths or group heatmaps are useful for informing initial interpretation of which visual details are most likely to make it into later visual processing (e.g. perception, encoding and long term memory representations) but care has to be taken in falsely assuming that fixation equals awareness (Smith, Lamont, & Henderson, 2012). Also, the visual form of gaze heatmaps vary widely depending on how many participants contribute to the heatmap, which parameters you choose to generate the heatmaps and which oculomotor measures the heatmap represent (Holmqvist, et al., 2011). For example, I have demonstrated that unlike during reading visual encoding during scene perception requires over 150ms during each fixation (Rayner, Smith, Malcolm, & Henderson, 2009). This means that if fixations with durations less than 150ms are included in a heatmap it may suggest parts of the image have been processed which in actual fact were fixated too briefly to be processed adequately. Similarly, heatmaps representing fixation duration instead of just fixation location have been shown to be a better representation of visual processing (Henderson, 2003). Heatmaps have an immediate allure but care has to be taken about imposing too much meaning on them especially when the gaze and the image are changing over time (see Smith & Mital, 2013; and Sawahata et al, 2008 for further discussion). As eyetracking hardware becomes more available to researchers from across a range of disciplines we need to work harder to ensure that it is not used inappropriately and that the conclusions that are drawn from eyetracking data are theoretically and statistically motivated (see Rayner, 1998; and Holmqvist et al, 2013 for clear guidance on how to conduct sound eyetracking studies).

Given that Robinson, Stadler and Rassel (this issue) manipulated the critical factor, i.e. the presence of audio the question now is whether their study tells us anything new about the AV influences on gaze during film viewing. To examine the influence of audio they chose two traditional methods for expressing the gaze data: area of interest (AOI) analysis and dispersal. By using nine static (relative to the screen) AOIs they were able to quantify how much time the gaze spent in each AOI and utilise this measure to work out how distributed gaze was across all AOIs. Using these measures they reported a trend towards greater dispersal in the mute condition compared to the audio condition and a small number of significant differences in the amount of time spent in some regions across the audio conditions. However, the conclusions we can draw from these findings are seriously hindered by the low sample size (only four participants were tested meaning that any statistical test is unlikely to reveal significant differences) and the static AOIs that did not move with the image content. By locking the AOIs to static screen coordinates their AOI measures express the deviation of gaze relative to these coordinates, not to the image content. This approach can be informative for quantifying gaze exploration away from the screen centre (Mital. Smith, Hill & Henderson, 2011) but in order to draw conclusions about what was being fixated the gaze needs to be quantified relative to dynamic AOIs that track objects of interest on the screen (see Smith & Mital, 2013). For example, their question about whether we fixate a speaker's mouth more in scenes where the clarity of the speech is difficult due to background noise (i.e. their "Indistinct Dialogue" scene) has previously been investigated in studies that have manipulated the presence of audio (Võ, Smith, Mital, & Henderson, 2012) or the level of background noise (Buchan, Paré, & Munhall, 2007) and measured gaze to dynamic mouth regions. As Robinson, Stadler & Rassel correctly predicted, lip reading increases as speech becomes less distinct or the listener's linguistic competence in the spoken language decreases (see Võ et al, 2012 for review).

Similarly, by measuring gaze dispersal using a limited number of static AOIs they are losing considerable nuance in the gaze data and have to resort to qualitative description of unintuitive bar charts (figure 4). There exist several methods for quantifying gaze dispersal (see Smith & Mital, 2013, for review) and even open-source tools for calculating this measure and comparing dispersal across groups (Le Meur & Baccino, 2013). Some methods are as easy, if not easier to calculate than the static AOIs used in the present study. For example, the Euclidean distance between the screen centre and the x/y gaze coordinates at each frame of the movie provides a rough measure of how spread out the gaze is from the screen centre (typically the default viewing location; Mital et al, 2011) and a similar calculation can be performed between the gaze position of all participants within a viewing condition to get a measure of group dispersal. Using such measures, Coutrot and colleagues (2012) showed that gaze dispersal is greater when you remove audio from dialogue film sequences and they have also observed shorter amplitude saccades and marginally shorter fixation durations. Although, I have recently shown that a non-dialogue sequence from Sergei Eisenstein's *Alexander Nevsky* (1938) does not show significant differences in eye movement metrics when the accompanying music is removed (Smith, 2014). This difference in findings points towards interesting differences in the impact *diegetic* (within the depicted scene, e.g. dialogue) and *non-diegetic* (outside of the depicted scene, e.g. the musical score) may have on gaze guidance. It also highlights how some cinematic features may have a greater impact on other aspects of a viewer's experience than those measureable by eyetracking such as physiological markers of arousal and emotional states. This is also the conclusion that Robinson, Stadler and Rassel's (this issue) come to.

Listening to the data (aka what is eyetracking good for?)

The methodological concerns I have raised in the previous section lead nicely to the article by William Brown in this special issue entitled *There's no I in Eye-tracking: How useful is Eye-tracking to Film Studies* (this issue). I have known William Brown for several years through our attendance of the Society for Cognitive Studies of the Moving Image (SCSMI) annual conference and I have a deep respect for his philosophical approach to film and his ability to incorporate empirical findings from the cognitive neurosciences, including some references to my own work into his theories. Therefore, it comes somewhat as a surprise that his article openly attacks the application of eyetracking to film studies. However, I welcome Brown's criticisms as it provides me with an opportunity to address some general assumptions about the scientific investigation of film and hopefully suggest future directions in which eyetracking research can avoid falling into some of the pitfalls Brown identifies.

Brown's main criticisms of current eyetracking researchare 1) eye tracking studies neglect "marginal" viewers or marginal ways of watching movies; 2) studies so far have neglected "marginal" films; 3) only provide "truisms", i.e. already known facts; and 4) have an implicit political agenda to argue that the only "true" way to study film is a scientific approach and the "best" way to make a film is to ensure homogeneity of viewer experience. I will address these criticisms in turn but before I do so I would like to state that a lot of Brown's arguments could generally be recast as an argument against science in general and are built upon a misunderstanding of how scientific studies should be conducted and what they mean.

To respond to Brown's first criticism that eyetracking "has up until now been limited somewhat by its emphasis on statistical significance – or, put simply, by its emphasis on telling us what most viewers look at when they watch films" (Brown, this issue; pg 1), I first have to subdivide the criticism into 'the search for significance' and 'attentional synchrony', i.e. how similar gaze is across viewers (Smith & Mital, 2013). Brown tells an anecdote about a Dutch film scholar who's data had to be excluded from an eyetracking study because they didn't look where the experimenter wanted them to look. I wholeheartedly agree with Brown that this sounds like a bad study as data should never be excluded for subjective reasons such as not supporting the hypothesis, i.e. looking as predicted. However, exclusion due to statistical reasons is valid if the research question being tested relates to how representative the behaviour of a small set of participants (known as the *sample*) are to the overall population. To explain when such a decision is valid and to respond to Brown's criticism about only 'searching for significance' I will first need to provide a brief overview of how empirical eye tracking studies are designed and why significance testing is important.

For example, if we were interested in the impact sound had on the probability of fixating an actor's mouth (a la Robinson, Stadler and Rassel, this issue) we would need to compare the gaze behaviour of a sample of participants who watch a sequence with the sound turned on to a sample who watched it with the sound turned off. By comparing the behaviour between these two groups using inferential statistics we are testing the likelihood that these two viewing conditions would differ in a population of all viewers given the variation within and between these two groups. In actual fact we do this by performing the opposite test: testing the probability that that the two groups belong to a single statistically indistinguishable group. This is known as the *null hypothesis*. By showing that there is less than a 5% chance that the null hypothesis is true we can conclude that there is a statistically significant chance that another sample of participants presented with the same two viewing conditions would show similar differences in viewing behaviour.

In order to test whether our two viewing conditions belong to one or two distributions we need to be able to express this distribution. This is typically done by identifying the mean score for each participant on the dependent variable of interest, in this case the probability of fixating a dynamic mouth AOI then calculating the mean for this measure across all participants within a group and

their variation in scores (known as the *standard deviation*). Most natural measures produce a distribution of scores looking somewhat like a bell curve (known as the *normal distribution*) with most observations near the centre of the distribution and an ever decreasing number of observations as you move away from this central score. Each observation (in our case, participants) can be expressed relative to this distribution by subtracting the mean of the distribution from its score and dividing by the standard deviation. This converts a raw score into a *normalized* or *z-score*. Roughly ninety-five percent of all observations will fall within two standard deviations of the mean for normally distributed data. This means that observations with a z-score greater than two are highly unrepresentative of that distribution and may be considered *outliers*. However, being unrepresentative of the group mean is insufficient motivation to exclude a participant. The outlier still belongs to the group distribution and should be included unless there is a supporting reason for exclusion such as measurement error, e.g. poor calibration of the eyetracker. If an extreme outlier is not excluded it can often have a disproportionate impact on the group mean and make statistical comparison of groups difficult. However, if this is the case it suggests that the sample size is too small and not representative of the overall population. Correct choice of sample size given an estimate of the predicted effect size combined with minimising measurement error should mean that subjective decisions do not have to be made about who's data is "right" and who should be included or excluded.

Brown also believes that eyetracking research has so far marginalised viewers who have atypical ways of watching film, such as film scholars either by not studying them or treating them as statistical outliers and excluding them from analyses. However, I would argue that the only way to know if their way of watching a film is atypical is to first map out the distribution of how viewers typically watch films. If a viewer attended more to the screen edge than the majority of other viewers in a random sample of the population (as was the case with Brown's film scholar colleague) this should show up as a large z-score when their gaze data is expressed relative to the group on a suitable measure such as Euclidean distance from the screen centre. Similarly, a non-native speaker of English may have appeared as an outlier in terms of how much time they spent looking at the speaker's mouth in Robinson, Stadler and Rassel's (this issue) study. Such idiosyncracies may be of interest to researchers and there are statistical methods for expressing emergent groupings within the data (e.g. cluster analysis) or seeing whether group membership predicts behaviour (e.g. regression). These approaches may have not previously been applied to questions of film viewing but this is simply due to the immaturity of the field and the limited availability of the equipment or expertise to conduct such studies.

In my own recent work I have shown how viewing task influences how we watch unedited video clips (Smith & Mital, 2013), how infants watch TV (Wass & Smith, in press), how infant gaze differs to adult gaze (Smith, Dekker, Mital, Saez De Urabain, & Karmiloff-Smith, in prep) and even how film scholars attend to and remember a short film compared to non-expert film viewers (Smith & Smith, in prep). Such group viewing differences are of great interest to me and I hope these studies illustrate how eye tracking has a lot to offer to such research questions if the right statistics and experimental designs are employed.

Brown's second main criticism is that the field of eyetracking neglects "marginal" films. I agree that the majority of films that have so far been used in eyetracking studies could be considered mainstream. For example, the film/TV clips used in this special issue include *Sherlock* (2010), *Up* (2009), and *Saving Private Ryan* (1998). However, this limit is simply a sign of how few eyetracking studies of moving images there have been. All research areas take time to fully explore the range of possible research questions within that area.

I have always employed a range of films from diverse film traditions, cultures, and languages. My first published eyetracking study (Smith & Henderson, 2008) used film clips from *Citizen Kane* (1941), *Dogville* (2003), *October* (1928), *Requiem for a Dream* (2000), *Dancer in the Dark* (2000), *Koyaanisqatsi* (1982) and *Blade Runner* (1982). Several of these films may be considered "marginal" relative to the mainstream. If I have chosen to focus most of my analyses on mainstream Hollywood cinema this is only because they were the most suitable exemplars of the phenomena I was investigating such as continuity editing and its creation of a universal pattern of viewing (Smith, 2006; 2012). This interest is not because, as Brown argues I have a hidden political agenda or an implicit belief that this style of filmmaking is the "right" way to make films. I am interested in this style because it is the dominant style and, as a cognitive scientist I wish to use film as a way of understanding how most people process audiovisual dynamic scenes. Hollywood film stands as a wonderfully rich example of what filmmakers think "fits" human cognition. By testing filmmaker intuitions and seeing what impact particular compositional decisions have on viewer eye movements and behavioural responses I hope to gain greater insight into how audiovisual perception operates in non-mediated situations (Smith, Levin & Cutting, 2012). But, just as a neuropsychologist can learn about typical brain function by studying patients with pathologies such as lesions and strokes, I can also learn about how we perceive a "typical" film by studying how we watch experimental or innovative films. My previous work is testament to this interest (Smith, 2006; 2012a; 2012b; 2014; Smith & Henderson, 2008) and I hope to continue finding intriguing films to study and further my understanding of film cognition.

One practical reason why eyetracking studies rarely use foreign language films is the presence of subtitles. As has been comprehensively demonstrated by other authors in this special issue (Kruger, Szarkowska and Krejtz, this issue) and earlier in this article, the sudden appearance of text on the screen, even if it is incomprehensible leads to differences in eye movement behaviour. This invalidates the use of eye tracking as a way to measure how the filmmaker intended to shape viewer attention and perception. The alternatives would be to either use silent film (an approach I employed with *October*; Smith & Henderson, 2008), remove the audio (which changes gaze behaviour and awareness of editing; Smith & Martin-Portugues Santacreau, under review) or use dubbing (which can bias the gaze down to the poorly synched lips; Smith, Batten, & Bedford, 2014). None of these options are ideal for investigating foreign language sound film and until there is a suitable methodological solution this will restrict eyetracking studies to experimental films in a participant's native language.

Finally, I would like to counter Brown's assertion that eyetracking investigations of film have so far only generated "truisms". I admit that there is often a temptation to reduce empirical findings to simplified take-home messages that only seem to confirm previous intuitions such as a bias of gaze towards the screen centre, towards speaking faces, moving objects or subtitles. However, I would argue that such messages fail to appreciate the nuance in the data. Empirical data correctly measured and analysed can provide subtle insights into a phenomenon that subjective introspection could never supply. For example, film editors believe that an impression of continuous action can be created across a cut by overlapping somewhere between two (Anderson, 1996) and four frames (Dmytryk, 1986) of the action. However, psychological investigations of time perception revealed that our judgements of duration depend on how attention is allocated during the estimated period (Zakay & Block, 1996) and will vary depending on whether our eyes remain still or saccade during the period (Yarrow et al, 2001). In my thesis (Smith, 2006) I used simplified film stimuli to investigate the role that visual attention played in estimation of temporal continuity across a cut and found that participants experienced an overlap of 58.44ms as continuous when an unexpected cut occurred during fixation and an omission of 43.63ms as continuous when they performed a saccade in

response to the cut. As different cuts may result in different degrees of overt (i.e. eye movements) and covert attentional shifts these empirical findings both support editor intuitions that temporal continuity varies between cuts (Dmytryk, 1986) whilst also explaining the factors that are important in influencing time perception at a level of precision not possible through introspection.

Reflecting on our own experience of a film suffers from the fact that it relies on our own senses and cognitive abilities to identify, interpret and express what we experience. I may feel that my experience of a dialogue sequence from *Antichrist* (2010) differs radically from a similar sequence from *Secrets & Lies* (1996) but I would be unable to attribute these differences to different aspects of the two scenes without quantifying both the cinematic features and my responses to them. Without isolating individual features I cannot know their causal contribution to my experience. Was it the rapid camera movements in *Antichrist*, the temporally incongruous editing, the emotionally extreme dialogue or the combination of these features that made me feel so unsettled whilst watching the scene? If you are not interested in understanding the causal contributions of each cinematic decision to an audience member's response then you may be content with informed introspection and not find empirical hypothesis testing the right method for you. I make no judgement about the validity of either approach as long as each researcher understands the limits of their approach. Introspection utilises the imprecise measurement tool that is the human brain and is therefore subject to distortion, human bias and an inability to extrapolate the subjective experience of one person to another. Empirical hypothesis testing also has its limitations: research questions have to be clearly formulated so that hypotheses can be stated in a way that allows them to be statistically tested using appropriate observable and reliable measurements. A failure at any of these stages can invalidate the conclusions that can be drawn from the data. For example, an eyetracker may be poorly calibrated resulting in an inaccurate record of where somebody was looking or it could be used to test an ill formed hypothesis such as how a particular film sequence caused attentional synchrony without having another film sequence to compare the gaze data to. Each approach has its strength and weaknesses and no single approach should be considered "better" than any other, just as no film should be considered "better" than any other film.

Conclusion

The articles collected here constitute the first attempt to bring together interdisciplinary perspectives on the application of eyetracking to film studies. I fully commend the intention of this special issue and hope that it encourages future researchers to conduct further studies using these methods to investigate research questions and film experiences we have not even conceived of. However, given that the recent release of low-cost eyetracking peripherals such as the EyeTribe[vii] tracker and the Tobii EyeX[viii] has recently moved eyetracking from a niche and highly expensive research tool to an accessible option for researchers in a range of disciplines I need to take this opportunity to issue a word of warning. As I have outlined in this article, eyetracking is like any other research tool in that it is only useful if used correctly, its limitations are respected, its data is interpreted through the appropriate application of statistics and conclusions are only drawn that are based on the data in combination with a sound theoretical base. Eyetracking is not the "saviour" of film studies , nor is science the only "valid" way to investigate somebody's experience of a film. Hopefully, the articles in this special issue and the ideas I have put forward here suggest how eyetracking can function within an interdisciplinary approach to film analysis that furthers our appreciation of film in previously unfathomed ways.

ACKNOWLEDGEMENTS

Thanks to Rachael Bedford, Sen Redmond and Craig Batty for comments on earlier drafts of this article. Thank you to John Henderson, Parag Mital and Robin Hill for help in gathering and visualising the eye movement data used in the Figures presented here. Their work was part of the DIEM Leverhulme Trust funded project (https://thediemproject.wordpress.com/). The author, Tim Smith is funded by EPSRC (EP/K012428/1), Leverhulme Trust (PLP-2013-028) and BIAL Foundation grant (224/12).

Bibliography

Anderson, Joseph. 1996. *The Reality of Illusion: An Ecological Approach to Cognitive Film Theory*. Southern. Illinois University Press.

Batty, Craig., Perkins, Claire., & Sita, Jodi. In Press. "How We Came To Eye Tracking Animation: A Cross-Disciplinary Approach to Researching the Moving Image", *Refractory*

Banks, M.S., Read, J.C.A., Allison, R.S., and Watt, S.J., 2012. "Stereoscopy and the human visual system." *SMPTE Mot. Imag. J.,* 121 (4), 24-43

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. 2008. "The pupil as a measure of emotional arousal and autonomic activation." *Psychophysiology*, *45*(4), 602-607.

Branigan, E., 1984. *Point of View in the Cinema: A Theory of Narration and Subjectivity in Classical Film.* Berlin: Mouton.

Brown, W. In Press. "There's no I in Eye-tracking: How Useful is Eye-tracking to Film Studies?", *Refractory*

Buchan, J. N., Pare´, M., & Munhall, K. G., 2007. "Spatial statistics of gaze fixations during dynamic face processing." *Social Neuroscience*, 2, 1–13.

Coutrot, A., Guyader, N., Ionescu, G. & Caplier, A., 2012. "Influence of Soundtrack on Eye Movements During Video Exploration"*, Journal of Eye Movement Research 5*, no. 4.2: 1-10.

Cutting, J. E., DeLong, J. E., & Nothelfer, C. E., 2010. "Attention and the evolution of Hollywood film." *Psychological Science*, 21, 440-447.

Dyer, T., In Press. "From Subtitles to SMS: Eye-Tracking, Texting and Sherlock", *Refractory*

Dyer, A. Pink, S. In Press. "Movement, attention and movies: the possibilities and limitations of eye-tracking?", *Refractory*

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

Henderson, J. M., 2003. "Human gaze control during real-world scene perception*." Trends in Cognitive Sciences,* 7, 498-504.

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.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. 2011. *Eye Tracking: A comprehensive guide to methods and measures*. Oxford, UK: OUP Press.

James, W. 1890. *The principles of psychology* (Vol.1). New York: Holt

Kruger, J. L., Szarkowska, A. & Krejtz, I. In Press. "Subtitles on the Moving Image: An Overview of Eye Tracking Studies", Refractory

Le Meur, O., & Baccino, T. 2013. "Methods for comparing scanpaths and saliency maps: strengths and weaknesses." *Behavior research methods*, 45(1), 251-266.

Magliano, J. P., & Zacks, J. M. 2011. "The Impact of Continuity Editing in Narrative Film on Event Segmentation." *Cognitive Science, 35*(8), 1-29.

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

Rayner, K. 1998. “Eye movements in reading and information processing: 20 years of research”. *Psychological Bulletin,* 124(3), 372-422.

Rayner, K., Smith, T.J., Malcolm, G.L. and Henderson, J.M. 2009. “Eye movements and visual encoding during scene perception.” *Psychological Science*, 20, 6-10.

Raz, G., Jacob, Y., Gonen, T., Winetraub, Y., Flash, T., Soreq, E., & Hendler, T. 2014. “Cry for her or cry with her: context-dependent dissociation of two modes of cinematic empathy reflected in network cohesion dynamics.” *Social cognitive and affective neuroscience*, 9(1), 30-38.

Redmond, S., Sita, J. Vincs, K., In Press. “Our Sherlockian Eyes: the Surveillance of Vision”, *Refractory*

Robinson, J., Stadler, J. & Rassell, In Press. “Sound and Sight: An Exploratory Look at Saving Private Ryan through the Eye-tracking Lens”, *Refractory*

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

Sawahata, Y., Khosla, R., Komine, K., Hiruma, N., Itou, T., Watanabe, S., et al., 2008. “Determining comprehension and quality of TV programs using eye-gaze tracking.” *Pattern Recognition*, 41(5), 1610-1626.

Smith, M., 2011. “Triangulating Aesthetic Experience”, paper presented at the annual *Society for Cognitive Studies of the Moving Image conference, Budapest, June 8–11*, 201

Smith, T. J., 2006. *An Attentional Theory of Continuity Editing.* Ph.D., University of Edinburgh, Edinburgh, UK.

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., 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. pages 165-191

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

Smith, T. J., Batten, J., & Bedford, R., 2014. “Implicit detection of asynchronous audiovisual speech by eye movements.” *Journal of Vision*,*14*(10), 440-440.

Smith, T. J., Dekker, T., Mital, P., Saez De Urabain, I. R. & Karmiloff-Smith, A., In Prep. “Watch like mother: Motion and faces make infant gaze indistinguishable from adult gaze during Tot TV.”

Smith, T.J. and Henderson, J.M., 2008. “Edit Blindness: The relationship between attention and global change blindness in dynamic scenes”. *Journal of Eye Movement Research*, 2(2):6, 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., Levin, D. T. & Cutting, J., 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.

Smith, T. J. & Martin-Portugues Santacreu, J. Y, Under Review. “Match-Action: The role of motion and audio in limiting awareness of global change blindness in film.”

Smith, T. J. & Smith, M., In Prep. “The impact of expertise on eye movements during film viewing.”

Suckfull, M., 2000. “Film Analysis and Psychophysiology Effects of Moments of Impact and Protagonists”. *Media Psychology*, *2*(3), 269-301.

Vilaro, A. & Smith, T.J., 2011. “Subtitle reading effects on visual and verbal information processing in films.” Published abstract In Perception. ECVP abstract supplement, 40. (p. 153*). European Conference on Visual Perception*. Toulousse, France.

Velichkovsky, B.M., Dornhoefer, S.M., Pannasch, S. & Unema, P.J.A., 2001. “Visual fixations and level of attentional processing”. In A. Duhowski (Ed.), *Proceedings of the International Conference Eye Tracking Research & Applications,* Palm Beach Gardens, FL, November 6-8, ACM Press.

Wass, S. V. & Smith, T. J., In Press. “Visual motherese? Signal-to-noise ratios in toddler-directed television,” *Developmental Science*

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C., 2001. “Illusory perceptions of space and time preserve cross-saccadic perceptual continuity”. *Nature*, 414.

Zakay, D., & Block, R., 1996. *Role of Attention in Time Estimation Processes. Time, Internal Clocks, and Movement* . Elsevier Science.

Figure Captions

Figure 1: Figure from Vilaró & Smith (2011) showing the gaze behaviour of multiple viewers directed to own language subtitles (A) and foreign language/uninterpretable subtitles (B).

Figure 2: Gaze Heatmaps of participants' free-viewing a trailer for Lego Indiana Jones computer game (left column) and the Video Republic documentary (right column). Notice how both make copious use of text within the image, as intertitles and as extra sources of information in the image (such as the head-up display in A3). Data and images were taken from the Dynamic Images and Eye Movement project (DIEM; Mital, Smith, Hill & Henderson, 2010). Videos can be found here (http://vimeo.com/6628451) and here (http://vimeo.com/2883321).

Figure 3: Gaze heatmap of eight participants watching a series of shots and digital inserts from Hulk (Ang Lee, 2003). Full heatmap video is available at http://youtu.be/tErdurgN8Yg.

Figures

Figure 1

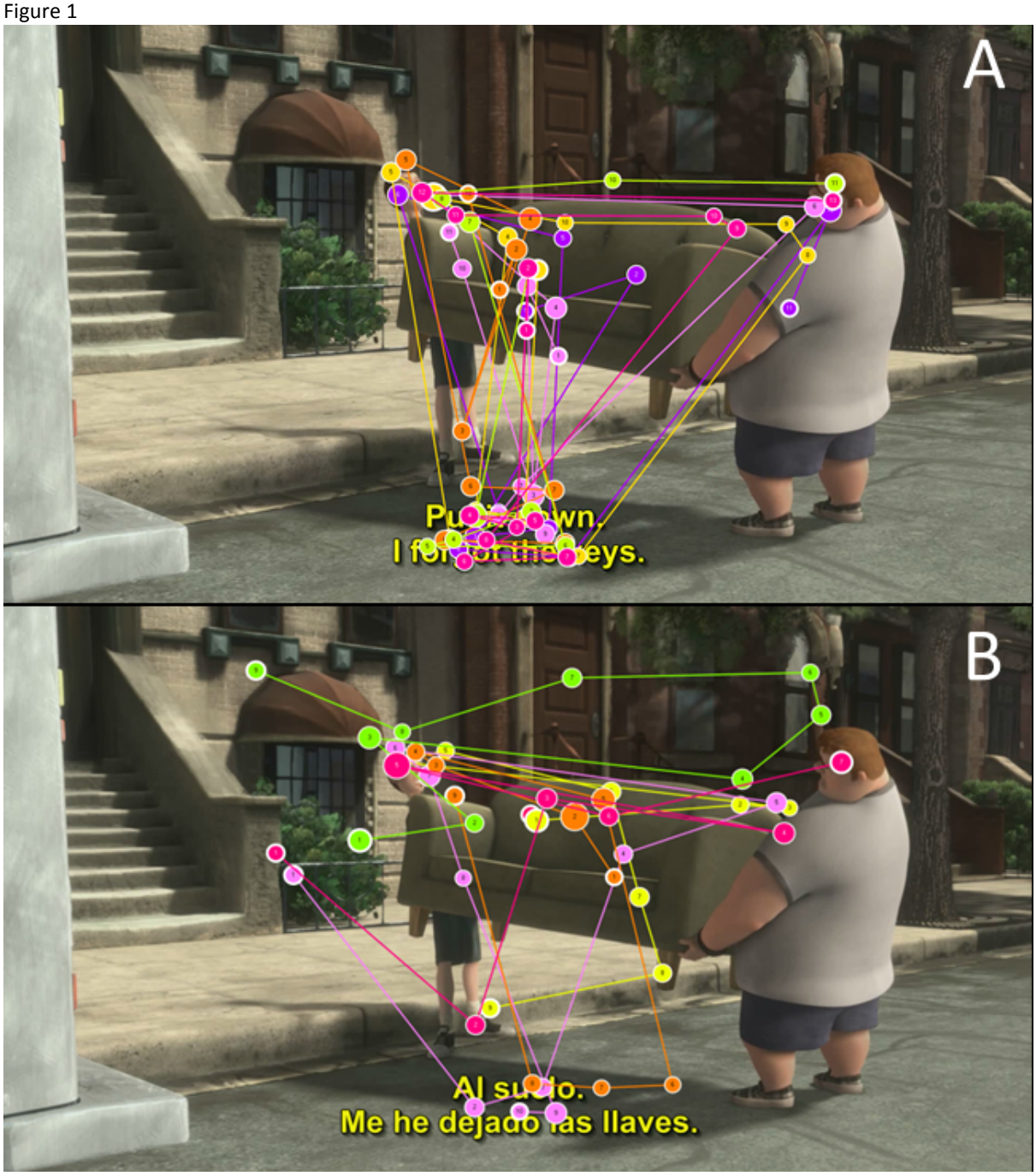

Figure 2

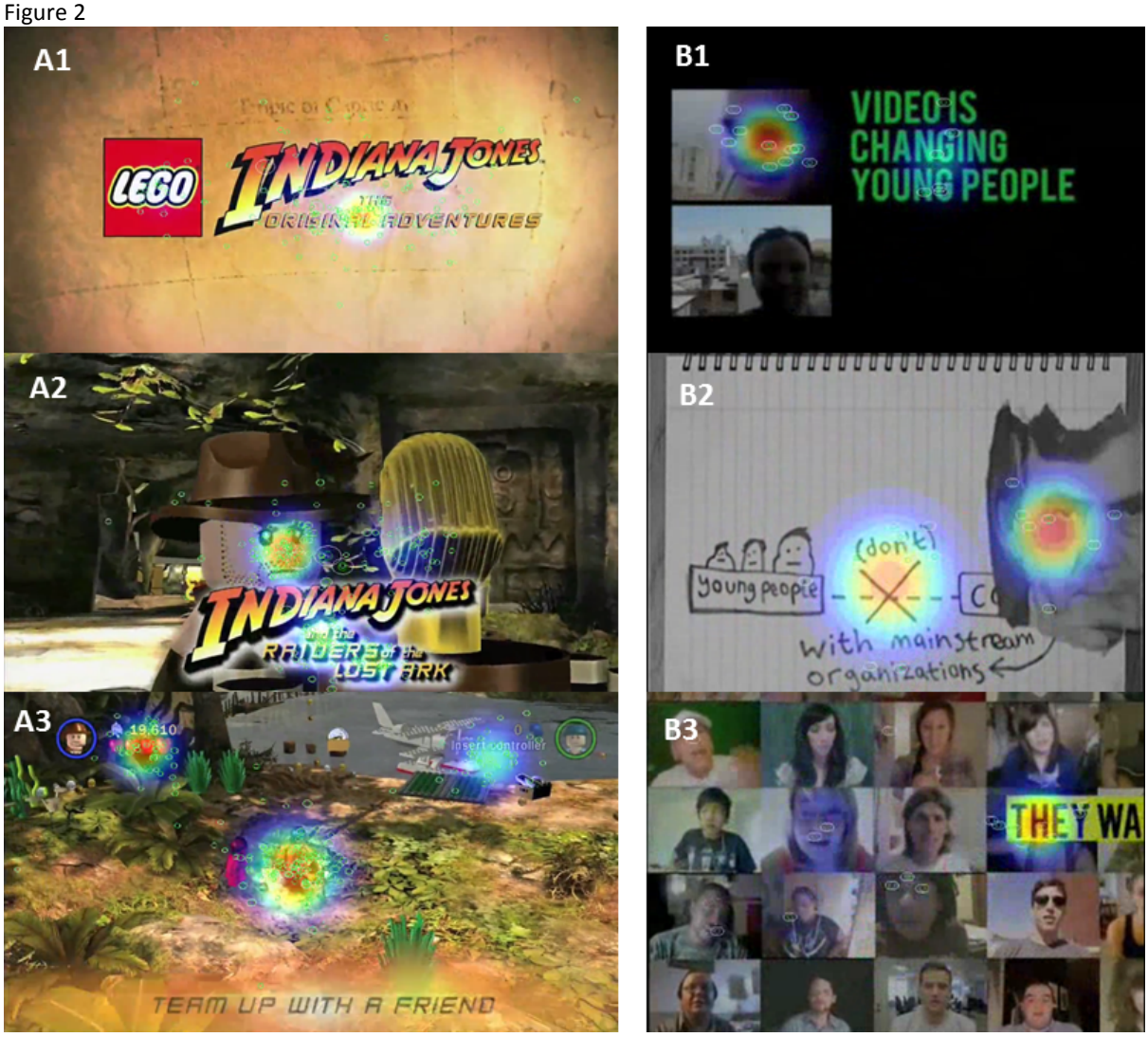
A1
LEGO
INDIANA JONES
THE ORIGINAL ADVENTURES
A2
INDIANA JONES
and the RAIDERS of the LOST ARK
A3
19,610
insert controller
TEAM UP WITH A FRIEND
B1
VIDEO IS
CHANGING
YOUNG PEOPLE
B2
young people
(don't)
with mainstream
organizations
B3
THEY WA

Figure 3

Endnotes

[i] Which is even more surprising considering how closely technological developments in film have been dependent on advances in other scientific disciplines including chemistry for the recording of light, computer graphics for special effects, visual optics, etc, and how film crew such as cinematographers, editors, sound designers, and even directors and script writers are often trained scientists.

[ii] An alternative take on eyetracking data is to divorce the data itself from psychological interpretation. Instead of viewing a gaze point as an index of where a viewer's overt attention is focussed and a record of the visual input most likely to be encoded into the viewer's long-term experience of the media, researchers can instead take a qualitative, or even aesthetic approach to the data. The gaze point becomes a trace of some aspect of the viewer's engagement with the film. The patterns of gaze, its movements across the screen and the coordination/disagreement between viewers can inform qualitative interpretation without recourse to visual cognition. Such an approach is evident in several of the articles in this special issue (including Redmond, Sita, and Vincs, this issue; Batty, Perkins, and Sita, this issue). This approach can be interesting and important for stimulating hypotheses about how such patterns of viewing have come about and may be a satisfying endpoint for some disciplinary approaches to film. However, if researchers are interested in testing these hypotheses further empirical manipulation of the factors that are believed to be important and statistical testing would be required. During such investigation current theories about what eye movements are and how they relate to cognition must also be respected.

[iii] Although, one promising area of research is the use of pupil diameter changes as an index of arousal (Bradley, Miccoli, Escrig, & Lang, 2008).

[iv] This technique has been used for decades by producers of TV advertisements and by some "pop" serials such as *Hollyoaks* in the UK (Thanks for Craig Batty for this observation).

[v] This trend in increasing pace and visual complexity of film is confirmed by statistical analyses of film corpora over time (Cutting, DeLong & Nothelfer, 2010) and has resulted in a backlash and increasing interest in "slow cinema".

[vi] Other authors in this special issue may argue that taking a critical approach to gaze heatmaps without recourse to psychology allows them to embed eyetracking within their existing theoretical framework (such as hermeneutics). However, I would warn that eyetracking data is simply a record of how a relatively arbitrary piece of machinery (the eyetracking hardware) and associated software decided to represent the centre of a viewer's gaze. There are numerous parameters that can be tweaked to massively alter how such gaze traces and heatmaps appear. Without understanding the psychology and the physiology of the human eye a researcher cannot know how to set these parameters, how much to trust the equipment they are using, or the data it is recording and as a consequence may over attribute interpretation to a representation that is not reliable.

[vii] https://theeyetribe.com/ (accessed 13/12/14). The EyeTribe tracker is $99 and is as spatially and temporally accurate (up to 60Hz sampling rate) as some science-grade trackers.

[viii] http://www.tobii.com/eye-experience/ (accessed 13/12/14). The Tobii EyeX tracker is $139, samples at 30Hz and is as spatially accurate as the EyeTribe although the EyeX does not give you as much access to the raw gaze data (e.g. pupil size and binocular gaze coordinates) as the EyeTribe.