Our Sherlockian Eyes: the Surveillance of Vision - Sean Redmond, Jodi Sita and Kim Vincs (http://refractory.unimelb.edu.au/2015/02/07/redmond-sita-vincs/)

12

Abstract

For this inter-disciplinary article, we undertook a pilot case study that eye-tracked the 'Holmes Saves Mrs. Hudson' sequence from the episode, A Scandal in Belgravia (Sherlock, BBC, 2012). This small-scale empirical study involved a total of 13 participants (3 males and 10 females, mean age was: 27 years), comprised of a mixture of academics and undergraduate students at La Trobe University in Melbourne, Australia. The article examines its findings through a range of threaded frames – neuroscience, forensics, surveillance, haptics, memory, performance-movement, and relationality – and uniquely draws upon the interests of the authors to set the examination in context. The article is both a reading of Sherlock and a dialogue between its authors. We discover that the codes and conventions of Sherlock have a direct impact on where viewers look but we also discover eyes emerging in the periphery of the frame, and we account for these ways of seeing in different ways.

My Sherlockian Eyes

Sean Redmond

I have always been fascinated, perhaps even obsessed, with my eyes. I have often *felt* them looking into things, as if they had their own embodied consciousness that I was entirely, simultaneously, conscious of. It was as if we, my eyes and I, saw the world separately and together, possessing a double vision, one set within the meaty windows of my sockets, and the other looking outside, grasping the world with a replete hapticity, sending shivers across my pupils and retinas as they did so.

I have found myself trying to catch my eyes out, to second guess their movements, their sightlines, and their interests. I must be a sight for sore eyes on the rush hour train, wrestling with what I will allow my eyes to see. I often try to resist my conforming eyes, to make them look towards the cultural periphery, to the aesthetic margins, and to the haphazard shards of broken, refracted light on oily windows that few others see as they go about their busy, and sometimes dreary lives. I have a deep yearning to see my eyes politicised, to turn them completely into organs of touch (Marks, 2000), and to feel them wander freely across the intricate layers of the film and television screen. I want Sherlockian eyes.

I have held a rather romantic notion about my viewing eyes, and the eyes of some viewers: that they sometimes wander freely across the spaces, objects, lights, colours, bodies, movements and sounds of the diegetic world they are presented with. Narrative action may be centre frame, and all the elements of the mise en scène may be attempting to draw one's eyes to this interaction, but I will catch myself looking to the far left of the screen, to hold my sight on an obscure pattern on a wall, or to search for the

origins of a distant minor or insignificant sound just off-screen. I want to see inside and outside the narrative simultaneously. I imagine my eyes as Sherlock-like, searching for narrative clues, new plot developments, and for the sensuous expression of character, mood and feeling. But I also see them loosing or freeing themselves; my eyes (unconsciously) float within all the elements of filmic or televisual material as they happen on the screen.

I see in Sherlock's eyes this double vision: the ability to have foresight, to see into the margins of things, and to be consciously aware of the vision within and all around him. As Sherlock sees into the finest grain of things, so do my eyes and I. My Sherlockian eyes are forensic, haptic, self-processing and are blessed with twenty twenty vision – they have the power to see into all things clearly. Sherlock mirrors or rather embodies the very qualities of the cinema machine (Metz, 1982), and of the surveillance regimes (Foucault, 1977), that emerged at the time the first *Sherlock Holmes* books were written (1887-1927). *Sherlock* is a text that already embodies the eye tracking experience.

But is this so, or just a fictive longing? What evidence do we have that our eyes do what we say they do? What evidence do we have that viewers possess a double vision? This romantic, phenomenological notion of the viewing, carnal, haptic eyes, then, we wanted to test, to explore, to see in action and interaction…

The Science of the Sherlockian Eye

Jodi Sita

I can often be found staring off into space, deep in thought, looking at nothing in particular. If I were being eye tracked it would look like I was staring at *something*. Where people look and, more particularly, why they are looking there, are questions that fascinate me and make me think about the phenomenon of blank stares. Human beings have fascinating eyes that, because they are housed independently, with their own localized environments, need to and can move about quite a lot. A tiny spot on the back of the eye, the retina, houses the receptors for high visual acuity, and this spot must be directed at the object we want to see, in order for us to see it clearly, and see its fine detail. People tend to move their eyes to aspects of a scene which are interesting or useful. The visual system in the brain directs these movements; they are not random. Bottom–up control processes (see; Itti & Koch, 2000) help direct some shifts in visual gaze and involve features that are thought to attract attention due to their ability to be noticeable. These include salient features such as luminosity, colour and movement. More importantly, in the human gaze we know that top-down processes are at play when viewing complex or meaningful scenes; our eyes employ feature selection which are based on our understanding of the scene and our internal expectations about where important things are or are likely to occur (Torralba et al., 2006; Birmingham et al., 2008, and Vincent et al., 2009).

What I am curious to learn more about is how our viewing behavior is shaped by what we are doing, by how we are interacting with the world and how our brains are responding to that and shaping that encounter. Thus, my involvement in this work comes from these curiosities, however, it also stems from my own forensic tendencies. It develops from my own need to ask the Sherlockian questions about viewers viewing *Sherlock*.

My early research had me investigate a branch of forensic science; handwriting and signature examination. At the time the area had a lot of practitioners, many quite experienced and successful, and there had been a substantial amount written about it, yet very little objective evidence for the fields' claims had been produced. What the field needed were studies which produced *hard evidence* to support or dispute its original claims. My work was part of a large and ongoing body of work, where the field's ideas and claims are tested objectively, and whose results can be used as evidence to support existing notions or derive new ones. It was within this area of research that I started using eye tracking, yet it has also led me to want to bring eye tracking to this moving-image field; where the focus of the eye tracker, shining like an objective lens over some of the theories of the area, can help bring to it some other method for its practitioners to use – to examine how viewers watch and are *involved* with what they are watching.

The Optometry of *Sherlock*

Kim Vincs

The science of the gaze—of how eyes fixate or fail to fixate—has always been of great interest to me, firstly in my original career as an optometrist, and more latterly as a choreographer and then a transmedia dance artist. As an optometrist, I was less concerned with where people looked than with whether they could look, and with the accuracy and resolution of the sensory information they received and interpreted when they did look. What do I mean by this? In considering whether people could look, I am referring **to** whether they were able to accurately fixate the static and moving targets they wanted and needed to. Fixating, that is, aiming one's eyes at a static or moving target, is a function of attention, and integrated as action by the sensory and muscular systems of the eye and brain. There are many pathological conditions that interfere with the capacity to fixate a static visual stimulus quickly, accurately and efficiently. As an optometrist, I was primarily concerned with detecting these conditions and referring patients who had them for appropriate treatment. I was, in a very real sense, perfectly happy to allow my patients to decide for themselves what to fixate on. My job was simply to ensure that, should they wish to, they would be capable of locating and tracking something. This willingness to allow dissociation between capacity and will, between ability and decision, is something I consider foundational to the ways in which I have pursued my subsequent research into creative practices. I have never, as a choreographer or an interactive / transmedia artist, wished to dictate to people where they should look or what they should perceive. I consider my job to be to place appropriate objects / events / movements within a context in which they can be perceived should people so choose.

This outlook has had some specific implications for my art practice. As a choreographer, I have never thought to ask what someone watches when they observe a dancer moving. Cognitive psychologist Kate Stevens' seminal work on eye movements in dance has demonstrated a classic novice/expert shift in the way that observers view dance. As with many other fields of expertise such as airline pilots, and driving instructors, experts make significantly fewer saccades, that is, changes in fixation, watching a dance performance than do novices, where experts are people with professional experience in dance and novices are people with no particular prior experience of the artform (Stevens et al., 2010). The implication of these results is that experts do not need to change fixation as many times as novices because they are able, to some extent, to predict where the dancing body will move. In essence, they understand what they are looking for, and are therefore able to maximize the efficiency of their fixation choices.

What Steven's work does tell us is which movement features most attract fixation when watching a dancing body. My own work in motion capture analysis of dance movement provides me with a theory about why this might be a difficult thing to measure. Dance, at the movement level, comprises movement of some 33 major joints, each of which may make movements of entirely different velocity, acceleration and magnitude to achieve an overall aesthetic effect. The dancing body essentially has no 'centre of focus' that can be interpolated from movement data such as the speed, momentum or even position of specific body parts, because the semantics of dance movement are only meaningful in relation to the composition across the body. As I have argued previously, (Vincs, 2014) the semantic significance of a movement bears no relationship to its metrics, such as amplitude or speed. In some aesthetic contexts, tiny movements of the fingers may be essential to the meaning and feeling tone of the movement. In others, such as large virtuosic or acrobatic movement forms, hand gestures may contribute relatively little to a movement's significance.

Dance grammars are aesthetically and culturally, rather than anatomically determined. I think that this fact has also contributed to my attraction to the notion of Sherlockian eyes. As a choreographer, I am always a detective, seeking potential significances in movement rather than predetermined ones. I value the opportunity to go looking for the dancing body, browsing, shuffling, wandering through the multiple and complex joint actions that comprise a single 'step,' looking for something of newness and emotional value rather than assuming I know what it is and where I will find it. Yet I am always aware that my aesthetic search is underpinned by a neurosensory apparatus that is primed to respond selectively to human movement (Hagendoorn, 2004, Vincs, 2009). I am therefore armed, at least potentially, with an inherent 'grammar' that is defined by the morphology and physical capacity of the human body, and I am curious as to what predilections and biases my visual sensory system imposes on my seemingly adventurous gaze.

What now follows is an exploration of our different approaches to the eye tracking data that we generated. Jodi is first off, outlining our empirical method and undertaking a close reading of the preliminary results. Jodi shows how the results begin to tell us that the viewers' gaze patterns and fixations are closely clustered together, and she situates these findings in relation to the science of the eye, the importance of the face in human communication, and to the visual and narrative codes and conventions of *Sherlock*. Sean then explores the results in terms of haptic visuality and the surveillance gaze, drawing upon phenomenology and the discourses of conspiracy to argue that the vision in *Sherlock* is marked by touch, texture, and control. Kim examines the results in terms of movement and relationality, examining the eye tracking data in terms of the way it supports and confirms the necessary nature of vision in seeing into moving things. Kim shows that even though there is a high degree of direction in terms of where viewers are being asked to look, visual perception allows or enable the eyes to wander. Finally, we conclude our article together, drawing together our voices to offer an interdisciplinary way forward.

Eye Tracking *Sherlock* (the objective viewing): Methods and Preliminary Results

Jodi Sita
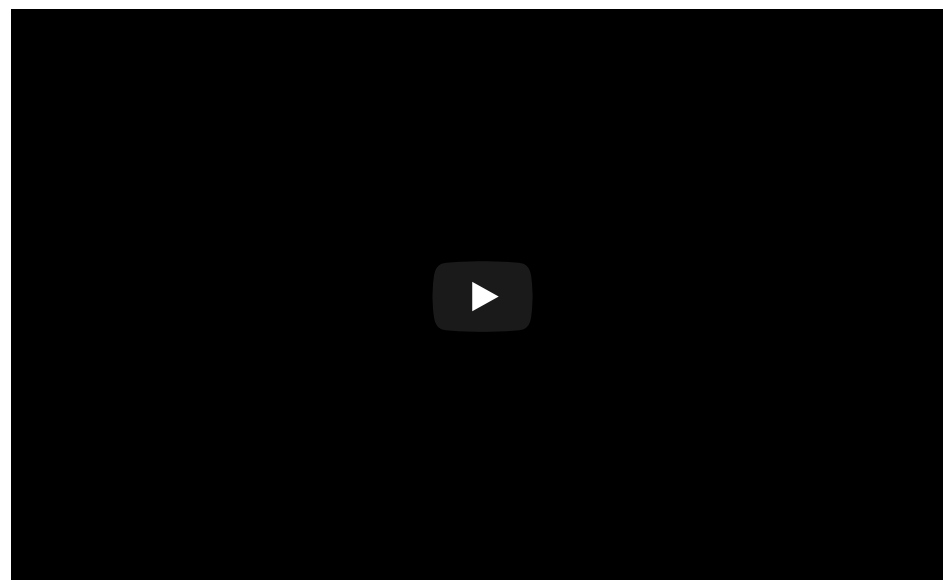
We undertook a pilot case study that eye-tracked the 'Holmes Saves Mrs. Hudson' sequence from the episode, A Scandal in Belgravia (*Sherlock*, BBC, 2012). This small-scale empirical study involved a total of 13 participants (3 males and 10 females, mean age was: 27 years), comprised of a mixture of academics and undergraduate students at La Trobe University in Melbourne, Australia.

A Tobii X-120 remote eye tracker (Tobii Technology, Stockholm, Sweden) was used to record participants eye movements which has an accuracy of 0.5º of visual angle and allows a moderate amount of free head movement (30 x 22 x 30 cm at 70 cm (Width x Height x Depth)). This data collection technique uses reflected infrared light from the eye to determine participants viewer gaze positioning and allows for natural head movements and natural human responses to screened material. The eye tracker was connected to a PC running an Intel ® core ™ i7 CPU 'Cool Master' hard-drive. The eye tracker used Tobii Studio 2.3.2 professional edition software for the presentation of the movie scene stimuli and recording eye movements. The eye tracker was set up on a desk, situated below a Dell PC monitor (1680×1050), which was utilised by the participants to view the *Sherlock* sequence. Participants were seated on a sturdy chair between 55-65cm away from the eye tracker and between 65-75cm from the viewing screen. A second screen (Dell; 1920×1080) was utilised by the researchers to view, in real time, the eye movements of the participants as they were being tracked and calibrated, although all computer analyses and statistics reported here was based on stored data.

Participants were recruited via posters advertising the study at La Trobe University, with ethics approval (Ethics approval number: FHEC13/101). Participants were required to be at least 18 years of age to be considered eligible. People whom expressed interest in taking part in the study were contacted via email to attend a single recording session. In preliminary tests participants were introduced to the study and screened for exclusion criteria such as taking medications (e.g. benzodiazepines) that may potentially affect their eye movement, known neurological conditions, disorders or injuries that could potentially

affect their eye movement. All participants were screened for normal or corrected to normal near visual acuity of N8 or better on the Designs for Vision near sighted visual acuity test, and with a pen-torch eye movement excursion test to screen for symmetrical movement of the eyes. Participants who were ametropic were allowed to wear their glasses to watch the stimuli.

Prior to eye movement data being collected, the eye tracker was configured for each participant using a 9-point on screen calibration test within the Tobii Studio recording software. Participants were told only that they would watch short segments from a variety of films. Recording sessions typically lasted between 15-25 minutes, and each participant was tested individually.

First, we found that our viewers' eyes were strongly drawn to follow movement and directional cues and signs. This included camera and character movement. In the opening scene, where Mrs. Hudson's fingers scrape along the wall, followed by Sherlock's fingers retracing her steps (03-010 seconds), we see all viewers making strings of successive fixations – each following these finger movements (see Figures 1 and 2). The sound of these fingers scraping along the wall was heavily amplified, and fully sychronised, and we suggest, then, that sound was also an aesthetic device being employed to direct where viewer's looked. These results confirm previous findings where camera movement, sound, character behavior, and editing patterns are seen to inform gaze patterns and fixations (see Smith, forthcoming, Smith and Mital, 2011).

In one brief shot in the middle of this scene, we cut to a close-up of Mrs. Hudson's face, full of anguish. All the subjects discern this face in amongst the movement and chaos of the surrounding action, as seen by their fixating to its features. Her face is captured in the center of the screen, making it central to the scene's visualisation. However, the face is known to be a strong attractor of what the eye attends to (Treuting (2006)), and with it being such an important narrative component in this scene, would have been a strong attentional cue.

(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Image-1-Finger-drag-2-subjects.jpg)
Figure 1: Finger drag: 2 subjects

(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Image-2-Finger-drag-13-subjects.jpg)
Figure 2: Finger drag: 13 subjects

Second, we observed an alignment in vision with regards to where Sherlock was looking. This sight co-proximation is referred to as 'joint attention,' in which what one attends to seems to shift automatically to where another is looking (Birmingham et al, 2009). Interestingly, this is a common misdirection trick used by magicians (Kuhn, et al, 2009).

In particular, we observed that Sherlock's' point of view in the scene very often produced a close proximity in viewers focus and attention (participant's looked where Sherlock looked, and with the same overall gaze patterning, see from 1.05 to 1.14). This also supports the findings reported by other film scholars using eye tracking methods, such as Rassell et al. (forthcoming, 2015) who found that a character's point of view and subjective experiences have an influence on where viewers look.

The trends for this short sequence support the idea that *Sherlock* is a character-driven drama in which his vision is not only foregrounded but given omnipotent and omniscient power. Thus, viewers are not only being positioned to observe from his authorial position but to trust where he looks and what he discovers there. There are recognizable genre codes and conventions also in play, structuring the looking patterns we have observed. This is a detective-thriller series that repeats a series of camera and editing motifs that become familiar to audiences (Neale, 1990).

(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Sherlock-Image-3.jpg)
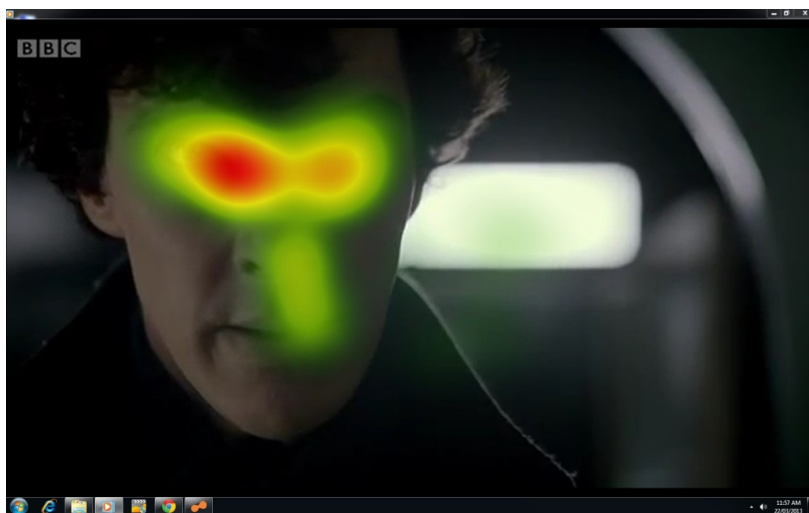
Figure 3: A heap map showing the hot spots where viewer's gazed; red indicating longer dwelling time

(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Sherlock-Image-4.jpg)
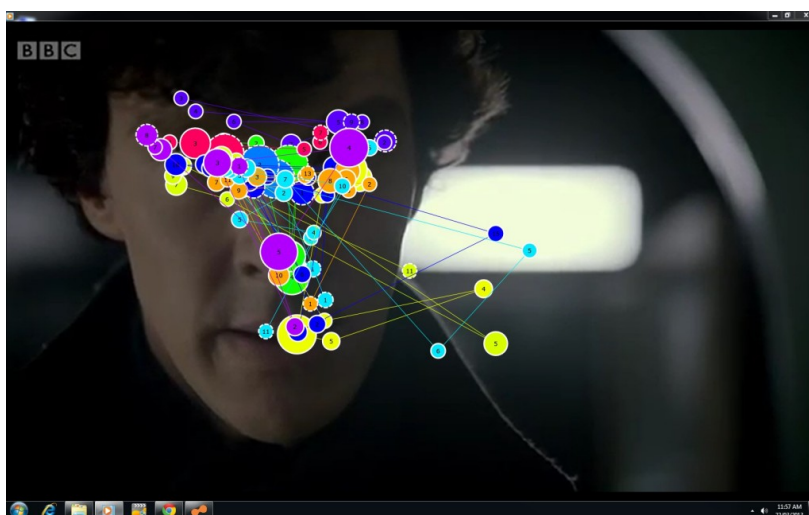
Figure 4: The gaze plots showing the sequence of looks that viewers made over Sherlock's face

Third, we found that viewers focused heavily on the characters faces, both in scenes with dialogue and those without. In scenes where Sherlock was clearly putting together the evidence, viewers focused heavily on his eyes, dwelling there for almost the entire shot. (Figures 3 and 4).

Viewers fixated back and forth between the eyes, face, and mouth of the central characters. These viewing patterns are characteristic of the movements made in facial and emotional recognition (Ekman & Friesen, 1971; Hernandez, et al, 2009) and show some indication that viewers were paying attention to the different character's in the scene, working out the role of each character and what their intentions and emotions

might be. These patterns of eye movements suggest that viewers are engaging with the scene as they would in a normal face-to-face encounter, using eye movements to verify who people are and what they are feeling. It is interesting to note that people who are not able to perform these socially informative tasks, such as those with the disorder schizophrenia, and with some traumatic brain injuries, do not show the same eye movement behaviors (Watt & Douglas, 2006; Loughland et al, 2002; Williams,et al, 1999).

Our viewers clearly followed narrative cues in line with the dialogue exchanges, looking back and forth between the character's interpersonal relays (Figures 5, 6 and 7). These results are similar to those of Treuting (2006), who eye-tracked 14 participants viewing short clips from such films as *Shawshank Redemption* (1994) and *Harry Potter and the Philosopher's Stone* (2001). Treuting found that gaze clusters emerged in and around the central character's faces involved in dialogue and moments of heightened drama (see also, Redmond, Sita, 2013).
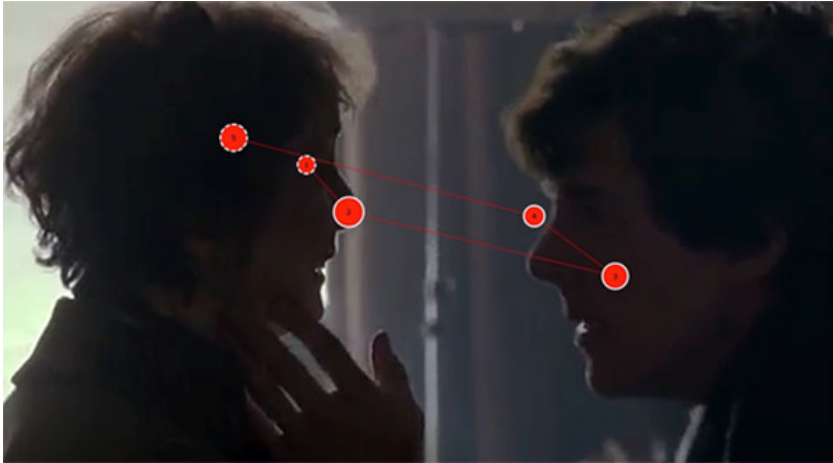
(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Sherlock-5.-Single-viewer-charecter-alighnment.jpg)
Figure 5: Single viewer character alignment

(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Image-6-6-subjects-and-the-relay-of-looks-on-eyes-mouths-and-faces.jpg)
Figure 6: 6 subjects and the relay of looks on eyes, mouths and faces.

(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Image-7-Final-scene-last-12-seconds-searching-for-information.jpg)

Figure 7: Final scene, last 12 seconds, searching for information: 13 subjects (most of the fixations are falling over the faces of the 2 central character's in dialogue)

Fourth, we saw evidence that viewers searched for narrative information and cues: this included fixating on aspects of the background wall before Sherlock first enters the scene (from 0.33-0.36 seconds), then moving between the image of a smile seen on the wall and Holmes' face, spending time 'reading' the shop window signs and the note on the front door (Figure 8) as Watson arrives at the scene (from 2.22 to 2.37). One can understand such scanning as influenced by the meticulous work of the mise en scène: where all the elements have been carefully placed to enact this type of searching for narrative cues (see Smith, forthcoming).

(http://refractory.unimelb.edu.au/wp-content/uploads/2015/02/Clip-8.jpg)
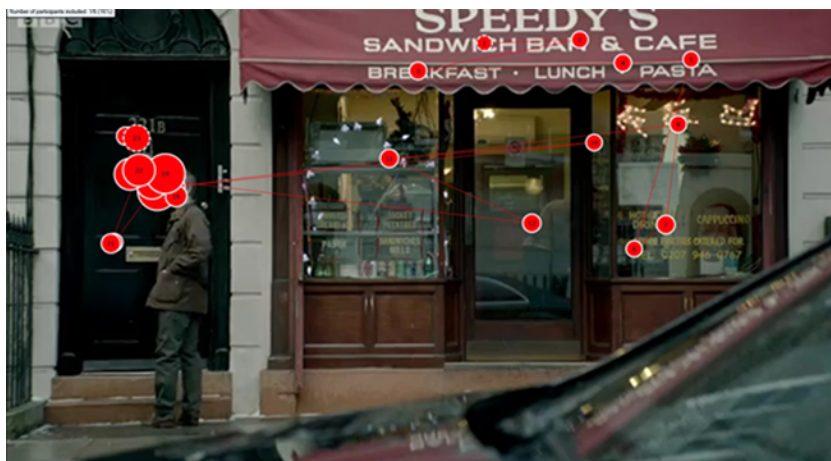
Figure 8: Searching for narrative information

Finally, albeit in relation to our last point, we observed that certain viewers looked at more elements of the mise en scène (Figure 9, shows gaze patterns for 4 of the 13 viewers), including the interior lights, the computer, and furniture, even as the more dramatic moments of the scene were taking place.

These findings were interesting but not totally unexpected; we would hope that not all people viewing the same scene would watch it in the same way (this is something that is discussed further below). Insights like this allow us to see that even though there are some aspects that are strong attention grabbers, such as faces and movement within a scene, other aspects can captivate and draw attention away from those areas of interest. For example, the scene shown below (Figure 9) involves a particularly emotive exchange between two key characters, Mrs. Hudson and Dr. Watson. The fact that 4 of the viewers were attending elsewhere helps us to see these aspects of interest outside the main narrative at play. Why certain viewers look to the margins of the screen, to the more 'insignificant' elements of the mise en scène remains of great interest. One possibility would be that these viewers were not fully engaged with the exchange between the characters, and their attention therefore drifted to other elements in the scene. Another possibility is that these scenic elements drew particular interest because of their pattern, colour, etc. Further testing is needed to begin to tease out whether this response is scene-dependent, or a characteristic of these particular observers.

Figure 9: A slightly different patterning – 4 subjects – and wider viewing

It should be noted, nonetheless, that these observations come from only a very small sample (13 participants to date), which will be increased, and which still needs to undergo further data analysis and interpretation.

In summary, what have we see in these results so far? Evidence of the eyes being held to attention by narrative cues, by camera and character movement, faces, dialogue, point of view and performance. These were elements to be expected and add to the growing body of eye tracking evidence that supports much of current film and television theory, particularly those working in the cognitive tradition such as David Bordwell (2007) and Noël Carroll (1996). The results equally support the results of other studies into; narrative centered visual texts (see Batty et al, forthcoming, 2015); and how sound and movement affect gaze patterns (Rassell et al, forthcoming, 2015) Further, they speak to the way viewers are pulled seamlessly into the diegetic worlds they believe and invest in.

In this article we would now like to apply two different theoretical filters to the results just summarized; the first will be an examination of the gaze, by Sean, and the second will be an examination of the physiological and perceptual processes of the eye in relation to movement, by Kim. Both filters will attempt to make deeper sense of the results from the traditions in which the scholars operate from. Following this analysis, a summary conclusion will draw their approaches together to make further inter-disciplinary sense of Sherlock's eyes.

Sherlock's Gaze

Sean Redmond

The concept of the gaze has a long and contentious history in Film Studies if much less so in the study of television. In fact, John Ellis has suggested that the domestic context in which television viewing has historically taken place, with a host of likely distractions, and in a context of constant programme flow and segmentation, produces a glance aesthetic whereby the image isn't looked into deeply or for a sustained period of time (1982: 138). *Sherlock*, of course, contests this idea since the programme is heavily built around the details of forensic gazing.

Most notably, the idea of the gaze has been employed in psychoanalytical film theory to make the argument that the cinema looking apparatus is patriarchal and heterosexual, and viewers are positioned as 'male' subjects through which masculine identifications emerge (Mulvey, 1975, 1989). Its main male characters, and male writers and directors of course control the vision regime in *Sherlock*, although its objects of focus are rarely to-be-looked-at female characters.

Critical race theory, by contrast, has employed the concept of the gaze to demonstrate how the racial Other is fixed in inferior, marginal and fetishized subject positions (Hall, 2001). *Sherlock* can be read as a post-colonial text enacting a present England that centres whiteness and 'invisibly' marginalizes the Other from its panopticon empowered centre (Cuningham, 2004). The Other in *Sherlock* of course extends to those who sit outside the bourgeois social centre; there are particular class dimensions to the way crime is surveyed and defined (Jann, 1990).

In terms of surveillance discourse, film has been read as a vision machine set within a invasive visual culture that promotes:

> the normalizing gaze, a surveillance that makes it possible to qualify, to classify and to punish. It establishes over individuals a visibility through which one differentiates and judges them (Foucault, 1977: 25).

Sherlock can be read as a text that carries out this normalizing gaze, defining the parameters of law and order and the way the criminal can be discovered, classified, and ultimately disciplined. That is not to say that the visual excesses of the programme do not at times undermine its simple binaries. To the contrary, *Sherlock* is constantly troubled by its own dominant discourses particularly through the way Sherlock is also a maverick outsider.

Finally, film phenomenology has made use of the gaze to demonstrate how looking and seeing is always embodied, experiential, and depending on the text, haptic and synesthetic – where 'the eyes themselves function like organs of touch' (Marks, 2010, 162). *Sherlock* creates the conditions of both embodied presence and haptic visuality through the way the gaze is employed to see deeply into things, while the programmes textural mise en scène 'demands' to be attended to.

What I would now like to do is analyse two particular aspects of the way the gaze can be understood in *Sherlock*, relating my reading back to the eye tracking results that we have, and to eye tracking technology itself. First, I will explore *Sherlock* through its forensic gazing and the way this creates the particular conditions for the way viewers become locked into particular viewing patters and relations. Second, I will explore *Sherlock* through its haptic elements whereby the viewer is understood to gaze at and touch (things) simultaneously.

The Forensic Gaze

In *Sherlock* one can argue that camera movement and position **are** motivated by the following factors. First, to reveal narrative information such as a new location, or setting; character relations and their relative physical proximity; time, and temporal detail; and moments of revelation where a new angle or focus reveals something previously hidden or a new 'enigma' emerges. Second, as a dramatic device: the camera is re-positioned to signal and cue moments of narrative development, crisis, reaction, and activation. Third, there are repeated and recognised televisual conventions of the programme: one can locate and expect certain camera motifs to function in *Sherlock*, such as the way we enter Sherlock's mind's eye to see what he is unearthing in microscopic close-up. Finally, camera movement and position signals certain emotional states and modes of feeling. The cut to a close-up, for example, a moment of affecting intensity, such as is the case with the fingers being scraped along the wall in the scene under analysis in this article.

When one takes these Sherlockian codes and conventions into consideration one can make better sense of the eye tracking results that we have gotten. The eyes of the viewer seem to be relentlessly led and directed. Viewers familiar with the programme's codes can be expected to have expectations of its visual tapestry, and to make predictions about where to look (see Rassell et al. forthcoming). This would explain both the way that viewers seem closely aligned with the looking operations of the scene (figures 1-7), and the way that viewers scan shots for narrative information (figures 8-10).

However, I also think there is something more telling to discover here – one around a consistent forensic looking regime where the text and the viewer align. This is the 'double vision' we refer to in my introduction to this article. Viewers come to embody the gazing powers that Sherlock possess and look at the diegetic world through his eyes even where no direct or imagined point of view is in operation. Viewers experience their very own form of social surveillance becoming detectives and snoopers in the process. *Sherlock*, then, can be read as a text of and for paranoid surveillance, fuelled by the constant search for facts, omissions, falsehoods, and half-truths. At a more general cultural level, trust is at issue here in what is perceived to be an age of 'faithless' activity and widespread corruption, where politicians are regarded to be as corrupt as the criminals they covertly support. Sherlockians ultimately become part of this age of conspiracy (Knight, 2002).

As does, in a very real sense, eye tracking technology and the data it produces. Sherlock is his very own eye tracker – he creates his own heat maps and relays and through this inbuilt biotechnology he sees into everything. Eye tracking technology *is* Sherlocklian and the data it produces allows us to see into everything the viewer sees. Or, at least mostly…

The Haptic Gaze

Laura U Marks (2010) has written that haptic visuality is a more intimate form of looking, where the eyes, 'move over the surface of its object rather than plunge into illusionist depth, not to distinguish form so much as to discern texture" (162). For Marks, film and video may be "thought of as impressionable and conductive, like skin." (2000: xi-xii) and this sensory materiality is heightened by it containing:

> Grainy, unclear images; sensuous imagery that evokes memory of the senses (i.e. water, nature); the depiction of characters in acute states of sensory activity (smelling, sniffing, tasting, etc.); close-to-the-body camera positions and panning across the surface of objects; changes in focus, under- and overexposure, decaying film and video imagery; optical printing; scratching on the emulsion; densely textured, effects and formats such as Pixelvision… and alternating between film/video. (Totaro, 2002)

The gaze found in *Sherlock* is very often a haptic one. The programme's entire mise en scène evokes the activity and memory of sensation. Lights, objects, clothes, furniture, exteriors are given deep and layered textures. Sherlockian environments are populated with objects and qualities that are themselves sensory driven (poison, oil, tactile fabrics, beads of sweat, cigarette smoke, wet soil). The camera very often dwells on these, picks them out, and tracks and pans over them in close and proximate detail. Sherlock of course is a master of haptic visuality – his eyes touches the things that he observes or that he conjures up in his imagination. In many respects, then, the viewer is also invited to see *Sherlock* through a haptic lens.

If one was to return to the eye tracking results on the scene what we might be observing is not just an alignment in vision, and the search for narrative information, but eyes that have been turned into organs of touch and deep sensual appreciation. For example, in image 1 and 2 viewers are not just following the fingers that scrape along the wall but touching (with) them, and in touching them feeling them as if it is their fingers suffering this pain. In figures 5-7 viewers are not just following the relay of looks between the two characters but 'touching' faces, eyes and mouths. In figures 8-10, viewers are not just searching for narrative information and clues but are actively seeing into the textures, lights, objects, items that populate those scenes. The heat maps that eye tracking technology can generate may be more apt than we imagine since the suggestion of temperature, of body-heat, may well give truth to the embodied and carnal nature of vision. This is one of the limitations of eye tracking technology, however, since it cannot tell us what people are *feeling* when watching a film or television text.

I would like to make one final observation about the eye that searches the mise en scène for narrative information or clues, as in figures 8-10. This is a point about the privacy or individualism of watching a screen text so co-dependent it is on personal memory, biography, and the contexts one finds one self-viewing something. After the viewing of the scene one of our subjects remarked that they had actually spent much of the time trying to figure out whom the actor was playing Dr Watson. Any number of 'personal' factors might get in the way of the looking regime of the text and for why we might scan a particular text.

Roland Barthes (1981) has usefully employed the concept of the punctum (a Latin word derived from the Greek word for trauma) to viewing photographs. He argues that the still image inspires an intensely private meaning, one in which an affecting 'partial object' emerges from its centre to 'prick' or 'wound' the viewer. The punctum is personal and as soon as it emerges it holds the viewer's gaze. Although Barthes is singularly writing about the photograph I think the idea of the punctum can be applied to the moving image text, to *Sherlock*. Although a dynamic media, television and film still settle on images and representations that reach out into the private realm of the viewer; and the viewer still finds their memories, traumas, life events activated and mobilised in the fictive worlds constructed. The wandering eyes of figures 8-10 are caught in their own biographical exploration, looking for objects that may ultimately wound them, their haptic eyes. In *Sherlock* we are not just being positioned as objects and subjects of surveillance but as carnal beings.

The Eyes of Sherlock; the Eyes of the Viewer

Kim Vincs

Coming, as I do, from the perspective of a dancer and choreographer, I read these eye tracking results not exactly as touch, but as a search for relationality. Erin Manning, in her seminal work on the philosophy of movement (Manning, 2009) emphasizes the incipiency of movement—movement as something in the act of becoming something, of reaching towards something or someone—over its positionality. That is to say, movement, in Manning's terms, is always a process of relationality or reaching towards the world, rather than simply a series of coordinates on a grid. Bodies, or, more precisely "bodies-in-the-making," are a means of thought rather than simply of action because they define, by the possibilities embedded in the moment of pre-articulation, a relationship or set of relationships within the world (Manning, 2009: 78).

Our eye movement data defines very clearly the relationships between the protagonists. In figures 1 and 2, 'finger drag,' the fixations map the pathway of Mrs. Hudson's fingers along the wall. The relationships between Mrs. Hudson's fingers and the texture of the wall are, in fact, the only potential human relationships within the scene. These relationships are not 'positional'— that is to say, fingers defining points on the wall—but, by virtue of their spatial distribution, an articulation of the trajectory of Mrs. Hudson's hand in relation to the wall.

Figures 3—6 reveal fixation patterns that are concentrated around the face, and in particular the eyes. These fixations address the origin of relationality in the scenes within the eyes, as if the eyes are understood to reveal the incipient thought of the characters. In Manning's terms, incipiency the moment in which a movement is in a state of 'pre-acceleration,' organizing and mobilizing itself, yet still capable of any number of actual outcomes, is the most potent aspect of a movement. The predominance of fixations on the eyes and face, while perfectly interpretable in as simply a biological reflex designed to respond to and recognize human faces, also speaks of the process by which relationality is thought into being.

Figures 7—10, with their fixations distributed between human-to-human gaze and gesture, (Figure 7) non-human scenic elements that lend cognitive elements or 'clues,' (Figure 8) and poetic elements such as the look and texture of surrounding objects in Figures 8 (curtain texture) and 10 (lens flare), articulate an expanded notion of relationality in which human and non-human elements are implicated within a web of actions.

For me, these results suggest a mutability between human and non-human elements that is reminiscent of Manning's understanding of relationality as something that can have both interpersonal and person-world dimensions, and also points to the kind of 'Sherlockian body' I seek as a dance artist. A purely 'narrative' approach to the filmic bodies analysed here would suggest that only human factors would feature in the gaze analysis. Similarly, a biologically driven notion of human movement perception as pre-wired to detect human shapes and actions would not predict migrations to and from the bodies to surrounding objects.

I read these results as revealing a semantic ambiguity that can form the basis for a Sherlockian search for relationality that is produced and constructed by the viewer as much as it is dictated by the film-maker. These scenes offer relatively few visual details from which to construct such a relationality, and this, no doubt, is indicative of a film-making style designed to direct the eye, and hence the mind, to very specifically and deliberately arranged narrative events. However, despite the seeming prescriptiveness of these images, they offer the viewer an opportunity, to construct relational scenarios across like and non-like (human and non-human) scenic elements. They therefore demonstrate at least the potential for a Sherlockian approach to the visual perception of movement.

Conclusion

So, what have we seen with your eyes? On the one hand, we have demonstrated how *Sherlock's* narrative and mise en scène pulls the viewer into taking certain (emotive, forensic, relational) viewing positions. *Sherlock* is tightly bound by a number of codes and conventions and its palette and composition are highly constructed, creating spaces and interactions that focus our eyes. On the other hand, we have looked at the way haptics and relationality open up the possibility for better understanding the synaesthetic and organic/inorganic ways through which the vision of movement takes place. A televisual text such as *Sherlock* intends to marshal our viewing experience but as we have also seen, eyes wander; they search, their own movement and the poetics of movement opening up textual encounters not always pre-determined by the scenes deliberate

operations. *Our eyes escape themselves*. Finally, we have noted how memory and biography can impact on where and why we look, or why we might look away. One of our respondents searched Watson's face in the hope of remembering the actor's name. The brutal violence metered out on Mrs. Holmes may be looked at (*felt*) differently by someone who has themselves undergone such misery. This is ultimately about our-being-in-the-world, and about how we make sense of beings-in-the-fictive world. Vision is never disembodied but full of the drink, food and love of life itself.

Vision needs to be understood as something that *can only be fully understood* through combining different theoretical and methodological frameworks. What we have found in this article, and in the broader work of the *Eye Tracking the Moving Image Research* group – see the introduction to this special edition for more on this group – is that it is in the conversations and deliberations, provocations and discussions between vision scientists, neuroscientists, anatomists, choreographers, film makers, ethnographers, screen theorists, and screen writers where insights are best made, conclusions thickened, and arguments enriched and extended. What we have found in this article is that *Sherlock* exists at the eye of the ArtsScience nexus, and it is at this nexus where the authors would like to situate their work.

The language of film and television constantly creates these spaces of vision and for *seeing* to take place, whether this be; the embodied point-of-view shot which allows us to become the character; aerial cinematography that brings wide open exteriors and bejeweled cityscapes into view; the furtive camera that glimpses into dark corners, allowing us to happen on to what is supposedly hidden; or the interiorized gaze that expressionistically captures the nightmare visions of the lost, the hunted, and the alien. With a close-up shot one can trace the undulating valleys of emotion on a character's face and feel their affecting eyes reaching out and into yours. In *Sherlock,* the very force of his vision, mobilized through special effects and the power of digital photography, enables his/our eyes to create mathematical formulations out of thin air and to re-visit crime scenes as if one is witnessing, experiencing it all over again. *Sherlock* establishes the sense that all vision is embodied and personally engineered, and to the wider conceit that television is an artform that gives a miraculous, omnipotent and omnipresent vision to its ever watchful viewers.

References

Barthes, Roland. 1981. *Camera lucida: Reflections on photography*. Macmillan.

Batty, Craig, Dyer, Adrian, Perkins, Claire and Sita, Jodi. 2015. "Seeing Animated Worlds: Eye Tracking and the Spectator's Experience of Narrative." In *Making Sense of Cinema*, edited by CarrieLynn D. Reinhard and Christopher Olson, New York: Bloomsbury, forthcoming.

Birmingham, Elina, Bischof, Walter. F., and Kingstone, Alan. 2008. "Gaze selection in complex social scene." *Visual Cognition* 16 (2–3): 341–355

Bordwell, David. 2007. *Poetics of Cinema*. London: Routledge.

Carroll, Noël. 1996. *Theorizing The Moving Image*, Cambridge, Cambridge University Press.

Cuningham, Henry. 1994. "Sherlock Holmes and the Case of Race". *The Journal of Popular Culture*, 28: 113–125. doi: 10.1111/j.0022-3840.1994.2802_113.x

Ellis, John. 1982. *Visible Fictions*, London: Routledge.

Ekman, Paul and Friesen, Wallace, V. 1971. "Constants across cultures in the face and emotion." *Journal of Personality and Social Psychology* Vol 17 (2): 124-129.

Foucault, Michel. 1977. *Discipline and Punish: the Birth of the Prison*, New York: Random House.

Itti, Laurent and Koch, Christof. 2000. (http://www.sciencedirect.com/science/article/pii/S0042698910000416#bib18)"A saliency-based search mechanism for overt and covert shifts of visual attention." *Vision Research* 40 (10–12): 1489–1506.

Jann, Rosemary. 1990. "Sherlock Holmes codes the social body". *ELH*: 685-708.

Knight, Peter. (Ed.). 2002. *Conspiracy nation: The politics of paranoia in postwar America*. New York: New York University Press.

Kuhn, Gustav, Tatler, Ben, and Cole, Geoff. 2009. "You look where I look! Effect of gaze cues on overt and covert attention misdirection." *Visual Cognition* 17 (6/7): 925-944.

Hagendoorn, Ivar. 2004. "Some speculative hypotheses about the nature and perception of dance and choreography". *Journal of consciousness studies,* 11, 79 – 110.

Hall, Stuart. 2001. :The Spectacle of the Other". *Discourse theory and practice: A reader*, Routledge: 324-344.

Hernandez, Nadia, Metzger, Aude, Magné, Rémy, Bonnet-Brilhault, Frédérique, Roux, Sylvie, Barthelemy, Catherine, and Martineau, Joëlle. 2009. "Exploration of core features of a human face by healthy and autistic adults analyzed by visual scanning." *Neuropsychologia* 47(4): 1004-1012.

Loughland, Carmel, M., Williams, Leanne. M., and Gordon, Evian. 2002. "Visual scanpaths to positive and negative facial emotions in an outpatient schizophrenia sample." *Schizophrenia Research,* 55: 159–170.

Manning, Erin. 2009. *Relationscapes: movement, art, philosophy*. Kindle edition Cambridge, Mass: MIT Press.

Marks, Laura U. 2000. *The Skin of the Film: Intercultural Cinema, Embodiment and the Senses*, Duke University Press.

Metz, Christian. 1982. *The imaginary signifier: Psychoanalysis and the cinema*. Indiana University Press.

Mulvey, Laura. 1975. "Visual Pleasure and Narrative Cinema". *Screen*, 16(3), 6-18.

Mulvey, Laura. 1989. *Visual and Other Pleasures*. London: Macmillan.

Neale, Steve. "Questions of Genre". *Screen*: 31(1) (1990): 45-65.

Rassell, Andrea, Redmond, Sean, Robinson, Jenny, Stadler, Jane, Verhagen, Darrin and Pink, Sarah. 2015. "Seeing, Sensing Sound: Eye Tracking Soundscapes in Saving Private Ryan and Monsters, Inc.". In *Making Sense of Cinema*, edited by CarrieLynn D. Reinhard and Christopher Olson, New York: Bloomsbury, forthcoming.

Redmond, Sean and Sita, Jodi. 2013. "What eye tracking tells us about the way we watch films (https://theconversation.com/what-eye%20tracking-tells-us-about-the-way-we-watch-films-19444)", 5th December, *The Conversation*. Accessed 5th September 2014.

Smith, Tim, J. 2013. "Watching You Watch Movies: Using Eye Tracking to Inform Cognitive Film Theory." In Psychocinematics: Exploring Cognition at the Movies, edited by Arthur P. Shimamura, Oxford University Press.

Smith, Tim. J. and Mital, Parag. K. 2011. "Watching the world go by: Attentional prioritization of social motion during dynamic scene viewing." [conference abstract]. *Journal of Vision, 11*(11): 478.

Stevens, Catherine, Winskel, Heather, Howell, Claire, Vidal, Lyne, Latimer, Cyril, and Milne-Home, Josephine. 2010. "Perceiving Dance: Schematic Expectations Guide Experts' Scanning of a Contemporary Dance Film. *Journal of Dance Medicine & Science*, *14*(1), 19 – 25.

Torralba, Antonio, Oliva, Aude., Castelhano, Monica. S., Henderson, John, M. 2006. "Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search." *Psychological Review,* 113 (4): 766–786.

Totaro, Donato. 2002. "Deleuzian Film Analysis: The Skin of the Film (http://www.horschamp.qc.ca/new_offscreen/skin.html%20)". *Off-Screen*, (accessed August 1st, 2011).

Treuting, Jennifer. 2006. "Eye tracking and cinema: A study of film theory and visual perception." *Society of Motion Picture and Television Engineers, 115*(1): 31-40.

Vincent, Benjamin.T., Baddeley, Roland, Correani, Alessia, Troscianko, Tom, and Leonards, Ute. 2009. "Do we look at lights? Using mixture modelling to distinguish between low- and high-level factors in natural image viewing". *Visual Cognition*, 17 (6–7), 856–879.

Vincs, Kim, and Barbour, Kim. 2014. "Snapshots of Complexity: using motion capture and principal component analysis to reconceptualise dance". *Digital Creativity, 25*(1) 62-78.

Vincs, Kim, Schubert, Emery, and Stevens, Catherine. 2009. "Measuring responses to dance: is there a 'grammar' of dance?" *Proceedings of the World Dance Alliance Global Summit*, Brisbane, July 14-16, 2008.

Watts, Amber, J., and Douglas, Jacinta, M. 2006. "Interpreting facial expression and communication competence following severe traumatic brain injury". *Aphasiology,* 20, 707–722.

Williams, Leanne, M., Loughland, Carmel, M., Gordon, Evian. and Davidson, Dean. 1999. "Visual scanpaths in schizophrenia: Is there a deficit in face recognition?". *Schizophrenia Research,* 40: 189–199.

Bios

Sean Redmond is an Associate Professor in Media and Communication at Deakin University. He has research interests in film and television aesthetics, film and television genre, film authorship, film sound, stardom and celebrity, and film phenomenology. He has published nine books including *The Cinema of Takeshi Kitano: Flowering Blood* (Columbia, 2013), and *Celebrity and the Media* (Palgrave, 2014), and with Su Holmes he edits the journal *Celebrity Studies*. Sean Redmond and Jodi Sita set up the Eye Tracking the Moving Image Research group in 2011.

Jodi Sita is an academic working within the areas of neuroscience and anatomy with expertise in eye tracking research. She has extensive experience, with multiple project types using eye tracking technologies and other biophysical data. Her current research uses eye tracking to study viewers gaze patterns while watching moving images; to examine expertise in Australian Rules Football League coaches and players and to examine the signature forgery process.

Professor Kim Vincs is Director and founder of Deakin Motion.Lab, at Deakin University. Kim integrates scientific and artistic approaches through research. She is currently working on a three-year project, supported by the Australian Research Council's Discovery program. Her collaborations with mathematicians, biomechanists and cognitive psychologists span Deakin and the Universities of Sydney, Western Sydney and New South Wales.

👍 **Like** 24   💬 **Send**

Related

Editorial - Seeing Into Things: Eye Tracking the Moving Image - Sean Redmond & Craig Batty (http://refractory.unimelb.edu.au/2015/02/06/redmond-batty/)
February 6, 2015
In "Digital Media/Internet"

From Subtitles to SMS: Eye Tracking, Texting and Sherlock - Tessa Dwyer (http://refractory.unimelb.edu.au/2015/02/07/dwyer/)
February 7, 2015
In "Animation"

Volume 25, 2015 (http://refractory.unimelb.edu.au/2015/02/06/volume-25-2015/)
February 6, 2015
In "Browse Past Volumes"