Abstract

This essay draws from capabilities theory to argue that visual literacy, which combines technical skills, knowledge, and ethics, is an essential human capability in the digital age. It builds on the idea of democratic voice to posit that individuals need to control the way they are seen and to visually account for their lives in order to achieve a balanced system of “surveillance.” Human beings are surveilled at unprecedented rates in the digital age. In conversation with Gates, who describes the way forensic specialists are learning to mine visual archives to craft coherent crime-solving narratives; Ristovska, who points out that court officials remain overly wedded to logocentric logic; and Spiesel, who reminds us of the dangers of naive realism, this essay focuses on the right of those who are watched to craft their own sousveillance while also understanding the power, limits, and ethical implications of visual communication.

Introduction

Discussions about surveillance video usually center around the tension between its evidentiary potential and privacy concerns. The sheer enormity of the image data available in the digital age disrupts this binary and, as the essays in this Dialogue Section evince, forces us to contend more carefully with the nature of the surveillance image itself: how it is made, rendered meaningful, and used in civil society. Gates (in this issue) describes the extensive work necessary by forensic specialists who sift through visual information in order to craft narratives useful to a criminal justice investigation. Ristovska (in this issue) argues that courts continue to impose a logocentric paradigm on images as evidence, rightly encouraging legal operatives to consider the different ways that images convey information. Spiesel (in this issue) underscores the persistence—and harmful impact—of naive realism and the way that it can too easily put great power into the hands of a few.

This essay also speaks to the problems of naive realism, logocentric approaches to evidence, and the way surveillance images, indeed any images, are incorporated into narrative, and argues that today’s media environment presents an urgent need for individuals to be able to participate meaningful in the visual sphere. Thinking about the way citizens “count,” in a democratic system, often conceptualized as having “voice,” takes on new dimensions in this hypermediated, and hyper-visual age. Today, voice is a matter of being heard and seen. This requires tangible and educational access to technology as well as the freedom to speak to power, monitor one’s surroundings, and account for one’s life. In what follows, I will start by reviewing the notion of democratic voice before exploring its requirements in terms of skills and technological affordances in support of my larger thesis, which posits that, in our hyper-surveilled world, individuals have an essential human right to use the same tools as their surveillors: to be heard and seen as part of a mutually responsible, ethical media system.
**Voice, Witnessing, and New Visibilities**

The democratic public sphere, as conceived by Habermas (1989), requires that citizens have a voice that is heard, considered, and valued (see also White 1995). The US and other industrialized democracies have yet to fully realize this ideal, yet it remains an idea for global human rights efforts. To have a voice is more than the freedom to express oneself; an individual must be both recognized and respected. Couldry (2010: 45) defined this form of voice as “our capacity to make, and be recognized as making, narratives about our lives and the world within which we act.” Further, he argued, voice has two dimensions. It is a process—a social, reflexive, and material activity—but for it to matter democratically, it must also be valued (Couldry 2010). Here “value” has many dimensions, some of them normative, practical, and moral. On paper, everyone counts, and yet patriarchy, racism, and class permeate public debate and neoliberal media policies have exacerbated social inequality. A person might have a voice in one context but be ignored in another; the voices of certain groups may not be considered morally worthy.

Visibility is another dimension of voice, one that civil rights activists have long used strategically in order to garner public support. Marches and protests are often constructed as street spectacle (Neumayer and Rossi 2018). Martin Luther King and his allies were strategic in their relationships with photojournalists from *Life* magazine and other outlets that published images of black oppression—a move believed to have made a significant, positive impact on white support for the movement (Gallagher and Zagacki 2007). Images are most powerful when combined with the moral weight of testimony, that is, when the images serve as evidence (Azoulay 2008; Peters 2001). While it is impossible to determine just how much influence an image might have in terms of public action, mediated witnessing can change the way the public understands a situation, and certain images become indelibly part of civic discourse (Azoulay 2008; Hariman and Lucaites 2007).

Today, such mediated visual testimony is possible for anyone with a smartphone. As Mann and Ferenbok (2013: 22) wrote, mobile video and wearable cameras provide users with memory aids and *proof*. Black Americans, who have complained about police abuse for decades, have been able to show the world examples of violence that would not have been believed years ago. Who would have believed an airline would pull a doctor from his seat and drag him off of an airplane (Victor and Drew 2017) without video? Absent a visual record, complaints about violence and abuse in a Florida prison might have gone nowhere, but a smuggled phone proved to the public that conditions were inhumane (Ellenbogen 2019). Whether as organized “cop-watching” or spontaneous recording, individuals at every level of society have used the evidentiary power of video to voice concerns about the conditions around them.

The power to film and to show, or “sousveillance,” presents a counterbalance to the power of the state and its corporate superstructure. Mann and Ferenbok (2013) used the word “veillance” to describe the way power is exercised by views from above and below. In the US, several federal appeals courts have upheld the right of everyday individuals to shoot video of public activity in public places, based on the First Amendment (Regan 2015). The Supreme Court has not yet ruled on the subject, and related issues are unsettled. For instance, while an individual may have the right to film, do they also have the right to upload video of another person in distress to the web? What about the right to film in non-public places when there is a public interest, as with the smuggled prison phone? Voice might be inextricably bound with visibility, yet the ethical and legal issues surrounding image-based communication remain underdeveloped.

**Digital Participation as Capability**

If being able to be seen and heard is a human right, it follows that individuals need to be capable of participating in the digital public sphere; they need access to technology, yes, but also audio-visual skills, ethical literacy, and policies that protect individual interests. This means more than narrowing the digital divide, as this capability is not merely a matter of giving everyone a smartphone and Wi-Fi access. To consider digital participation an essential human capability means prioritizing it in a larger social structure that balances individual and social interests and pursuing the development of new legal and ethical norms.
The capability approach originated with economist Amartya Sen (2005), who argued for a model of national development that considered more than gross domestic product (GDP) and that examined how many citizens had power over their own lives with access to education, health care, and a clean environment. Nussbaum (2000) extended Sen’s theory with a focus on women’s rights by specifying key “capabilities” that allow individuals to live well, including the ability to control one’s health and body, exercise one’s imagination and thought (namely, pursue education), and to have control over one’s material and political environment (see also Terzi 2014). The capabilities approach is not a media theory, yet it offers a useful way to consider democratic voice and visibility in the digital age.

If voice is both a matter of word and image in the digital age, then it follows that participation in multimedia communication is an emerging human capability. Those left out of the digital public sphere are denied the ability to account for their lives—in word and image—to the rest of the polity. Visual communication has become an essential component of democratic discourse online.

**Ethical Visual Discourse**

As a capability, audio-visual participation in the public sphere requires not just access to technology but a type of media literacy bound to ethical principles. Hobbs (2017: 15; emphasis added) defined her subject as “a process that involves accessing, analyzing, creating, reflecting and taking action.” Many recent media literacy campaigns focus on how to assess the credibility of messages, and, while this is important, an understanding of ethical participation is equally essential, especially in the case of images. After all, photos and video are powerful not only as evidence in human “tellings.” They are more memorable than words and pack more emotional punch (Barry 1997). In the courts, images are generally used in support of verbal testimony, and the legal community continues to struggle with the best way to fairly incorporate visual communication (Feigenson and Spiesel 2011). Even journalism remains logocentric, treating images carelessly and relegating photographers to lesser roles in the newsroom (Bock, Lough, and Fadnis 2017). As a result, the contemporary agora is everything Orwell envisioned and more, as trolls, bots, and other propagandists use images to effectively advance dis- and misinformation (Phillips 2015).

Therefore, camera rights are only part of the capability to participate in the democratic digital sphere. Whether in the form of state panopticism, organized sousveillance, or the everyday “accounting for one’s life,” visual communication permeates the digital public sphere, but a normative understanding of visual ethics in democratic discourse is largely inchoate. Most media literacy campaigns today focus on helping the audience critically analyse messages, not on the ethical crafting of user-generated messages. The media is no longer an institution apart from the users; we are the media and the media are us.

**Conclusion**

Each of the essayists in this Dialogue Section is concerned about the way surveillance imagery is rendered meaningful and used to solve crimes, advance human rights, or construct policy. I’ve argued that, in order to maintain balance in the visual sphere, all individuals need to pay heed to Gates, Ristovska, and Spiesel, to reject naive realism and learn the way images are created, narrativized, and rendered meaningful. In the digital environment, voice and visibility are tightly entwined. Human beings need access to the technologies of digital participation and media literacy that cultivate an ecology of mutual responsibility in word and image. To apply Azoulay’s (2008) phrase, photography has emerged as part of the “civil contract.” If visibility is to be tied with voice as an essential human right, then visual communication needs to be woven into larger discussions of ethics in democratic speech.

In the end, the camera is not, and never has been, the answer. If there is to be an era of balanced “veillance,” what matters most is how humans use cameras, what they say with images, and how they use visual evidence to account for their lives.
References