Papers 4 - Beyond the Interview: New Methods for Collecting Data

# Eye Tracking in Medical Ethnography: Evaluating Evidence for Perception, Action, and Collaboration in Healthcare Professions

LARRY S. MCGRATH

Design Science Consulting, Inc.

Using eye tracking in ethnographic research poses numerous theoretical and practical challenges. How might devices originally intended to record individuals' vision of two-dimensional planes be useful in interpersonal contexts with dynamic visual interfaces? What would the technology reveal about collegial environments in which different levels of knowledge and expertise come together and inform decision-making processes? Why would pupil movement show us anything that conventional ethnographic methods could not? In this paper, I argue that these challenges are not intractable. When tailored to specific questions about perception, action, and collaboration, eye trackers can reveal behaviors that elude ethnographers' gaze. In so doing, the devices enrich the observational and interview-based methods already employed in ethnographic studies of workplace dynamics.

Hospitals are a fruitful context in which to test the value of eye-tracking evidence. Healthcare professionals look, interpret, and act on wide-ranging streams of visual information. The focus of this paper is to explore the possibility of carrying out potential fieldwork in pulmonology suites. I put forward a hypothetical eye-tracking framework to interrogate: 1) the different perceptual skillsets exhibited by pulmonologists; 2) the collaborative decision-making practices by which medical teams interpret monitors' bronchoscopic imagery and perform lung biopsies. According to this model, optical evidence (i.e., fixation durations, scan paths, and heat maps) could demonstrate not only how healthcare professionals look at visual interfaces; the evidence could be especially useful to show when and where vision runs aground. Such gaps in eye-tracking data illuminate additional sensory systems which facilitate information consumption and exchange. When brought to bear on potential fieldwork, the eye-tracking model offers an incisive point of departure for ethnographers to investigate, observe, and present questions about peoples' perceptual expertise.

Far from opening a glimpse "inside" peoples' heads, eye trackers add one perspective among others. When accommodated to the mobile conditions of immersive research, the technology offers a modest, though useful, contribution to ethnographers' toolkit.

## INTRODUCTION

"Feel my hand," says a young pulmonology fellow as she extends her right arm toward me. Her fingers quiver; they cup inward, bearing the trace of a grip held tight for over an hour. We are in a major research hospital located on the West Coast of the United States and the fellow has just completed a bronchoscopic lung biopsy on an elderly man suspected to have cancer. I asked what it was like to carry out the procedure. Her hand tells the whole story. Bronchoscopes are 115cm-long tubes with a small probe at the end. It generates a live image of the airways on a monitor. The tube slides down a patient's throat and snakes through the bronchi thanks to the guidance of a handle grasped in the pulmonologist's hand. Manipulating a bronchoscope demands a high degree of manual dexterity. Awkward twists of the wrist and shifts of the body rotate the instrument's direction while the thumb and index finger depress to flex the probe and extract tissue samples. Now the fellow is exhausted. The chief pulmonologist, under whom she is studying, handed her the reigns to conduct the biopsy from beginning to end. Although it's not her first time, she's hardly experienced by medical standards. Behind her are four years of medical school and another four of residency. The pathway to mastering the technique will involve many more throbbing hands over the next three years of the fellowship until she learns to loosen her grip, feel the shape of the airways' anatomy, and rely ever so slightly less on the monitor's visualizations before her eyes.

Experienced pulmonologists often take for granted just how difficult it can be to cultivate the necessary skillset. Monitors' streams of imagery compete for attention with the physical demands of bronchoscopes. How does one strike the right balance? The skills feel innate after years of practice. Such was the line of thought conveyed by the chief pulmonologist. She reflected: the more you do, the more you forget what was hard about it when you first started; over time that becomes seamless and you don't remember until you're trying to teach someone. Teaching fellows entails an intensive mentorship, which is made no easier when inculcating what comes across as an inscrutable knack. How might such an exacting but crucial skillset for bronchoscopy be made transparent, intentional, and helpful for those still developing their practice?

The scenario is ideal for eye tracking. The technology renders implicit practices explicit, which is key to understanding the perceptual skillsets that distinguish experts from novices. This is especially the case in light of the established theory that rational reflection plays a minimal role in expert perception.<sup>1</sup> Eye trackers have been used to study, for instance, how professional pianists' vision hovers between the keyboard and sheet music; both remain in view. By contrast, the tyro's gaze frequently jumps back and forth from notation to the keys.<sup>2</sup> Similarly, psychologists have tracked the pupil movement of elite cricket players and found that they do not follow the ball throughout its flight when batting. Their vision is predictive; it fixates on the path that the ball is likely to follow.<sup>3</sup> Yet, ethnographers like ourselves who study workplace dynamics often shy away from eye trackers. Although the devices could be used to analyze perceptual skillsets and make their constitutive elements intelligible, disciplinary reasons stand in the way. Anthropology examines public behaviors in shared cultural contexts.<sup>4</sup> Apart from winking, pupil movement is a private activity (or at least it appears so at first blush). Eye tracking might thus seem alien to the immersive and interactive environments in which ethnographic inquiry unfolds. In the sections that follow, I argue the contrary is true. I set about establishing principles to adapt eye tracking to potential ethnographic studies. When addressed to specific questions about looking, feeling, and acting, eye tracking can be used to transform private practices into public knowledge.

Exploring perceptual behavior is not new to workplace ethnography. Charles and Marjorie Goodwin showed how airport personnel look at airplanes to make quick decisions about transferring baggage. Despite their brevity, baggage loaders' momentary glances are structured by vast organizational practices of information sharing.<sup>5</sup> The study used cameras to record personnel from multiple angles while they interacted with airplane code sheets and identified haphazardly parked airplanes. Lucy Suchman carried out similar video studies. She recorded airport operations rooms in which people coordinate plane schedules using computer screens, radio frequencies, and telephone lines.<sup>6</sup> Eye trackers add another camera to the mix.

My argument is that eye tracking stands to enrich ethnographers' observational and interview-based methods so long as the technology is adapted to anthropological methods and applied to specific problems at the intersection of perception and action. The task is not innocent. It involves theoretical and practical challenges. Eye trackers were originally designed to evaluate stationary observers' gaze of two-dimensional surfaces. Effectively incorporating the devices into field research hinges on whether ethnographers can accommodate them to the dynamic conditions of three-dimensional environments. Ultimately, such challenges are not insurmountable. An aim of this article is to present methodical insights from psychology, philosophy, and anthropology, which help to wrest eye tracking from its conventional uses in superficial visual studies.

At Design Science, we brought these insights to bear on field research in American hospitals. Our goal was to understand how diverse levels of expertise inflect perception and action, particularly where consuming and communicating dense flows of visual information is a routine aspect of work. In hospitals, effectively interpreting and promptly acting on networks of monitors, charts, x-rays, and other imagery is of vital concern to patient safety. The stakes heighten given the fact that healthcare professionals' perceptual skillsets often diverge widely. Nurse practitioners exchange metrics and orders with new interns; advanced radiologists interpret fluoroscopy scans alongside fresh technicians. Yet, all parties engage with the same visual interfaces.

The focus of our model is pulmonologists. After observing them in pulmonology suites, we reconstructed their perceptual repertoire in an effort to create educational resources for the less experienced. Learning to wield a bronchoscope involves trial and error. Understanding skilled perception in action can make the process more manageable.

Moreover, pulmonology suites are a useful context in which to test the limits of eye tracking. Even though the purpose of bronchoscopy is to visualize the airways for the sake of diagnostic or therapeutic interventions, we found it remarkable how integral sensory systems other than vision are to the procedure. One pulmonologist noted that with practice, he cultivated the tactile ability to know where he was in the airways without seeing them clearly on the bronchoscopic monitor. Pulmonologists see less than they perceive. The models used to interpret their pupil movement should, therefore, point beyond the visual field. More generally, the limits of eye-tracking evidence could reveal opportunities to delve deeper into peoples' perceptual skillsets.

#### EYE TRACKING TWO DIMENSIONS IN TRANSITION

Although fast-paced workplace environments like hospitals are ripe to be examined through the lens of eye trackers, the technology remains beholden to lingering methodological assumptions. These have persisted since its initial application over 70 years ago.<sup>7</sup> In 1947, Paul Fitts and his colleagues conducted pioneering eye-tracking studies in airplane cockpits.<sup>8</sup> A mounted camera filmed the pupil movements of 40 pilots as they sat and scanned the array of controls, buttons, and dials while flying C-45 planes. The study's objective was to determine how perceptual practices diverged among pilots with various skill levels. Two consequences followed. First, experienced pilots tended to dwell on instruments for less time. They exhibited shorter fixation durations, which Fitts took to mean that the experienced pilots retrieved information quicker than did their less experienced counterparts. Second, he calculated the frequency with which pilots shifted their gaze among instruments and divided the frequency by the total number of transitions. The result revealed which instruments were linked in the pilots' visual field. By evaluating divergences in perceptual expertise, Fitts and his team optimized the arrangement of instruments in cockpits to facilitate smooth decision-making processes. The study laid the foundations of modern usability research but also predisposed eye-tracking methods to two-dimensional visual fields.

Eye tracking has come a long way. Fitts and his colleagues manually classified pilots' eye movement after the fact using film reels. Mobile eye-trackers now sit atop observers' heads and automatically generate fixation data. The technology at work is pupil center corneal reflection. An infrared light illuminates the eye to bring about visible reflections. A small camera captures these reflections, which identify the light source on the cornea and in the pupil. Software calculates the vector of the angles of the reflection patterns between the cornea and pupil; the vector is then used to determine the direction of the observer's gaze. Discrete time stamps with x/y coordinates are assigned to these movements—what are called "gaze points."

Unto themselves, gaze points offer little information. They become meaningful once rendered fixations. When an observer looks over a visual plane, his or her eyes move rapidly before settling on certain points. That moment when the eyes halt their movement and achieve relative stability is a fixation. Fixations are quick. They last between 50 and 600 milliseconds. They're also a useful unit of analysis because fixations serve as indications of information consumption. They mark the minimum threshold of ocular stability necessary for an observer to notice visual content. Eye-tracking systems include a fixation filter, which uses an algorithm to translate gaze points into fixations. What's left out are saccades: the eyes' motion between one fixation and the next. The average duration of a saccade is 20 to 40 milliseconds.

Fixations are made useful when correlated with the space where they're directed. Researchers take a sample image of the visual interface and divide it into areas of interest. Each corresponds to a discrete region of the surface where one's gaze might fall. An automapping program is then used to calculate the sequence, frequency, and duration of fixations across the various areas of interest. What results is a bank of data falling under categories such as fixation duration, visit count, average visit duration, and time to first fixation.

Although technologies have evolved since the time of Fitts' study, much remains the same. Eye-tracking evidence did not directly entail either of Fitts' conclusions. He inferred the meaning of pilots' perceptual activity from their pupils' movement. Fitts correlated fixation duration with the difficulty of information extraction and fixation frequency with the importance of the area seen. Today as well, eye-tracking evidence is only as valuable as the inferences we draw. When contexts shift, our categories do as well. While exploring an art museum, for example, my lingering gaze would likely be indicative of the paintings' appeal—not their complexity. Eye-tracking metrics do not speak for themselves.

Two-dimensional eye tracking is useful in certain aspects of medicine and healthcare. At Design Science, the technology allows us to collect ocular data and to draw inferences about the usability of medical devices. We use the Tobii Pro Glasses 2. This mobile eye tracker rests on a study participant's nose and ears like normal glasses (Figure 1). The glasses weigh 45 grams and include four sensors emitting infrared light to the retinas at a 100-Hz sampling rate. Along the bridge is a high-definition scene camera directed outward to the

visual field; its recording angles reach 82° horizontally and 52° vertically. The infrared reflections are superimposed on the camera footage to generate our eye-tracking metrics.

It's worth noting that the recording angles as well as the pupil center corneal reflection technology are artificial apparatuses. Their purpose is to focus on the center of the visual field. The ocular movement tracked is foveal. This is the region of the eye extending one degree from the center (the macula), which contains cone photoreceptors governing directional vision of objects. Eliminated from the eye-tracking sensors are para-foveal vision (extending to five degrees from the center) as well as peripheral vision (occupying the retina beyond five degrees). By isolating foveal vision from the natural spectrum of vision, eye trackers can be used to analyze separate fixations, each of which lands on identifiable content.



Figure 1. IFU usability study with simulated task completion.

As an example of our usability research, we use the Tobii Pro Glasses 2 to examine and help optimize Instructions for Use (IFUs). Medical devices and products are sold with IFUs, which inform users and lay caregivers about proper uses, risks, and benefits. In the United States, the Food and Drug Administration requires that all instructions and labels pass usability testing to certify that they are easily readable.<sup>9</sup> This is a service we offer. During IFU studies, we observe users as they read an IFU and carry out simulated tasks. Perhaps the participant administers an injection. Our researchers observe participants with an eye to use errors. We then ask follow-up questions to ascertain their root cause. Arriving at the fundamental causes underlying errors is crucial to finding design remedies that would correct possible future use errors. With eye tracking, we can plunge even deeper and ask questions whose answers often elude traditional root-cause analysis. Which elements of an IFU stand out to users? Do they ignore, read, or skim certain elements? Do users rely more on images or text to comprehend the IFU? (Figures 2 and 3 below indicate the latter.) What do users observe when performing critical steps? Our reiterative research process serves to evaluate

and compare multiple versions of IFUs. It thus becomes easier to eliminate design flaws and to reorganize IFUs' visual architecture for intuitive reading.

After conducting IFU comprehension studies, we use Tobii Pro Lab's automapping program to generate a wide array of eye-tracking data. First, a sample image of the IFU is divided into areas of interest, each corresponding to distinct tasks (and further divided into images and text). Second, the program calculates correlations between each area and participants' fixations. We've found three kinds of data to be useful: fixation duration metrics (indicating the distribution of fixations on each area), gaze pathways (visualizing the sequence in which users read IFU sections), and heat maps (showing which areas solicit users' fixations).



Context is key. The data reveal little by themselves. By focusing on participants' intentional tasks, such as preparing the skin for injection, our researchers draw probabilistic inferences from eye-tracking data. Whether a participant fixates on certain IFU areas, and – more importantly – passes over others, is more likely to be indicative of perceptual acts of attention – and not, for instance, curiosity or distraction. Addressing goal-oriented ocular activity allows us to deliver tailored design recommendations. We might reorganize texts and images, add visual prominence at critical steps, or highlight overlooked areas. The end products are more intuitive IFUs that meet FDA guidelines.

## TRACKING EYES IN IMMERSIVE CONTEXTS

It may seem that eye tracking is best suited for narrow scenarios in which the space of action and perception comprises only two planes. I'd like to suggest, instead, that the technology is pliable. Its utility hinges on the methods and settings brought to eye-tracking studies. Adapting eye trackers to ethnographic fieldwork turns in large part on researchers' willingness to eschew the assumptions of two-dimensional user research. The interfaces examined in much of that research is accessible at a glance – as if the interface were projected from a stable point. Participants also remain stationary while observing. This setup is alien to the immersive and interactive contexts in which ethnographic inquiry unfolds. Those contexts are three-dimensional and unstable. Therein, perception operates via active exploration of the environment and not via passive observation of a surface. People do not look *at* the visual field. A goal of eye tracking in ethnographic research is to understand how people situate themselves *in* the field.

Adapting eye trackers to immersive contexts is not seamless. The technology has proved to be congenial to two-dimensional interfaces in large part because researchers rely (often implicitly) on two interlocking theories of perception. The first is that perceptions derive from sense data; the second is that perceptions constitute mental representations. According to the first theory, the sensible qualities of objects explain the content of perceptual experience.<sup>10</sup> The shape of the font on an IFU or the color of dials in a cockpit all comprise sense data. One is directly aware of each datum. Moreover, sense-data theory counts among perceptual experience only those qualities directly sensed. That I don't see the reverse side of the IFU or a lever hidden by the dial means that whatever lies occluded before my eyes is not a sense datum.<sup>11</sup> The theory accords so well with the experimental setup of much usability research because automapping programs measure individual fixations in so far as they correlate with what is immediately accessible to vision. Just as sense data are interposed between objects in the world and visual experience, so too do areas of interest on a sample image serve to bridge the visual surface and the participant's perceptual activity.

According to the second theory, perceptions are episodes in the mind which correspond to (and thereby represent) a state of affairs. It follows that when we perceive objects in the world, we perceive them by way of representations. A representation is typically thought to be a sensory state. For the weak version of the theory, this sensory state is a relation one has to objects – my representation, for example, is of the white color in the IFU's background and of the black font in the foreground.<sup>12</sup> For the strong version, a sensory state is a mental item with semantic properties – e.g., my visual perception "that there is a fourth and fifth step in the IFU."<sup>13</sup> Both versions construe representations as pictures projected before vision. And much like a picture, representations include contents which one can describe. The theory nicely accommodates two-dimensional eye-tracking research because fixations function like representations. Researchers use fixations to infer referential acts of attention from the pupils' movement.

Plenty of ink has been spilled contesting both sense-data and representational theories of perception. Resolving those disputes is not my aim. Rather, I'd like to identify and elaborate on theories of perception more appropriate to the methodologies and set up unique to eye tracking in ethnographic research. Sense-data and representational theories inform much of current eye-tracking studies because those theories construe perception in analogous terms. Both take visual perception to function like a camera. For the sense-data theorist, pictures are projected onto our field of view. For the representational theorist, vision is akin to looking at pictures. In either case, visual perception is figured as if it occupied a plane orthogonal to an observer's point of view. Ethnographers would find this approach unbefitting to their work. When studying immersive contexts, our subjects do not see what is projected from a specific place; they see what is accessible in an entire space.

Over the past half-century, psychologists, philosophers, and neuroscientists have contributed to a constellation of theories which posit visual perception as an active engagement with the environment. Taking as their point of departure that perceivers are embedded in an interactive context, these theories claim that vision unfolds in two directions. On the one hand, people rely on sensorimotor and cognitive skills to perceive a space visually. The eyes' vision is of a piece with the body's movement. As James Gibson succinctly wrote, "we must perceive in order to move, but we must also move in order to perceive."<sup>14</sup> On the other hand, space is made accessible only partly due to vision. Such is obviously true in the trivial sense that a variety of sensory systems (i.e., tactile, auditory, and proprioceptive) also facilitate spatial navigation. More importantly, one often subdues some visual elements in order to perceive others in the visual field effectively. Maurice Merleau-Ponty put it thus, "To see the object, it is necessary not to see the play of shadows and light around it."<sup>15</sup> In both cases, vision does not happen *to* someone; people make the environment visible. In neither case does perceptual activity take place, as it were, between the ears.

If we take seriously the idea that visual perception is a form of action in the environment, then a few implications follow. First, what people see depends on where their bodies stand. Given that vision reflects a local perspective in space, how someone positions his or herself is integral to the constitution of the visual field. Whereas superficial evetracking studies control participants' perspective by stabilizing their position, ethnographic eye-tracking research is an inescapably perspectival affair. We might, therefore, consider eye trackers to record only a part of vision-namely, the trace left by peoples' embodied movement. Second, people move their bodies in order to act on the environment. In making use of the environment, it appears not simply as a stable source of sensory stimuli, but as a malleable resource which serves our purposes. Visual perception is goal-oriented in the sense that it consists of more than what is actually visible. It also includes what is potentially visible. One moves through space sometimes to gather information, but more often to change the space. Eye-tracking video could therefore be used to reveal how subjects actively transform the visual field. Third, the utility found in the visual field depends on the perceptual skills brought to it. What is useful might not be immediately visible. Objects' color, shape, or location often obfuscate vision. How people alter the visual field to their advantage hinges on a repertoire of habitual, cognitive, and sensory techniques. Alva Noë puts it thus: "Perceptual experience just is a mode of skillful exploration of the world."<sup>16</sup>

Eye trackers offer one tool for visualizing the multi-faceted skillsets with which people explore the environment—particularly in medical facilities where visual displays often require years of training to comprehend. Visual perception is incomplete. It involves competencies and sensory systems that lie beyond what can be seen. Yet, that is not at all a reason to dispense with eye tracking. The technology can supplement the tools already available to ethnographers, including close observation, behavior mapping, and participant interviews. Eye tracking might be most useful, in fact, where vision cedes to other modalities of spatial navigation. Consequently, some of the data categories used in conventional eye-tracking research are inadequate to ethnographic fieldwork. Since they are not stable, immersive contexts cannot easily be divided into fixed areas of interest. So, tracking fixations' duration and frequency is often a baffling task.

Nonetheless, gaze pathways, heat maps, and stability tracking can offer useful evidence. Gaze pathways reveal the sequence in which people make use of visual resources. Visual experience does not come together at once. Tracing how people gradually maneuver a space over time can shed light on the activities by which they make the environment useful. Heat maps can also be useful when the visual field remains relatively stable. These visualizations indicate not only where vision, but also where the body, moves in space. In addition, researchers can track the stability of pupils' gaze, which often represents the level of expertise involved in visual perception. Those familiar with the environment exhibit smooth ocular movements. By contrast, erratic ocular movements tend to be found among perceivers less skilled in discerning, retrieving, and anticipating visual resources.<sup>17</sup> Data for neither category are generated automatically. Gaze pathway and stability offer two frames of reference through which to analyze eye-tracking recordings.

Far from revealing a hidden stratum of subjective experience, eye tracking actually offers ethnographers another context to observe, describe, and interpret. The inferences drawn from the evidence are only as accurate as the tools ethnographers already bring to their fieldwork. As I suggest in the following section, eye tracking offers a fruitful context when it shows not only peoples' visual navigation of space, but also the limits of their visual perception.

### PERCEPTION AND ACTION IN BRONCHOSCOPY

The setting is a pulmonology suite in a research hospital in the Northeast United States. It's a busy environment. Below, webs of wires are taped to the floor so that nobody trips. An anesthesiology workstation sits in the far left corner; adjacent is a mobile nursing station with two screens on which patients' vital signs flow from left to right; cabinets of medical supplies line the walls; a back table where instruments are prepared extends across the right side; near us stands a four-foot tall bronchoscopy power station. It runs wires to a large wall-mounted monitor. This is the focal point of the room. Everyone's gaze is directed at it. The monitor visualizes the interior of the respiratory passage. It is the portal through which the medical team will navigate their way to a cluster of lesions in the patient's upper left lung, which are suspected to be cancerous.

About 230,000 Americans are diagnosed with lung cancer each year.<sup>18</sup> A bronchoscopic lung biopsy is a standard diagnostic procedure. But it is not straightforward. Bronchoscopes are long, flexible devices inserted through the throat. The device includes a white-light probe, which produces a live image of the airways on the monitor. Today, the image will help orient the pulmonologist to the lesions, which are challenging to reach at the edge of the mediastinum (the central region of the thoracic cavity containing the heart, trachea, and lymph nodes). She will proceed to insert a 22-gauge fine needle aspirate through the bronchoscope's inner tube and remove small tissue samples from the lesions. They can be moving targets. When the patient coughs, his lungs move – making the procedure no easier. Finally, the samples will be handed over to a cytopathologist, whose job is to analyze them under a microscope and detect malignancies.

The objective of our observations is to understand the pulmonologist's navigational technique. Guiding a bronchoscope through the airways is a delicate process. It demands not only that the pulmonologist perceives visual imagery on the monitor but also that she manually contorts the bronchoscope's handle, twisting her wrist and shifting her body, while feeling tactile impressions of the bronchial walls. The pulmonologist is trying to see inside the patient's airways in order to help him. But she doesn't rely on sight alone. Pulmonologists develop a muscular memory of the airways' anatomy. The monitor is there to offer verification.

Understanding this perceptual skillset would allow us to make bronchscopic expertise transparent, particularly for the sake of those learning the practice. In this respect, ethnography offers a useful educational tool, which imparts concepts to the pulmonologist's otherwise immediate perceptual process—what Hubert Dreyfus calls "absorbed coping." Learners' apprenticeship is cognitive; they rely on explicit rules to comprehend new techniques. Experts don't. Dreyfus elaborates, "although many forms of expertise pass through a stage in which one needs reasons to guide action, after much involved experience the learner develops a way of coping in which reasoning plays no role. Then, instead of relying on rules and standards to decide on and justify her actions, the expert immediately responds to the current concrete situation."<sup>19</sup> Our aim is to develop a hypothetical eye-tracking model that could disentangle the mélange of perceptual skills. What feels natural to the expert pulmonologist might thus be made deliberate for the fellow.

The procedure is about to begin. An elderly man lying on his back in a mobile patient bed is rolled into the room through doublewide doors. Casual banter with the pulmonologist sets him at ease while the nurse anesthetist administers moderate sedation. Semi-conscious, the patient is intubated and the bronchoscope soon makes its way into his mouth, over his tongue, and down the trachea. An endoscopic image appears on the wall-mounted monitor. It depicts bronchial rings sliding downward as the bronchoscope advances toward the left main bronchus. The pulmonologist's eyes remain glued to the monitor. She confirms that the left upper lobe is near, about two centimeters past the main carina.

Outside her view, the pulmonologist's hands do all of the work. They rotate left then twist right to orient the bronchoscope's tip. At one point, she shifts her body 90 degrees; her right shoulder now faces the monitor. The pulmonologist's own body is the vehicle by which she comprehends that of the patient.<sup>20</sup>

Once the lesion comes into view, its time to activate the ultrasonic probe. The bronchoscope is equipped with endobronchial ultrasound (EBUS), which visualizes structures beyond the airways walls. This second image allows the pulmonologist to see what the white-light probe can't: inside the lesion. She passes a needle into the bronchoscope and attempts to drive it through the side of the airway and into the lesion. But the cartilage lining the airway is thick. The needle deflects. Feedback appears on the monitor, distorting the ultrasonic image (Figure 4).



Figure 4. EBUS monitor with ultrasonic image in the center (with feedback) and endoscopic image in bottom left

The respiratory technician rushes over. He holds the base of the bronchoscope, just above the patient's mouth, helping the pulmonologist to secure the ultrasonic probe in place. Four hands now grip the device. Their manual contortions give rise to what appears in the monitor's dual images. The imagery connects the movement of the medical team with that of the patient. The situation makes literal Merleau-Ponty's words: "there is an immediate equivalence between the orientation of the visual field and the awareness of one's own body as the potentiality of that field."<sup>21</sup> Although the two images on the screen are of the patient's bronchus, they also lay bare the teams' collaborative corporeal efforts to drive through it.

Puncturing the lesion is the most fragile part of the procedure. Pushing the needle too far could result in trauma. So, the pulmonologist gently feathers the needle while verifying its placement on the monitor. The task demands immense attention. She shifts her gaze between two images: the ultrasonic image of the lesion in the center and the endoscopic image of the airway in the bottom left. Her eyes also look elsewhere.

Our simulated eye-tracking model, generated thanks to Tobii Pro Lab's automapping program, reconstructs a triangular pattern: fixations transition between the dual images and outward beyond the left edge of the monitor (Figure 5).

Why would a pulmonologist look away from the monitor, especially in moments that demand heightened attention? What is taking place when the gaze in the model repeatedly shifts outward?

Following the model, the optical data are valuable for what they reveal as much as conceal. Although the heat map indicates the regions of the visual field where the pupils fixate, those regions also indicate sensory systems other than vision. Our inference is that the gaze would shift away from the monitor as a pulmonologist shifts her sensory emphasis from visual to tactile perception. She uses her own haptic impression of the bronchi to guide the needle through the cartilage and into the lesion. It's not so much the she looks at nothing. She feels something. She directs her attention to the feeling of the bronchial anatomy in her grip of the bronchoscope's handle.



Figure 5. Hypothetical model of heat map over EBUS monitor

More specifically, the eyes' trajectory aligns with the body's movement. The pulmonologist shifts her stance left. The hypothetical gaze pathway reconstructed in Figure 6 illustrates this spatial realignment. Gaze plots on the left mark where she would turn laterally, moving her body toward the left of the monitor. To be sure, my interpretation does not obtain on the basis of the gaze pathway alone. I also observed the pulmonologist in my role as ethnographer. She looked at the two-dimensional image in front of her while moving in the three-dimensional space around her. Both are perceptual skillsets essential to performing lung biopsies.



Figure 6. Hypothetical model of gaze pathway over EBUS monitor

What results is a multi-sensory circuit activating visual and tactile perception. Sliding her focus between either, the pulmonologist orients her own body via two spatial configurations: those of the room and of the patient's airways. In the model, my naked eye supplements the

eye-tracking data; both offer an occasion to pull apart the pulmonologist's entangled perceptual skillsets. In a follow-up interview, she elaborated that the skillset involves a combination of tactility and vision. What matters when initially performing bronchoscopy is one's ability to know how the device works and to know how one needs to move to make it go where it needs to go. One uses both the visual cues and the hand's movement. In short, it's seeing and feeling.

Touching and seeing flow together. Yet, eye-tracking evidence makes apparent that they do not always flow proportionately. The evidence illuminates when and where pulmonologists' deploy sophisticated kinesthetic competencies, shifting perceptual systems back and forth. Conventional ethnographic methods of observation and participant interviews fill in the blanks, as it were, and reveal the pathway to mastery behind the medical art. Another pulmonologist tells me that all of the incoming first-year fellows go through a two-hour session about how to hold and manipulate the bronchoscope, and how to watch the monitor. Everyone learns to use his or her control arm and support arm. And in the beginning, the fellows usually grip the device too hard because they're overly focused on the monitor. Vision shrouds feeling.

Pulmonologists develop supple techniques over time before they learn to collaborate with medical teams. The intricate skillsets at work in a lung biopsy is a strong reason why the procedure likely could not be carried out by machines. Eye tracking offers one tool to help ethnographers document these skillsets. The findings provide insights for people and organizations (e.g., device manufacturers, medical schools, and value analysis committees) to improve the ensemble of coordinated practices involved in bronchoscopy.

# CONCLUSION: PUSHING THE BOUNDS OF ETHNOGRAPHIC EYE TRACKING

Although challenging, adapting devices meant for recording individuals' vision of twodimensional planes to interpersonal and immersive contexts with three-dimensional visual interfaces is not only feasible; it's also useful for ethnographers. Workplace ethnography broaches many questions. Eye trackers are not dispositive. In fact, they're incomplete when used in insolation. But when tailored to address specific questions about practices of looking, feeling, and acting, the technology offers a launching pad from with to hone ethnographers' lines of inquiry. I hope to have offered one ensemble of methods and a modest model for doing so. We continue to refine our model at Design Science in order, eventually, to integrate eye trackers in actual medical facilities.

The contexts where ethnographic fieldwork takes place are quite unlike the cockpits in which Fitts and his team had originally studied pupil movement in 1947. They configured the cockpit along two planes while participants remained seated. Fast-paced workplaces are nothing of the sort. That ethnographers might be lucky enough to find people willing to wear eye trackers while going about their business is hardly given. The devices are cumbersome when worn by busy healthcare professionals communicating information and making quick decisions in medical facilities. Nevertheless, the questions that Fitts had posed half a century ago – namely, how does expertise inflect perception and action? – remain just as relevant today. Across collegial environments, peoples' disparate perceptual skillsets contribute to collaborative activities.

Eye trackers offer an additional perspective in such contexts. The devices visualize what escapes ethnographers' naked eyes. One of our aims has been to show how eye trackers could also be useful for their limitations, which expand opportunities to probe perspectives other than the eyes'. Clarity and opacity are both valuable sources of evidence. Ethnographers might thus delve into and analyze expert practices whose sensory modalities remain caught, as it were, in peoples' blind spots.

#### NOTES

Acknowledgments – An invaluable interlocutor throughout this article's revisions was Rebekah Park; I am also indebted to Christina Stefan for her editorial acumen and to Ranjan Nayyar for his persistent insights into human factors research.

1. See Hubert L. Dreyfus and Stuart E. Dreyfus, *Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer* (New York: Free Press, 1988).

2. "What Does a Pianist See?," *Fractal Media: Function.* Published Mar. 18, 2017. <u>https://www.youtube.com/</u>watch?v=GVvY8KfXXgE.

3. D.L. Mann, W. Spratford, and B. Abernethy B. "The Head Tracks and Gaze Predicts: How the World's Best Batters Hit a Ball," *PLoS ONE* 8, no. 3 (2013): e58289.

4. Consider Clifford Geertz's pithy yet canonical quip, "Culture is public because meaning is." *The Interpretation of Cultures* (New York: Basic Books, 1973), 12.

5. Charles Goodwin and Marjorie Goodwin, "Seeing as Situated Activity: Formulating Planes," in *Cognition and Communication at Work*, ed. Y. Engeström and D. Middleton (Cambridge: Cambridge University Press, 1996), 61-95.

6. Lucy Suchman, "Centers of Coordination: A Case and Some Themes," in *Discourse, Tools, and Reasoning: Essays on Situated Cognition*, ed. L.B. Resnick, R. Säljö, R., C. Pontecorvo, and B. Burge (Berlin: Springer Verlag, 1996), 41-62.

7. See E. Sumartojo, A. Dyer, J. Garcia Mendoza, and E. Gomez, "Ethnography through the digital eye: what do we see when we look?" in *Refiguring Techniques in Digital Visual Research*, eds. Edgar Gómez Cruz, Shanti Sumartojo, Sarah Pink (Switzerland: Springer, 2017), 67-80.

8. Paul Fitts, et al., "Eye movements of aircraft pilots during instrument-landing approaches," *Aeronautical Engineering Review* 9, no. 2 (1950): 24-29.

9. See U.S. Department of Health and Human Services: Food and Drug Administration, "Guidance on Medical Device Patient Labeling; Final Guidance for Industry and FDA Reviewers" (2001), page online: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm07080 1.pdf

10. Sense-data theory has its roots in the work of G.E. Moore; see *Some Main Problems of Philosophy* (London: George, Allen and Unwin, 1953).

11. One may, of course, be aware that the IFU has a reverse side. Sense-data theorists argue that one sees it indirectly – that is, in virtue of the side that one sees directly. See Frank Jackson, *Perception: A Representative Theory* (Cambridge: Cambridge University Press, 1977).

12. For the classical formulation, see Franz Brentano, *Psychologie vom empirischen Standpunkt* (Leipzig: Duncker & Humblot, 1874).

13. See Fred Dretske, *Perception, Knowledge and Belief* (Cambridge: Cambridge University Press, 2000); Jerry Fodor, *Representations* (Cambridge, MA: The MIT Press, 1981).

14. James J. Gibson, The Ecological Approach to Visual Perception (Hillsdale, NJ: Lawrence Earlbaum, 1979), 223.

15. Maurice Merleau-Ponty, The Primacy of Perception: and Other Essays on Phenomenological Psychology, the Philosophy of Art, History and Politics, trans. James Edie. (Evanston, IL: Northwestern University Press, 1964), 167.

16. Alva Noë, Action in Perception (Cambridge, MA: The MIT Press, 2004), 194.

17. Shane L. Rogers, et al. "Using dual eye tracking to uncover personal gaze patterns during social interaction," *Scientific Reports* 8, no. 4271 (2018). Page online. DOI: 10.1038/s41598-018-22726-7.

18. American Cancer Society, "Key Statistics for Lung Cancer." Page revised Jan. 4, 2018. https://www.cancer.org/cancer/non-small-cell-lung-cancer/about/key-statistics.html

19. Hubert L. Dreyfus, "Overcoming the Myth of the Mental," Topoi 25 (2006): 47.

20. A procedural guide distinguishes the airways' anatomy according to the tactile impression transmitted via the bronchoscope. For example, "The posterior part of the trachea is vertically structured (resembling long vertical plies) and easy to distinguish from the arch-shaped cartilage structures (the horizontal support of the bronchial tree)." See, Felix Hirth, et al. "Endobronchial Ultrasound-guided Transbronchial Needle Aspiration." *Journal of Bronchology* 13, no. 2 (2006), 13.

21. Maurice Merleau-Ponty, Phenomenology of Perception, trans. Colin Smith (London: Routledge, 1962), 239.