13. Captured Time: Eye Tracking and the Attention Economy

Alexander Monea

Abstract

This chapter examines the history how eye tracking came to stand in as a measurement for what people pay attention to. I argue that this connection between eye movement and attention is problematic and may lead to undesirable developments in the contemporary attention economy as it is implemented across digital platforms and smartphone apps in the near future—e.g., *ads that pause playing when you look away from the screen*. The chapter traces the emergence of eye tracking technology in the late nineteenth and early twentieth centuries and shows how its implementation to measure attention in the mid twentieth century cemented its purpose and has driven its development ever since.

Keywords: eye tracking, surveillance, attention, human-computer interaction, media studies, history of technology

It is curious that today there are nearly as many popular authors arguing that attention can be captured and sold reliably enough to form the bed-rock of the contemporary economy as there are authors arguing that with self-discipline we can harness our attention to either succeed in or escape from that very same contemporary attention economy.¹ The collective wisdom across these texts is something like the following: when an individual mechanism of attention capture works, it's zombifying capacities are near inescapable, but any individual mechanism of attention capture will not be very efficient and will be highly exploitable, subvertible, or even just ignorable. As such, the infrastructure of the attention economy has been one of pure bombardment, of inefficient spamming, which when collectively constant and immersive ensures that you are everywhere and

always ensnared by at least one or two of these mechanisms while you ignore the rest. There is thus a fundamental ambiguity in the term. Our attention has become both our greatest weakness and the source of our salvation in the information age.

The reason that attention presents us with this paradox today is because of its historical fusion with vision. It is this assumption—that the movement of the eyes constitutes a sufficient stand-in for mental focus—that renders the temporality of attention available for quantification and thus, in the terms of Lorraine Daston and Peter Galison, “mechanically objective” knowledge.² As we’ll see, the slippage between the qualitative temporality of everyday attention and the quantitative temporality of tracked eye movements is one reason for this paradox—the second will be the radical unpredictability of eye movements. Nonetheless, our culture is dominated by an idea that the patterning of lived temporal phenomena can be put into a circulatory relationship with what media theorists like Axel Volmar or Wolfgang Ernst have described as the microtemporalities of time-critical media like eye tracking sensors and the flow of electrons across the silicon circuit boards of computational devices.³ Through the mediation of eye tracking, the thoughts of mind and machine might be tethered across their temporal registers. Mind and machine might be made to correspond with and to pattern each other, which, in our most dystopic nightmare and the wildest dreams of platform shareholders, might lead to a command and control structure.

The first section examines the history of eye tracking technologies from their conception in the nineteenth century to their stabilization in the twentieth century. By mid-century, eye tracking technology was accurate, available, and affordable—it was a hammer waiting for a nail, so to speak. The second section shows how eye tracking came to be nearly synonymous with tracking attention. Cyberneticists and psychologists looking to measure attention turned to eye tracking as the only objective metric available for attention. This cybernetic legacy continually inspires the rhetoric in which eye tracking is positioned as a mechanism of command, control and communication. While this representation has always been fraught, eye tracking research often slips back into a reliance on science fiction as it acknowledges but then quickly forgets its own limitations. The final section argues that eye tracking technology currently constitutes the largest gap in the attention economy and that contemporary technologies make it increasingly feasible that attention merchants will try to capture that data. I posit two dystopic futures, one soft and one hard. In the soft dystopia, attention merchants are able to capture eye tracking data to, for instance,

make sure we pay for freemium content with our attention to advertisements. In the hard dystopia, the specificity of each individual's eye movements will become machine learnable and thus dynamic design processes will be able to modulate our perception—and perhaps attention—in a much more consistent way.

The future of the attention economy from digital platforms to smartphone apps could be inflected by the advent of eye tracking technology being placed by default in all our screen-based interfaces. The use of ubiquitous eye tracking technologies to pattern the temporal flow of attention will likely constitute a speed-up and convenience for some users while constituting a slow-down and burden for others. It is easy to imagine a future where affluent users access premium dynamic interfaces that increase their speed and productivity while a large portion of less privileged users have their attention trapped and monetized in exchange for access to apps, platforms, or the internet writ large—think Zuckerberg's "free" internet (via drones, satellites, or balloons) on steroids.

The Mechanization of Eye Tracking, or Why the Eye Can't be Controlled

Throughout the nineteenth century many attempts were made to measure eye movement. Louis-Émile Javal, for instance, used afterimages to determine that there was no vertical displacement of the eye as it glided horizontally across text during reading.⁴ This experiment is popularly understood as the origin point of modern eye tracking. A number of contemporaneous researchers were considering using feathers and bristles attached to the eye to record its movements.⁵ By the turn of the century, Edmund Burke Huey published a groundbreaking study on the psychology of reading. Huey had successfully used plaster-of-Paris to build a mold that fitted over the cornea that was connected by a light celloidin-covered glass lever to a kymograph (fig. 13.1). The glass lever maneuvered the aluminum pointer of the kymograph to record its movements on a smoked drum, and later on smoked paper, as subjects read varying texts (fig. 13.2). Huey writes, "During the reading, the reader was usually quite unconscious of there being an attachment to his eye, and the reading proceeded as glibly and easily as could be desired," perhaps because he rendered participants' eyes insensitive by the application of holocain or cocaine.⁶ Provided the subject's head was fastened between iron standards to keep it still, he could successfully record their eye movements (fig. 13.1). Thus, by limiting the

Figure 13.1. Huey's eye tracking apparatus

Figure 13.2. Scanpaths recorded by Huey

subject's mobility in space, Huey could capture the foci of the subject's eye movements in time, thus allowing what would retrospectively be enshrined as one of the first successful synchronizations of mind and machine.

Despite Huey's claims, the mechanical properties of his apparatus limited its ability to objectively capture eye movement, in particular because the inertia from the apparatus caused overshoots in the recorded traces.⁷ In 1901, Raymond Dodge and Thomas Cline wrote, "Out of this advance in the physiology of vision has developed a constantly growing group of psychological questions which only a quantitative knowledge of the eye movements can answer."⁸ Dodge and Cline worked to develop an apparatus capable of making *exact* measurements of eye movements, with a particular focus on their angle velocity. A successful apparatus would need to register eye movement without producing momentum or inertia.⁹ The only possible solution was *film*, the perfect medium for turning a sequence of captured spatial coordinates into a simulation of phenomenological temporality. In a move that would determine many of our contemporary eye tracking technologies, Dodge and Cline would catch reflected light off the cornea on a sensitive film. In an apparatus jerry-built with knitting needles, bicycle pumps, cardboard, and a modified 5 x 7 bellows camera, they would secure their subject's head and photograph their eye movements (fig. 13.3).

Their most important invention was a plate-holder inside the camera that would move the film vertically in a continuous and even motion immediately behind a narrow horizontal slit through which the reflected light off the cornea was received (fig. 13.4). In addition, they used a pendulum within the plate-holder to create a time record by having its oscillations periodically intercept the light allowed in by one side of the horizontal slit in the plate-holder (fig. 13.5).

In later studies, Dodge would come to call these exposures 'kinetograms' (fig. 13.6).¹⁰ With these kinetograms, a mechanically objective representation of the temporality of eye movements became available to researchers. The reliability and cost effectiveness of this new medium far exceeded the limited demands of eye tracking research to that date, still largely focused on reading and visual impairments. It was, in a sense, a hammer awaiting a nail that would not appear until eye movements became a stand-in for human attention writ large. While this conjuncture would be instituted largely by psychologists outside of ophthalmology laboratories, the stage was set by researchers who moved from analyzing scanpaths during reading to scanpaths during analysis of images.

Figure 13.3. Dodge and Cline's eye tracking apparatus

Figure 13.4. Dodge and Cline's film box

Figure 13.5. Dodge and Cline's film recording of eye movements

Figure 13.6. Dodge and Cline's film recording of eye movements

Figure 13.7. Buswell's eye tracking apparatus

The use of film to capture eye movements became the norm in eye tracking research for nearly 70 years and was quickly applied outside of the study of reading. George Malcolm Stratton would use a similar photographic technique to examine eye movements when subjects viewed simple geometrical patterns and line illusions. Stratton quickly found that he could reliably predict the scanpaths that his subjects' eyes would take when presented with these patterns,¹¹ and was particularly unnerved by how disharmonious eye movements could produce the aesthetic experience of symmetry.¹² As Wade notes, "Stratton's work is significant because it attempted to bridge the gap between visual phenomena (illusions), cognition (aesthetic judgments), and the underlying mechanisms (eye movements)," and further because Stratton shifted the focus of scientific study from the movement of the eyes to the locations selected for fixation.¹³ This focus on how we select particular locations to fixate on would become a central research question

Figure 13.8. Buswell's analysis of eye movements during viewing of painting

moving forward and would set the stage for the fusion of eye movement and attention research.

The next major advancement came in 1935 when Guy Buswell published the eye tracking records of nearly 2,000 individual scanpaths from 200 participants each of whom viewed multiple pictures through an apparatus very similar to Dodge and Cline's (fig. 13.7).¹⁴ Rather than focusing on the eye movements themselves, Buswell analyzed the fixation points, publishing density plots from all of his recorded scanpaths (fig. 13.8). By reducing the spatial mobility of the subject's head, Buswell was able to use film to capture the temporal dimensions of eye movements as sequences of fixations. This reduction then allowed for a spatialization of these temporal dimensions (fig. 13.8).

From this innovative work, Buswell made some startling discoveries about the uniqueness of his participants' scanpaths. People tended to have very different scanpaths when presented with the same picture, while only the first handful of fixation points had even a semblance of homogeneity. Additionally, the average duration of fixations for each picture varied widely across observers. Buswell examined whether these differences correlated well with other variables like artistic training, age, and race/ethnicity, but found that within group variation always exceeded the average variation between groups to such an extent that the correlations could not be considered significant.¹⁵ In essence, people's phenomenological temporalities were radically individualized. While mechanically objective recording of eye movements was now able to synchronize with the flow of phenomenological vision in its past and present foci, it remained unable to predict future foci.

Alfred L. Yarbus would eventually extend this research by fusing the film recordings with suction cup eye devices to investigate miniature eye movements during fixation, detailed kinematics of individual saccades, vergence, pursuit, and, perhaps most importantly, eye movements during perception of complex objects.¹⁶ For this last emphasis Yarbus famously analyzed the eye movements of subjects as they viewed Ilya Repin's painting *The Unexpected Visitor*. He found that while there was some homogeneity in the scanpaths, especially in the earlier points of fixation, the scanpaths were highly individualized. Additionally, when the same subject viewed

Figure 13.9. Yarbus's analysis of scanpaths when viewing painting based on priming.

the same painting a number of separate times, each successive viewing's scanpath would be different, although the difference in scanpaths was greater between different subjects than between successive viewings by the same subject.¹⁷ In perhaps the most enduring illustration of early eye tracking research, Yarbus extended Buswell's research into priming's effect on scanpaths (fig. 13.9). As Yarbus notes, "Depending on the task in which a

person is engaged, i.e., depending on the character of the information which he must obtain, the distribution of the points of fixation on an object will vary correspondingly, because different items of information are usually localized in different parts of an object."¹⁸

The role of priming and other cognitive factors in determining scanpaths, and in particular the sequence of selected points for fixation, is still central to eye tracking research and difficult to predict with much accuracy.¹⁹ Yarbus' study was recently reproduced and expanded to demonstrate that what he found after priming a single subject upon multiple viewings of *The Unexpected Visitor* was generalizable to multiple subjects.²⁰

Cybernetic Analysis of Black Boxes, or Why People Try to Control the Eye

It is in Paul Fitts' research from the 1940s that eye tracking first was welded to human attention. Here the foci of eye movements stand in for the foci of mental attention and thus the scanpaths give an approximation of the sequence and rhythm of human thought—particularly as a stream of consciousness. Fitts was a pioneer in human factors, served as the sixth president of the American Psychological Association, and developed important laws and formulae for how humans use visual and proprioceptive feedback to situate themselves in space and time. He explicitly positioned himself in a tradition of interdisciplinary military research that was meant to better adapt humans and machines for cooperative interaction in combat, with a particular focus on aviation.²¹ The goal of much of this collaborative research between the American Psychological Association and the military was to better select and train pilots, better design aviation technology (esp. cockpits), and thus to ensure optimal air force capacities. For Fitts, this was fundamentally an *informational* problem about the visual and proprioceptive feedback loops that allow a pilot to interact with an aircraft.²² This approach situates Fitts in a transitional discursive tradition between behaviorism and cognitivism that Paul Edwards has described as "cybernetic psychology."²³

Behaviorism understood the human mind to be a black box that could only be studied through measuring its inputs (stimuli) and outputs (behavior) with the ultimate goal of producing a stable theory of the functional relation between the two such that behavior might be controlled through the engineering of stimuli. Cybernetics largely shifted the metaphor of the black box onto electronics and used information theory to produce new methods of inferring the contents of black boxes through analysis of random noise inputs, and then

used the analogy of mind and machine to argue that the same mathematical analyses useful in understanding electronics could similarly be applied to the human mind.²⁴ The goal of cybernetic psychology was “to understand the processes of perception, memory, and language in terms of formalizable transformations of information and feedback circuits or control loops.”²⁵

While some psychologists directly equated human attention with a mechanical model of sensory inputs,²⁶ Fitts was more open about the fact that the connection between thought, attention, and vision was more a matter of convenience. He writes, “If we know where a pilot is looking we do not necessarily know what he is thinking, but we know something of what he is thinking about. In servo terminology, we know what error signal inputs he is operating on.”²⁷ What the eye is focused on is only roughly correlated to what the mind is paying attention to, which in turn is only roughly correlated to what the mind is thinking about. Eye movements and fixations were chosen as inputs to be studied because they were the only variables that could be objectively observed and mechanically reproduced. In short, eye movements became the stand-in for attention—and thus the temporality of thought—because eye tracking technology was sophisticated, affordable, and accurate enough to serve scientific and military purposes.

This is the point of fusion between a technology waiting for an application and a theory of mind looking for a method. This entanglement of thought, attention, and vision became essential to cybernetics as a whole.²⁸ Take, for example, the cybernetic understanding of teleology outlined by Arturo Rosenblueth, Norbert Wiener, and Julian Bigelow.²⁹ They understood humans and machines to pursue goals in the same way, by utilizing a cycle of predicting future states, taking actions, and self-correcting their course of action based on negative feedback received. Goals are “a final condition in which the behaving object reaches a definite correlation in time or in space with respect to another object or event.”³⁰ The sensory feedback received is primarily visual. Warren McCulloch went so far as to argue that the mind was dominated by the eye. He writes, “The eye is not only the most important of sense organs. It is the most complicated, being in reality an invaginated evagination of the brain itself.”³¹ Here McCulloch is arguing that in vision the brain is externalized and folded back in on itself; it touches itself as it touches the world. This is no more nor less than the cybernetic version of Maurice Merleau-Ponty’s insights about the chiasm—the entanglement of sensor and sensation—whose anatomical definition interestingly corresponds to the crossing optic tracts of the eyes that is so fundamental to binocular vision.³² As Rudolf Arnheim has shown at great length perception contains cognition, or as he puts it, “visual perception is

visual thinking.”³³ This is central to cybernetic psychology: we must examine sensory inputs—vision in particular—if we are to understand how the mind works as a mechanism. In the context of aviation, as Fitts notes, “we cannot study man’s motor system at the behavioral level in isolation from its associated sensory mechanisms.”³⁴

Eye tracking research continues to bear the imprint of this cybernetic legacy today. The textbook *Eye Tracking Methodology* describes the visual system repeatedly in terms of information processing, noise reduction, command and control systems, and feedback circuits and loops—including as an “attentional feedback loop.”³⁵ While the author is careful to note that he is speaking of *visual* attention and not attention writ large, the analogizing of mind and machine leads him to habitually conflate the two in practice. For instance, Duchowski refers to eye tracking data as “objective evidence provided by users’ gaze and hence attention.”³⁶ Wilson Geisler and Lawrence Cormack similarly describe attention with a cybernetic vocabulary, and Árni Kristjánsson goes so far as to argue that eye movements and attention largely share the same neural resources.³⁷ This is a discursive convention that companies producing eye tracking technologies like Tobii are all too happy to encourage. In their online marketing materials for Tobii Pro they write, “Eye tracking is a unique method to objectively measure consumers’ attention and spontaneous responses to marketing messages.”³⁸

As I’ve shown, this repeated analogizing of mind and machine was embedded in eye tracking discourse through cybernetics in the mid-twentieth century. The result of that legacy is that we are led to believe that human “systems” might be just as amenable to command, control, and communication as computational systems. However, as I’ve shown in the previous section, our scanpaths are too susceptible to priming and other idiosyncrasies for this analogy to ever hold true. Instead, just as Colin Milburn has shown in the context of nanotechnology, eye tracking is as much science fiction as science.³⁹ By this I do not mean that there is no real science behind the use of eye tracking. There is a great deal of innovative and rigorous research being conducted through these means. What I do mean is that the investment in and development of eye tracking technologies is as much driven by science fiction stories about successful applications *in the future* as it is by current applications. Regardless of the success or failure of any of these anticipated future applications, the discourse on eye tracking will continue producing science fiction so long as its cybernetic genealogy remains intact. Eye tracking seems forever destined to fail to live up to its hype because of the individual specificity of eye movements. Yet, like it is our political party or our compulsively philandering spouse, we can’t help

but believe its perpetual promises that it has turned a corner, that this time it will be different, that things are finally going to work out.

Computational Eye Tracking and Two Dystopic Futures

In 2016, Nvidia publicly announced ongoing research into the use of eye tracking technology to make gains in computer processing power for virtual reality. By taking advantage of foveal vision, Nvidia realized it could drastically reduce the processing power required to render VR graphics. In essence, Nvidia argues that VR headsets can render the graphics outside a designated range of the current eye fixation in lower resolution without the user noticing any ill effects. This may prove essential for the future of VR. As Simon Parkin writes, “When the player using the Nvidia system focuses on a new area of the scene, eye-tracking software shifts the focus of the rendering in kind. To render a full scene in VR at 90 frames per second, the lowest acceptable frame rate in VR before users begin to report feelings of nausea, four million pixels must be rendered at almost a hundred times a second. But by focusing the rendering only on the player’s line of sight, huge computational savings can be made.”⁴⁰ While Nvidia will not produce VR headsets with eye tracking technology, a number of companies are already moving in this direction. For example, the Kickstarter-backed Fove headset is meant to be completely controlled by the user’s eye movements,⁴¹ and companies like SensoMotoric Instruments and Tobii Tech are already working to adapt their eye tracking technologies to VR headsets. As people increasingly access visual information through VR media, the natural inclusion of eye tracking technology promises to offer a fresh new stream of user data.

As far back as 2011, Tobii was partnering with Lenovo to produce a laptop that could be controlled with eye motions, but these technologies have yet to sweep the consumer market.⁴² This may in part be due to the high cost, need to accommodate cumbersome hardware, and inaccuracy during real-world use that such systems have demonstrated. Recently we have seen rapid advances in software-based solutions that take advantage of preexisting technologies, like webcams and forward-facing cameras on tablets and smart phones, whose fixed position relative to the screen can be exploited for eye tracking purposes.⁴³ Machine learning has left eye tracking technology largely unimproved,⁴⁴ and this is likely due to limitations in the availability of large-scale datasets of captured eye movements—most have around 50 subjects.⁴⁵ By 2016, researchers were constructing “a mobile-based eye tracking dataset containing almost 1500 subjects from a wide variety of

background, recorded under variable lighting conditions and unconstrained head motion.”⁴⁶ They fed this data into a convolutional neural network and produced a system that could match state-of-the-art eye tracking performance. Further, they were able to use dark knowledge approaches to simplify their model enough for it to run in real-time on mobile devices.⁴⁷ Provided the ubiquity of eye tracking technology is solely a software problem, it is safe to assume that we are not far from eye tracking technology being a standard feature of all smartphones, tablets, and laptops.

The soft dystopic vision of the future is one in which the current trend of using click-through, typed input, scrolling, and time-spent-on-page data to make these attention capture mechanisms more efficient is supplemented with eye tracking data at web scale. Once our VR headsets, laptops, tablets, and smartphones incorporate eye tracking technology, it is inevitable that our eye tracking data will be collected as we navigate our digital worlds in exchange for access to new digital goods and services. The collection of vast stores of user scanpaths offers up a first solution to the individuality of our eye movements: we then might collect enough data per user to statistically model their individual eye movement behaviors in response to various stimuli and priming. As user interfaces and designs become increasingly modular and dynamic, it is not hard to envision a world in which this collection of our eye movements can lead to the personalization of our interfaces through ongoing dynamic iterations. For instance, every Google account, or even every Google search for every Google account, would generate a particular optimal placement for “sponsored content” so that it might draw the most attention and produce the most click-throughs. This will never be a world in which we become perfect ideological drones ready for ongoing hypnotic commands from global capital, but it will be one in which the effectiveness and efficiency of our current attention capturing mechanisms is vastly increased.

Even should dynamic or personalized interface designs not prove possible with eye tracking data at web scale, there are even easier to imagine dystopic futures that can arise when eye tracking technology is embedded in all of our screen technologies as a default. The business models of most of our digital platforms and the bulk of our “freemium” smartphone apps are based on advertising revenue. Imagine a world in which these companies could monitor the consumptive labor of viewing advertisements that their users are performing for them. Imagine a world in which *advertisements pause when you stop looking at them*. Where you can’t return to primary content until you’ve *paid* attention. Imagine how much more these companies could charge advertisers if they could guarantee eyeballs on their adds, literally. We

aren't far away from a Candy Crush app that tracks your eye movements to make sure you are watching the advertisements that unlock daily boosters, or from public WiFi hotspots that require you to focus on intermittent advertisements to stay connected. The most likely dystopia is one in which consumptive labor—i.e. watching advertisements—becomes a stable market and attention gets standardized into currency that you *pay with*. Here the phrase “paying attention” will finally take on its full meaning under capital.

In this first dystopic vision of the future one can already see the potential emergence of even greater class and gender differentiation in the temporal patterning of digital flows. As Axel Volmar and Kyle Stine have pointed out in their introduction to this volume and Sarah Sharma has analyzed in great detail,⁴⁸ there already exists a great class differentiation in our contemporary lived temporalities, as the optimization and acceleration of affluent people's everyday lives is a privilege borne on the backs of the underprivileged, whose lived temporalities are relationally overburdened, disrupted, and thrown out of sync with the world. We already have an instance akin to my imagined dystopia in the use of a platform like YouTube, which serves more ads to people accessing the site via mobile phones than on desktops and laptops—and keep in mind that the underprivileged people of the globe are much more likely to only have internet access through their mobile devices. One can easily envision the exacerbation of these divides by imagining the future of affluent, white, cis-gendered men paying to use optimized interfaces in the future to speed up their digital lives on platforms subsidized by the attentional labor of, for instance, working class, black, single mothers forced to pay attention to every second of the ads supporting their access to digital platforms. This situation will only be worse in areas of the globe without affordable broadband access. As platforms like Facebook look to offer drone, satellite, and balloon-based mobile internet access to their closed iterations of the “internet,” it is easy to imagine a future in which these people's entire internet access is facilitated by the extraction of attentional capital through eye tracking technologies. In this future, the digital temporality of large swaths of the globe might be patterned in accordance to these attentional extraction mechanisms.

In the hard dystopic vision of the future is one in which this process is much more successful than can be reasonably hoped for. Here we can envision a process whereby the personalization afforded by eye tracking data will lead to completely individualized interfaces and designs statistically optimized to capture attention by structuring the sequence of points of fixation in our scanpaths. This is a literal determination not only of what but of *how* we see. In his book *Visual Thinking*, Rudolf Arnheim argues that

the “cognitive operations called thinking are not the privilege of mental processes above and beyond perception but the essential ingredients of perception itself,” and also that “visual perception ... is not a passive recording stimulus material but an active concern of the mind.”⁴⁹ As such, we can understand such a capacity for dynamic iterative design based on personalized scanpath profiles to be a literal manipulation of mind. This hard dystopia fits in well with a long tradition in film criticism ranging from Horkheimer and Adorno’s woes over cinema relieving of us of any capacity to imagine for ourselves to William Connolly’s more recent arguments that the cinema can utilize affect to (re-)program its audience.⁵⁰ The primary difference here is that this will be a world in which the mechanism of attention capture is individualized beyond all possibility of communal experience, and thus short-circuits the capacity for building communities through shared experience. In his three-volume work *Technics and Time*, Bernard Stiegler worried that mass media have led to a world in which everyone’s lived temporality is synchronized with broadcast media in what he terms “the industrial temporalization of consciousness,” thus leading to a cultural leveling and homogeneity. In this hard dystopia we can see the opposite problem arising from network media and attentional modulation, where every single person’s lived temporality is synchronized with a radically individualized temporal patterning mechanism.⁵¹

The endpoint of eye tracking certainly does not need to be so grim. It is extremely promising as a method for creating user interfaces that make attention less intentional, so to speak, and thus less exhausting. This is precisely the promise that Walter Benjamin saw in film: in opposition to Horkheimer and Adorno, Benjamin argued that film was uniquely capable of “distracted reception,” where audiences could develop a critical understanding of the film’s contents without having to expend the conscious energy of intentionally directing their attention towards each nuance of the film’s contents.⁵² Personalization based on scanpath models could do something similar by making objects we need or ought to focus on easier to pay attention to—more absorbing, interesting, or intuitive. It could be the visual equivalent of Richard Thaler’s nudges.⁵³ It might make learning or communicating much easier or more efficient for many more people. The problem is with the potential weaponization of this data, which seems all but inevitable given the attentional infrastructure which such a technology would get plugged into. In such an instance, what used to be our primary defense against attention capture—namely, that eye movements and fixations were highly individualized with no clustering of behavior around demographic groupings—becomes our new primary weakness, which is

the personalization of attentional objects. Perfect echo chambers. Mass solipsism. Catatonia. A horrific monadology. The true boob tube. A real-life version of *Ringu* in which the husks of departed souls are found slouched before screens that just read as static to second parties.

Notes

1. For evidence of the former, see Thomas H. Davenport and John C. Beck, *The Attention Economy* (Cambridge, MA: Harvard Business Review Press, 2001); Mara Einstein, *Black Ops Advertising: Native Ads, Content Marketing and the covert World of the Digital Sell* (New York: OR Books, 2016); Nir Eyal, *Hooked: How to Build Habit-Forming Products*, ed. Ryan Hoover (New York: Penguin, 2014); Ben Parr, *Captivology: The Science of Capturing People's Attention* (New York: HarperOne, 2016); Jonathan Taplin, *Move Fast and Break Things: How Facebook, Google, and Amazon Cornered Culture and Undermined Democracy* (New York: Hachette, 2018); James G. Webster, *The Marketplace of Attention: How Audiences Take Shape in a Digital Age* (Cambridge, MA: MIT Press, 2016); Tim Wu, *The Attention Merchants: The Epic Scramble to Get Inside Our Heads* (New York: Vintage, 2017). For examples of the latter, see Chris Bailey, *The Productivity Project: Accomplishing More by Managing Your Time, Attention, and Energy* (New York: Crown Business, 2017); Matthew B. Crawford, *The World beyond Your Head: On Becoming an Individual in an Age of Distraction* (New York: Farrar, Straus and Giroux, 2016); Daniel Goleman, *Focus: The Hidden Driver of Excellence* (New York: HarperCollins, 2015); Neen James, *Attention Pays: How to Drive Profitability, Productivity, and Accountability* (Hoboken, NJ: Wiley, 2018).
2. Lorraine Daston and Peter Galison, *Objectivity* (Cambridge, MA: Zone Books, 2007).
3. Axel Volmar, ed. *Zeitkritische Medien* (Berlin: Kadmos, 2009); Wolfgang Ernst, *Digital Memory and the Archive*, ed. Jussi Parikka (Minneapolis: University of Minnesota Press, 2013).
4. Louis É. Javal, "Essai sur la physiologie de la lecture," *Annales d'Oculistique* 80 (1878): 240–274.
5. August Ahrens, *Die Bewegung der Augen beim Schreiben* (Rostock, Germany: University of Rostock, 1891); Edmund B. Delabarre, "A Method of Recording Eye Movements," *American Journal of Psychology* 9, no. 4 (1898): 572–572; Edmund B. Huey, "Preliminary Experiments in the Physiology and Psychology of Reading," *American Journal of Psychology* 9, no. 4 (1898): 575–586; Edmund B. Huey, *The Psychology and Pedagogy of Reading, with a Review of the History of Reading and Writing and of Methods, Texts, and Hygiene in Reading* (New York: MacMillan, 1908); E. Rähmann, "Über den nystagmus und seine ätiologie," *Archiv für Ophthalmologie* 24 (1978): 237–242.

6. Huey, "Preliminary Experiments in the Physiology and Psychology of Reading," 25–26.
7. Raymond Dodge and Thomas S. Cline, "The Angle Velocity of Eye Movements," *Psychological Review* 8, no. 2 (1901): 147.
8. Dodge and Cline, "The Angle Velocity of Eye Movements," 145.
9. Dodge and Cline, "The Angle Velocity of Eye Movements," 147.
10. Raymond Dodge, "Five Types of Eye Movement in the Horizontal Meridian Plane of the Field of Regard," *American Journal of Physiology* 8, no. 4 (1903): 307–329; Raymond Dodge, "The Participation of Eye Movements in the Visual Perception of Motion," *The Psychological Review* 11, no. 1 (1904): 1–14.
11. George M. Stratton, "Eye-Movements and the Aesthetics of Visual Form," *Philosophische Studien* 20 (1902): 343.
12. Stratton, "Symmetry, Linear Illusions and the Movements of the Eye," *The Psychological Review* 13, no. 2 (1906): 95.
13. Nicholas J. Wade, "Pioneers of Eye Movement Research," *i-Perception* 1, no. 2 (2010): 57.
14. Guy Buswell, *How People Look at Pictures: A Study of the Psychology and Perception in Art* (Chicago: University of Chicago Press, 1935).
15. Buswell, *How People Look at Pictures*, 131.
16. Alfred L. Yarbus, *Eye Movements and Vision*, trans. Basil Haigh (New York: Plenum Press, 1967).
17. Benjamin W. Tatler, Nicholas J. Wade, Hoi Kwan, John M. Findlay, and Boris M. Velichkovsky, "Yarbus, Eye Movements, and vision," *i-Perception* 1, no. 1 (2010): 7–27.
18. Yarbus, *Eye Movements and Vision*, 192. In reference to the figure, Yarbus writes: "Seven records of eye movements by the same subject. Each record lasted 3 minutes. The subject examined the reproduction with both eyes. 1) Free examination of the picture. Before the subsequent recording sessions, the subject was asked to: 2) estimate the material circumstances of the family in the picture; 3) give the ages of the people; 4) surmise what the family had been doing before the arrival of the "unexpected visitor"; 5) remember the clothes worn by the people; 6) remember the position of the people and objects in the room; 7) estimate how long the "unexpected visitor" had been away from the family" (174).
19. Tatler et al., "Yarbus, Eye Movements, and Vision," 24–25.
20. Marianne DeAngelus and Jeff B. Pelz, "Top-down Control of Eye Movements: Yarbus Revisited," *Visual Cognition* 17, no. 6–7 (2009): 790–811.
21. Paul M. Fitts, "German Applied Psychology during World War II," *American Psychologist* 1 (1947): 151–161.
22. Fitts, "The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement," *Journal of Experimental Psychology* 47, no. 6 (1954): 381–382.
23. Paul N. Edwards, *The Closed World: Computers and the Politics of Discourse in Cold War America* (Cambridge, MA: MIT Press, 1996).

24. Ross Ashby, *An Introduction to Cybernetics* (London: Chapman & Hall, 1956); Norbert Wiener, *Cybernetics: or Control and Communication in the Animal and the Machine*, 2nd ed. (Cambridge, MA: MIT Press, 1961).
25. Edwards, *The Closed World*, 179–180.
26. E.g. Donald E. Broadbent, “A Mechanical Model for Human Attention and Immediate Memory,” *Psychological Review* 64, no. 3 (1957): 205–215.
27. Paul M. Fitts, Richard E. Jones, and John L. Milton, “Eye Movements of Aircraft Pilots during Instrument-Landing Approaches,” *Aeronautical Engineering Review* 9, no. 2 (1950): 1–6.
28. Orit Halpern, *Beautiful Data: A History of Vision and Reason since 1945* (Durham, NC: Duke University Press, 2015).
29. Arturo Rosenblueth, Norbert Wiener, and Julian Bigelow, “Behavior, Purpose and Teleology,” *Philosophy of Science* 10, no. 1 (1943): 18–24.
30. Rosenblueth et al., “Behavior, Purpose and Teleology,” 18.
31. Warren S. McCulloch, “Information in the Head,” *Synthese* 9, no. 1 (1955): 235.
32. Maurice Merleau-Ponty, *The Visible and the Invisible*, trans. Alphonso Lingis, ed. Claude Lefort (Evanston, IL: Northwestern University Press, 1968).
33. Rudolf Arnheim, *Visual Thinking* (Berkeley and Los Angeles: University of California Press, 1969).
34. Fitts, “Information Capacity,” 281.
35. Andrew T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 3rd ed. (Cham, Switzerland: Springer International, 2017). Duchowski even describes the retina’s photoreceptors as “transducers” (18).
36. Duchowski, ix, 3–4.
37. Wilson S. Geisler and Lawrence K. Cormack, “Models of Overt Attention,” in *The Oxford Handbook of Eye Movements*, eds. Simon P. Libersedge, Iain Gilchrist, and Stefan Everling (New York: Oxford University Press, 2011), 439–454; Árni Kristjánsson, “The Intriguing Interactive Relationship between Visual Attention and Saccadic Eye Movements,” in *The Oxford Handbook of Eye Movements*, 455–470.
38. “Tobii Pro: Fields of Use,” Tobii, accessed May 12, 2018, <https://www.tobii.com/fields-of-use/>.
39. Colin Milburn, *Nanovision: Engineering the Future* (Durham, NC: Duke University Press, 2008).
40. Simon Parkin, “Nvidia’s Eye-Tracking Tech Could Revolutionize Virtual Reality,” *MIT Technology Review*, July 21, 2016, <https://www.technologyreview.com/s/601941/nvidias-eye-tracking-tech-could-revolutionize-virtual-reality/>.
41. Parkin, “Point, Click, and Fire in Virtual Reality—with Just Your Eyes,” *MIT Technology Review*, June 23, 2015, <https://www.technologyreview.com/s/538711/point-click-and-fire-in-virtual-reality-with-just-your-eyes/>.
42. Tom Simonite, “A Laptop That Knows Where You’re Looking,” *MIT Technology Review*, March 4, 2011, <https://www.technologyreview.com/s/423198/a-laptop-that-knows-where-youre-looking/>. In fact, whole conventions have since been held with prototypes demonstrating the future of commands in screen-based media. See Simonite, “PC Makers Bet on Gaze, Gesture, Voice,

- and Touch,” *MIT Technology Review*, January 10, 2013, <https://www.technologyreview.com/s/509751/pc-makers-bet-on-gaze-gesture-voice-and-touch/>.
43. Kyle Krafska, Aditya Khosla, Petr Kellnhofer, Harini Kannan, Suchendra Bhandarkar, Wojciech Matusik, and Antonio Torralba, “Eye Tracking for Everyone,” *arXiv preprint arXiv:1606.05814* (2016): 2176–2184.
 44. Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling, “Appearance-Based Gaze Estimation in the Wild,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), 4511–4520.
 45. Qiong Huang, Ashok Veeraraghavan, and Ashutosh Sabharwal, “TabletGaze: Unconstrained Appearance-Based Gaze Estimation in Mobile Tablets,” *Machine Vision and Applications* 28, no. 5–6 (2017): 445–461; Yusuke Sugano, Yasuyuki Matsushita, and Yoichi Sato, “Learning-by-Synthesis for Appearance-Based 3D Gaze Estimation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014): 1821–1828.
 46. Krafska et al., “Eye Tracking for Everyone,” 2.
 47. Geoffrey Hinton, Oriol Vinyals, and Jeff Dean, “Distilling the Knowledge in a Neural Network,” *arXiv preprint arXiv:1503.02531* (2015): 1–9.
 48. Sarah Sharma, *In the Meantime: Temporality and Cultural Politics* (Durham, NC: Duke University Press, 2014).
 49. Arnheim, *Visual Thinking*, 13, 37.
 50. Max Horkheimer and Theodor W. Adorno, *Dialectic of Enlightenment*, trans. Edmund Jephcott, ed. Gunzelin S. Noerr (Stanford, CA: Stanford University Press, 2007); William E. Connolly, *Neuropolitics: Thinking, Culture, Speed* (Minneapolis, MN: University of Minnesota Press, 2002).
 51. Bernard Stiegler, *Technics and Time, 1: The Fault of Epimetheus*, trans. Richard Beardsworth and George Collins (Stanford, CA: Stanford University Press, 1998); Bernard Stiegler, *Technics and Time, 2: Disorientation*, trans. Stephen Barker (Stanford, CA: Stanford University Press, 2008); Bernard Stiegler, *Technics and Time, 3: Cinematic Time and the Question of Malaise*, trans. Stephen Barker (Stanford, CA: Stanford University Press, 2010).
 52. Walter Benjamin, “On the Work of Art in the Age of Mechanical Reproduction,” trans. Harry Zohn, in *Illuminations*, ed. Hannah Arendt (New York: Schocken, 1968), 217–253.
 53. Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (New York: Penguin, 2009).

About the Author

Alexander Monea is Assistant Professor of English and Cultural Studies at George Mason University (4400 University Drive MSN 3E4, Fairfax, VA 22030, United States). His research focuses on the history and cultural impact of big data and machine learning. His upcoming book examines heteronormative bias in internet content moderation.