



Caught in Eye Trackers' Blind Spots: Adapting Vision Studies to Ethnographic Field Research

Larry S. McGrath^(✉), Lindsay A. Carrabine^(✉),
and Ranjan Nayyar^(✉)

Design Science Consulting, Inc., Philadelphia, USA
{larry.mcgrath, lindsay.carrabine,
ranjan.nayyar}@dsconsulting.com

Abstract. Although eye tracking has been used to analyze user interfaces, the technology poses numerous challenges when applied to environments in which people move, work, and interact. Aren't eye trackers limited to studying two-dimensional planes? How might the device help to understand activities that are not exclusively visual? These questions are integral to ethnography, the observational study of peoples' interactions with products, processes, and places. We offer answers by way of two bodies of research. The first was conducted to optimize Instructions for Use (IFUs) for medical devices. The second was an exploratory study conducted to adapt eye trackers to manual tasks. The objective was to cultivate interpretive methods for evaluating visual and non-visual perception in eye-tracking data (i.e., fixation duration, gaze plots, heat maps). We suggest that the latter body of research illuminates new opportunities for ethnographic fieldwork to examine how medical professionals move their bodies when simultaneously carrying out procedures and viewing visual systems.

Keywords: Ethnography · Vision · Medical devices · User experience · Design human factors · Instructions for Use · Hospitals

1 Introduction

Technological advancements in eye-tracking technology has contributed to widespread vision studies of user interfaces, particularly in human factors research. At Design Science, we use mobile eye trackers to support the design of Instructions for Use (IFUs). Medical devices and products are sold with IFUs, which inform users and lay caregivers about proper uses, risks, and benefits. Eye trackers generated insights into how people actually use a given set of instructions—revealing what the users see, what they skip, and where they stumble. We have used the results to produce intuitive and readable IFUs.

We are now revamping our eye-tracking methods in an effort to incorporate the device into our ethnographic field research in hospitals and clinics. Many researchers, however, have been slow to use eye trackers as a part of their toolkit because questions linger about the utility, methodology, and value of conducting visual research in medical environments where people move, work, and interact. Why would devices

designed to record superficial visual interfaces be useful in three-dimensional contexts? Given that medical professionals employ multiple perceptual systems when interacting with devices, how could eye tracking be adapted to examine modes of perception other than vision? Answers remain fleeting. As a result, ethnographic research gets lost in the blind spots of eye tracking.

In this paper, we show that this needn't be the case. Eye trackers stand to enrich ethnographic inquiry by empowering researchers to examine peoples' visual, tactile, and spatial interactions with medical devices.

Our goal is to outline interpretive methods to make eye tracking impactful in immersive contexts where users don't make use of visual perception alone. We do so by drawing on two bodies of research carried out at Design Science, in which researchers reconfigured vision studies for IFU optimization into experimental ethnographic models. Along the way, we found that the constraints of eye trackers in conventional visual-interface research generate criteria for improving ethnographic fieldwork.

The first body of research on IFUs identified what users read (and didn't) when performing tasks with medical devices. Eye trackers buoyed our reiterative research process, which serves to compare multiple versions of IFUs. Eye-tracking data made it easier for us to answer usability questions, such as:

- Which aspects of an IFU lead to use errors and difficulties?
- What elements of an IFU stand out to users?
- Do users rely more on images or text to comprehend an IFU?
- When confusion arises, is it due to the IFU or to the device?

Our human factors engineers and graphic designers used the results to eliminate design flaws and to reorganize IFUs' visual architecture for intuitive reading. We found, however, that the studies were suited to scenarios in which users primarily relied on vision to navigate superficial interfaces. Ethnography presents different parameters. In an effort to understand immersive contexts in which users do not only look *at* a visual field, but also interact with objects *in* the visual field, our researchers sought to devise alternative methods for ethnographic observation.

The second body of research is of users' interactions with various devices using visual and haptic feedback. In our study, tasks were designed to emulate interventional procedures in which, for example, pulmonologists or neurosurgeons engage with patients' bodies directly via medical devices (i.e., bronchoscopes and spinal screw-drivers) and indirectly via visual interfaces (i.e., bronchoscopy and fluoroscopy). By tracking participants' pupil movement while they performed manual tasks and observed a visual interface, our researchers tackled questions that are crucial to using eye trackers in future ethnographic research, such as:

- How is pupil movement indicative of users' spatial and tactile navigation of medical spaces?
- At which point do users shift their concentration from visual to tactile perception?
- How could users strike the right balance between tactile and visual engagements with medical devices?
- Which aspects of devices induce users to rely on perceptual systems other than vision?

The results demonstrated that tasks involving continuous physical exertion prompt users to diminish visual perception and instead to rely on tactile perception. By contrast, tasks that involve spatial coordination induce users to rely more on visual perception. Moreover, eye-tracking data can be interpreted to reveal both modes of perception. We suggest that distinct fixation patterns are associated with continuous physical exertion and spatial coordination. These results will empower our ethnographic researchers to use eye trackers in order to explore medical professionals’ embodied interactions with interventional procedures for the sake of generating device design recommendations.

For both bodies of research, we used the Tobii Pro Glasses 2. This mobile eye tracker rests on the nose and ears like normal glasses. The glasses weigh 45 g and feature four sensors, which emit infrared light to the retinas at a 100-Hz sampling rate. A high-definition scene camera is attached to the bridge and directed outward to the visual field.

2 Visual Studies for Optimizing IFUs

2.1 Conventional Eye-Tracking Metrics vs. Customized Analysis

Usability testing is critical to creating effective IFUs. Observing users as they read and carry out simulated tasks helps to optimize the layout. By adding eye-tracking technology to traditional usability testing, we followed participants’ optical behaviors using conventional metrics as well as our customized analyses (see Table 1). The results combine quantitative reports and qualitative insights to facilitate the design of intuitive and readable IFUs.

Table 1. Eye-tracking data

	Conventional metrics	Custom analysis
Data types	<ul style="list-style-type: none">• Heat maps• Gaze plots• Fixation duration	<ul style="list-style-type: none">• Parallel scan path visualizations• Reading-rate comparison• Task completion association
Key insights	<ul style="list-style-type: none">• Length of time participants observe IFU elements• Order in which participants navigate IFU content	<ul style="list-style-type: none">• Relation of use errors to gaze patterns• Relation of gaze patterns to IFU layout• Relation of reading rate (i.e., skimming vs. reading thoroughly) to IFU content

Out-of-the-box eye tracking generates gaze plots and heat maps, which visualize the IFU areas observed by users (see Figs. 1 and 2). These conventional metrics convey helpful information, yet they fail to provide a full context with meaningful clues to help optimize IFUs (see Table 2).

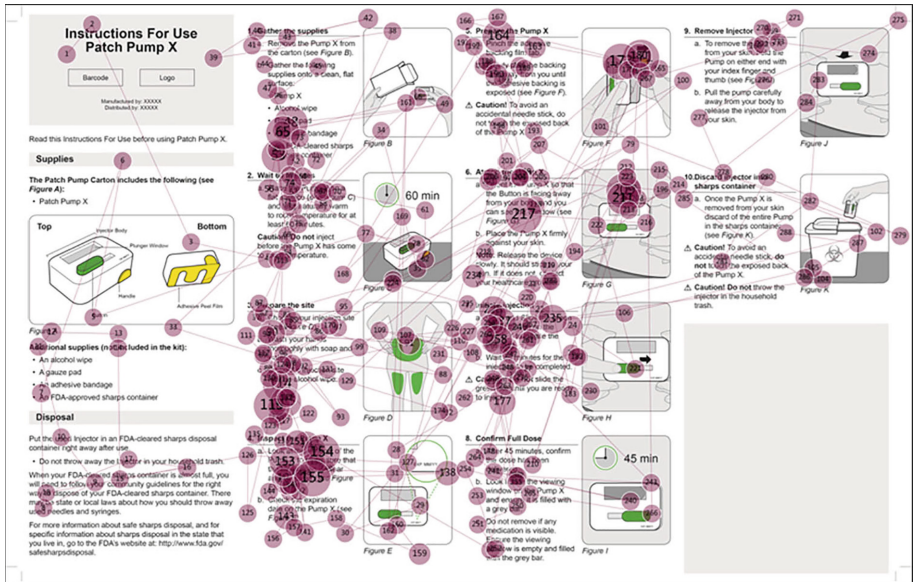


Fig. 1. Gaze plot

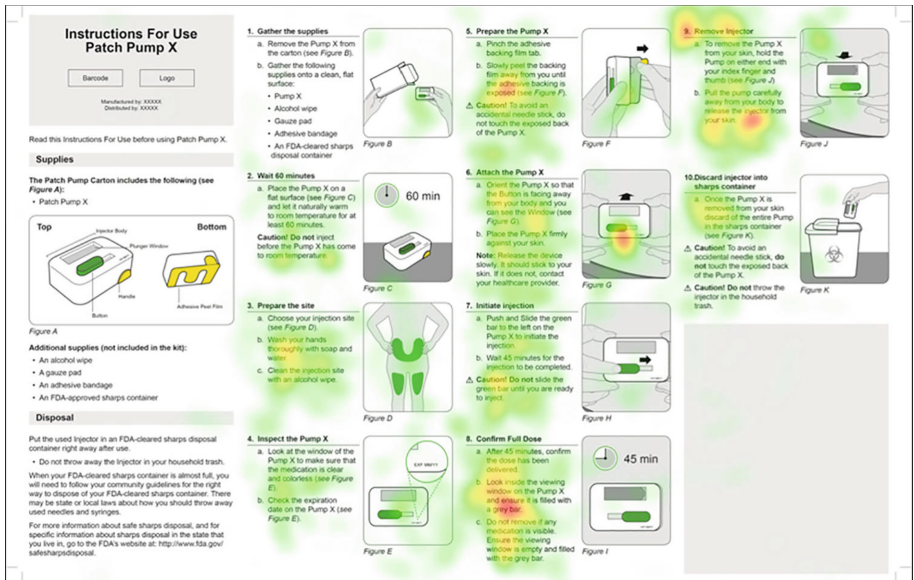


Fig. 2. Heat map

Table 2. Conventional eye-tracking metrics

	Key insights	What’s missing?
Gaze plots	Pathway of participants’ gaze across IFU content	Gaze plots are hard to interpret due to information overload
Heat maps	Degree of participants’ fixations on different IFU content indicative of which content attracted participants’ attention	No information about why the participant fixated on different areas of the IFU

Through our process of analyzing the eye-tracking data and uncovering the unmet needs of the conventional eye-tracking metrics, we created easy-to-read scan path visualizations (see Fig. 3). Effective IFUs guide participants to follow a determined operational order. Our visualizations illustrate the path of participants’ gazes: whether they skip steps, jump between text and images, or return to various IFU content areas.

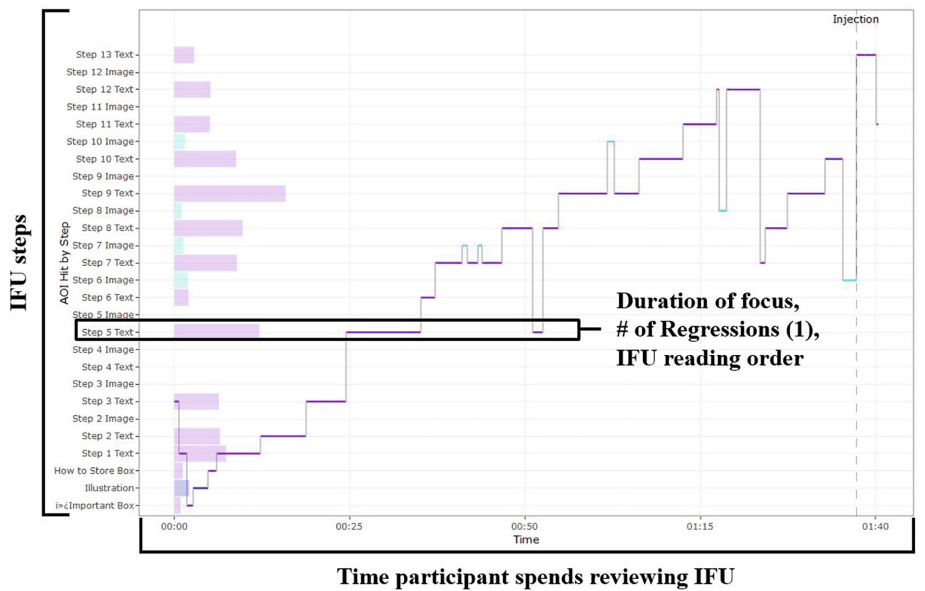


Fig. 3. Scan path visualization

2.2 Results

One of the challenges that confronts researchers conducting eye-tracking analysis is choosing and making sense of the myriad eye-tracking data. For a general understanding of participants’ use of an IFU, heat maps and gaze plots are useful because they are easy to generate and interpret. However, when the aim is to diagnose a recurrent issue, such as when participants make a mistake at a specific instructional step, it can be useful to look at scan path visualizations, fixation metrics, and raw video from the eye tracker.

Upon review of the customized scan path visualizations, we generated concrete design recommendations to improve the IFU (see Table 3).

Table 3. Examples of IFU design recommendations

Gaze pattern behavior	Usability result/feedback	IFU design implication
<ul style="list-style-type: none"> • User skims over Steps 10–13 • User does not observe images at Steps 10–13 	Use difficulty at Step 11: <ul style="list-style-type: none"> • User is confused by needle guard and assumes something has gone wrong 	<ul style="list-style-type: none"> • Highlight critical tasks in Step 10–13 • Improve illustration to highlight specific device element
<ul style="list-style-type: none"> • User skips step 4 (image and text) while reading IFU 	Use error at Step 4: <ul style="list-style-type: none"> • User does not mention anything about waiting 30 min before giving the injection 	<ul style="list-style-type: none"> • Improve layout and bring location of Step 4 from top of the page closer to the center • Highlight critical information at Step 4 • Repeat critical information located in Step 4 (wait 30 min before injecting) at Steps 6, 7

3 Visual Studies for Embodied Research

Although eye trackers generated insightful data for optimizing IFUs, we found that different considerations were in order when conducting vision studies with users who manually engage with devices. In such contexts, users do not rely exclusively on sight. Indeed, rarely do people use only their eyes when navigating everyday spaces. How, then, might eye trackers account for visual perception as well as sensorimotor modes of perception? We sought to answer the question by creating an experimental vision study in which eye-tracking data were used to track not only the pupils' movement but that of the body as well.

We observe these contexts in Design Science's field research. For example, neurosurgeons conduct spinal fusions by watching fluoroscopic images to orient the manual placement of pedicle screws. In so doing, neurosurgeons navigate multiple modes of perception: both the fluoroscopic imaging and the tactile impression of the screw driver in patients' vertebrae. We have found that neurosurgeons enjoy direct haptic feedback of the vertebrae, though only indirect visual feedback from the imaging. We have also found that pulmonologists encounter a similar perceptual asymmetry. When conducting lung biopsies, for instance, pulmonologists manipulate a bronchoscope through patients' airways and enjoy direct haptic feedback via the device, yet indirect visual feedback from bronchoscopic visualizations on a screen. Both scenarios are ripe for eye tracking. Neurosurgeons, pulmonologists, and potentially other medical professionals stand to benefit from a better understanding of the intimate and imbricated interactions between visual and tactile perception. This is especially the case for younger medical professionals, who have yet to develop the

delicate sense of balance when manipulating devices in patients' bodies and simultaneously watching images of patients' bodies.

Designing an experimental eye-tracking study to emulate medical professionals' perceptual asymmetry poses challenges. In lived experience, vision seamlessly complements tactility. Isolating distinct modes of perception is not straightforward. Instead of isolating either perceptual mode, we interpreted eye-tracking data to be indicative of both bodily and ocular activities. The pupils' movement reflects not only the eyes, but also an ensemble of embodied activities such as holding the head in place, shifting weight, and manipulating the hands [1]. We accounted for these activities by approaching eye-tracking data (notably, fixation durations, heat maps, and gaze plots) as indices of tactile perception. The gaze is one aspect of embodied movement.

3.1 Experimental Setup

For the study, 14 participants performed simple manual tasks with devices hidden under a box. Although the participants could not look directly inside the box, they could see indirectly via a video feed projected onto a television screen (see Fig. 4). This served as the visual interface. Akin to the parameters of interventional procedures mentioned above, participants encountered an asymmetry between visual and tactile perception.



Fig. 4. Study setup for Scenarios 1 and 2

In the first scenario, participants were presented with a screwdriver and wood block with eight holes, four of which had Philips-head screws. The moderator asked the participants to un-screw and re-screw (in a different hole) each of the screws.

Scenario 1 involved two distinct kinds of manual activities: spatial coordination and continuous physical exertion. In the image below, a participant performs the former. She coordinates her hands in space, aligning them with a screw (Fig. 5).



Fig. 5. Scenario 1 spatial coordination

For the second kind of manual activity, participants exercised exertive force to unscrew and re-screw using the screwdriver (see Fig. 6).



Fig. 6. Scenario 1 physical exertion

In the second scenario, participants were presented with a pill organizer. It had eight compartments, each of which contained three pills (see Fig. 7).

The moderator prompted participants to open select compartments, remove a specified number of pills, and arrange them in two rows along dotted lines. Unlike Scenario 1, the second scenario emphasized one manual activity: spatial coordination. Participants identified, selected, extracted, and arranged small objects (see Fig. 8). There was little expectation to exert sustained physical force.



Fig. 7. Scenario 2

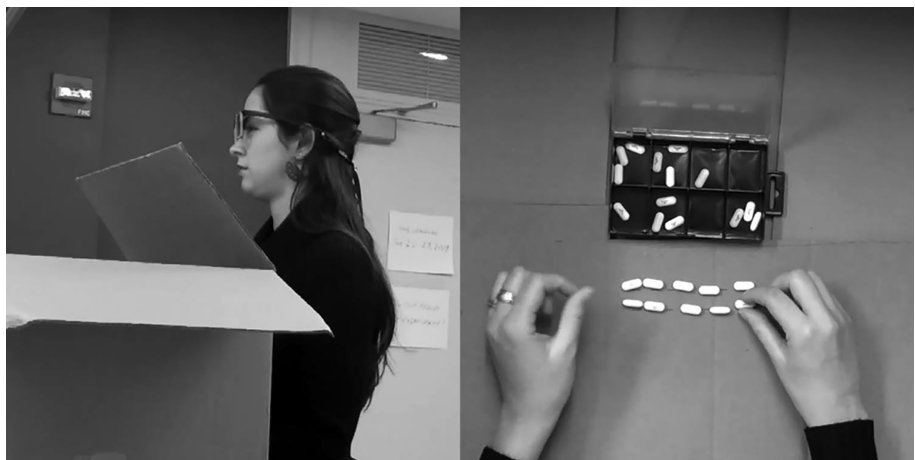


Fig. 8. Scenario 2 spatial coordination

Given the divisions between vision and tactility are porous in lived experience, we interpreted visual and tactile perception as heuristic categories. Neither are pure. Our focus was the moments at which participants shifted their attention from visual to non-visual perception: namely, when they looked away from the screen to perform a manual task. By having participants perform distinct kinds of manual tasks (i.e., coordinative and exertive), we compared tactile and visual perception relatively. Neither was wholly tactile nor visual; they differed according to their respective degrees of emphasis.

3.2 Results

Our results revealed that participants in Scenario 2 spent 84.65% of their time, on average, fixating on the screen. By contrast, the same participants in Scenario 1 spent 69.27% of their time, on average, fixating on the screen. Moreover, the difference was statistically significant; every participant spent more time looking away from the screen during Scenario 1 ($p=.006$ according to a two-tailed t-test).

Closer examination of the eye-tracking results reveals distinct fixation patterns associated with the Scenarios. Each involved different manual activities: spatial coordination and continuous physical exertion. When prompted to locate objects, participants directed their gaze at the screen in order to gather spatial information and to coordinate their hands. By contrast, participants tended to direct their gaze away from the screen when performing manual tasks involving intensive exertion.

The fixation pattern associated with spatial coordination was pronounced in Scenario 2. Participants identified, selected, and extracted pills from the small compartments of a pill organizer. Those participants' fixation patterns were confined to a narrow region on the screen around the visualization of manual activity. This is likely because the region offered the spatial information required to execute tasks.

A divergent fixation pattern associated with continuous physical exertion was peculiar to Scenario 1. Like Scenario 2, participants identified and selected screws, while the participants also observed the screen in order to gather spatial information. Once participants anchored a screw in place, however, they proceeded to apply exertive force to the screwdriver and shift their gaze beyond the screen. Moreover, the region to which participants directed their gaze remained consistent for each individual (though not across individuals).

What is the nature of this fixation pattern? Its purpose, we hypothesized, was not primarily visual. Participants in Scenario 1 did not aim to collect spatial information but instead to facilitate embodied effort. Just how the fixation pattern helped them in doing so demanded explanation.

3.3 Eye-Tracking Visualizations

Heat maps and gaze plots help to clarify the distinct perceptual modes involved in each fixation pattern. These visualizations represent the distribution of fixations across the visual field. Heat maps illustrate the concentration of fixations. Gaze plots illustrate their pathway. Together, the visualizations allowed us to make sense of the characteristics unique to each fixation pattern.

Participant 9 was exemplary. She fixated on the screen for 88% of the time during Scenario 2, which involved more coordination, and for only 70% of the time during Scenario 1, which involved more exertion.

The gaze plot of the participant's pupil movement in Scenario 2 illustrated that her gaze centered on the region of manual action in which she identified, located, and extracted pills from compartments of a pill organizer. This region offered ample visual resources for the coordination of the participant's manual tasks. The gaze plot of the same participant's pupil movement in Scenario 1 illustrated that her gaze repeatedly shifted above the screen and slightly to the left (see Figs. 9 and 10).

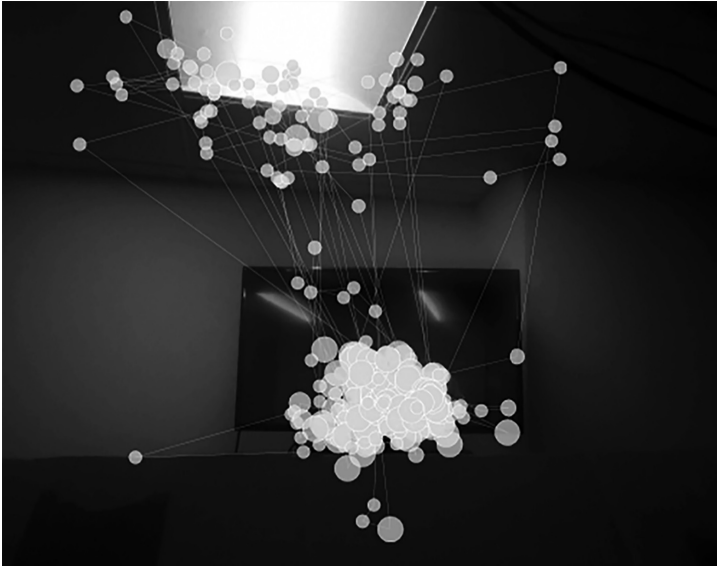


Fig. 9. Scenario 1 gaze plot



Fig. 10. Scenario 2 gaze plot

Whereas the gaze plot of Scenario 2 illustrates constrained pupil movement around the region of manual action, the gaze plot of Scenario 1 illustrates a mobile fixation pattern between two regions: the region of manual action on the screen as well as a relatively consistent region above the screen and to the top left. The latter region coincides with moments when the participant exerted force to manipulate the screw-driver. Embodied strain, and minimal spatial coordination, was involved.

Moreover, the fixation pattern associated with continuous physical exertion did not occupy a stable point (see Fig. 9). Pupil movements transitioned back-and-forth between the region of manual activity (the screw) and a region beyond the screen. The fixation pattern, therefore, involved a mobile circuit of visual perception.

3.4 Discussion

We interpret these divergent fixation patterns to represent a shift in emphasis between competing modes of perception: visual and tactile. Fixating on the screen was associated with visual perception, which participants emphasized to draw spatial information about the location of small objects. Shifting the gaze away from the screen was associated with tactile perception, which participants emphasized to concentrate on the screwdriver's haptic feedback. The latter scenario is representative of those we encounter in ethnographic observations of interventional medicine. For example, when straining to screw in a pedicle screw during spinal fusion procedures, neurosurgeons look away from key visual interfaces – such as fluoroscopic images – in order to concentrate on the task's tactile demands. Novice neurosurgeons stand to benefit from understanding the delicate balance struck by expert neurosurgeons when navigating overlapping perceptual systems.

Moreover, we found that the distinctive fixation pattern exhibited in Scenario 2 reflects participants' "attentional anchors" [2]. These are emergent (and not pre-given) constructs of vision. They offer a visual pivot point, as it were, upon which people exert perceptual leverage in order to concentrate on the exertive demands of manual tasks. The selection and formation of attentional anchors—that is, the range of space beyond the screen where participants fixated—was unique for each participant. Nonetheless, patterns emerged. In 9 of 11 cases (for which complete gaze plots were generated), participants' range was positioned on the side of the visual field opposite their dominant hand (i.e., right-handed participants tended to shift their gaze to the left of the screen). Numerous interpretations are possible. Because attentional anchoring secured embodied leverage, the associated fixation pattern may have facilitated the centrifugal force with which participants drove the screwdriver's handle. Alternatively, because the attentional anchor reflected repeated sensorimotor activity; its selection and location may have aligned with the diagonal movement of participants' bodies when applying continuous exertion to the screwdriver. In either case, the fact that fixation patterns tended to intersect with a region of the visual field opposite of participants' dominant hands suggests that the eyes' movement derived from the body's. Tactile perception lead; visual perception followed.

4 Conclusion

Different contexts call for different approaches to eye tracking. Not all usability challenges present the same parameters for vision studies. By comparing the distinctive aspects of two bodies of research, we hope to have shown that eye tracking is a malleable tool adaptable to diverse usability contexts. When examining IFUs, eye trackers facilitate a reiterative approach that optimizes their arrangement of information

and visual architecture. The result is more intuitive instructions for the sake of effective medical device use. When it comes to the manual practices by which people interact with devices, eye tracking can—and should—be interpreted for modes of perception that are not exclusively visual. By interpreting eye-tracking data as an index of tactile perception, ethnographic researchers can better adapt vision studies to the immersive contexts in which interactions between users and devices takes place. The results, we hope, will illuminate what has remained in eye trackers' *blind spots*: the embodied skillsets with which people exert force and coordinate movement when performing manual tasks.

References

1. Solman, G., Foulsham, T., Kingstone, A.: Eye and head movements are complementary in visual selection. *R. Soc. Open Sci.* **4**, 160569 (2017)
2. Abrahamson, D., Bakker, A.: Making sense of movement in embodied design for mathematics learning. *Cogn. Res.: Princ. Implic.* **1**(1), 33 (2016)