The Gestural Image:
The Selfie, Photography Theory, and Kinesthetic Sociability

PAUL FROSH
The Hebrew University of Jerusalem, Israel

Keywords: selfies, digital image, photography, visual culture, embodiment, social media

A selfie, whatever else it might be, is usually a photograph: a pictorial image produced by a camera. This banal observation informs widespread understandings of the selfie as a cultural category: “A photograph that one has taken of oneself” (Oxford Dictionaries Word of the Year, 2013, p. 1). Yet despite the selfie’s obvious photographic provenance, little scholarly research has drawn systematically on the intellectual resource most closely associated with the aesthetics of its host medium: photography theory.

In a way, this is unsurprising. The selfie is the progeny of digital networks. Its distinctiveness from older forms of self-depiction seems to derive from nonrepresentational changes: innovations in distribution, storage, and metadata that are not directly concerned with the production or aesthetic design of images. As the 2013 Oxford Dictionaries definition continues, the selfie is typically “taken with a smartphone or webcam and uploaded to a social media website.” It is these innovations that are thought to distinguish the selfie:

Instantaneous distribution of an image via Instagram and similar social networks is what makes the phenomenon of the selfie significantly different from its earlier photographic precursors. (Tifentale, 2014, p. 11)

Recent research on transformations in personal photography, in general, has been largely provoked by these nonrepresentational developments (Lasén & Gómez Cruz, 2009; Lister, 2013; Nightingale, 2007): Their influence has even prompted Gómez Cruz and Meyer (2012) to proclaim the arrival of a new “fifth moment” in photography history. And where aesthetic developments are foregrounded, they too appear to be driven by device functions not principally concerned with image production or design. Hence the immediacy, ephemerality, and incessant performativity of contemporary everyday photographs are primarily explained with reference to the combined ubiquity, mobility, and connectivity of smartphone devices (Murray, 2008; Van Dijck, 2008; Van House, 2011).

Once nonrepresentational technological changes are made analytically preeminent, what role remains for an aesthetically oriented and medium-specific intellectual tradition like photography theory? In

1 The author would like to thank Zohar Kampf, Ifat Maoz, Amit Pinchevski, Limor Shifman, and Julia Sonnevend for their readings of an earlier version of this article.

Copyright © 2015 (Paul Frosh, paul.frosh@mail.huji.ac.il). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at http://ijoc.org.
fact, the recent prominence of nonrepresentational practices echoes a recurrent tension in photography theory that has long divided scholars: an ontological commitment to the (largely semiotic) "essence" of the medium, which tends to privilege the discrete photographic image as an object of aesthetic analysis, versus historical conceptualizations of photography as a fluctuating constellation of devices, material cultural practices, and representational forms (Batchen, 1997). This tension between aesthetic object and sociotechnical practice is even evident in two of the adjective-noun compounds recently coined in contemporary photography theory itself: the networked image (Rubinstein & Sluis, 2008) and algorithmic photography (Uricchio, 2011).

The selfie affords a productive vantage point on this tension. On the one hand, it appears to constitute an aesthetic and representational innovation in everyday photography, potentially offering a degree of resistance to nonrepresentational emphases (similar claims could be made for popular applications like Instagram and Hipstamatic). Moreover, as a photographic genre, it invites attention to the pictorial conventions underpinning generic identity: After all, one cannot recognize an image as a selfie without looking at what it represents. Yet on the other hand, as with genre more broadly, representational criteria alone are insufficient (Mittell, 2001). Understanding that a particular image is a selfie (rather than just a photograph of, say, a face) requires viewers to make inferences about the nondepictive technocultural conditions in which the image was made (Frosh, 2001). It requires, among other things, that these viewers have been adequately socialized through having seen, taken, or heard tell of selfies.

The selfie thus foregrounds representational change while putting the very term "representation" in quotation marks as a contingent variable. This makes it a timely "theoretical object" (Verhoeff, 2009), a concrete phenomenon that is good to think with about contemporary photography and representation in general. The selfie prompts us to ask how it can be explained using concepts fashioned to illuminate the traditional aesthetics of photography and how it might reconfigure those concepts to forge new directions for theorizing both photography and digital culture.

In what follows, then, I propose to animate the conceptual fecundity of the selfie by engineering some brief encounters with a number of terms in photography theory: indexicality, composition, and reflection. Together these encounters will weave an argument—perhaps surprising, given its emergence from a visually oriented body of thought—that the selfie is a "gestural image" and that we should not understand its aesthetics purely in visual terms. Rather, selfies conspicuously integrate still images into a technocultural circuit of corporeal social energy that I will call kinesthetic sociability. This circuit connects the bodies of individuals, their mobility through physical and informational spaces, and the micro-bodily hand and eye movements they use to operate digital interfaces.

The need for at least basic representational criteria is evident from the sampling design of projects such as "Selfiecity" (http://selfiecity.net/#dataset): "To locate selfies photos, we randomly selected 120,000 photos (20,000–30,000 photos per city) from a total of 656,000 images we collected on Instagram. 2–4 Amazon’s Mechanical Turk workers tagged each photo. For these, we asked Mechanical Turk workers the simple question ‘Does this photo show a single selfie?”

2 The need for at least basic representational criteria is evident from the sampling design of projects such as "Selfiecity" (http://selfiecity.net/#dataset): “To locate selfies photos, we randomly selected 120,000 photos (20,000–30,000 photos per city) from a total of 656,000 images we collected on Instagram. 2–4 Amazon’s Mechanical Turk workers tagged each photo. For these, we asked Mechanical Turk workers the simple question ‘Does this photo show a single selfie?’”
Indexicality

Indexicality is the conceptual bedrock of traditional photography theory. Based on Peirce’s notion of the index as a sign that stands for its object through physical or causal connection, it designates the sense that photography is distinctive because what it depicts must have been located in front of the camera at the moment the photograph was taken. The photograph is described as an “emanation” (Barthes, 2000) of the referent, or a “quotation” (Sontag, 1977) from reality, since it is produced by light-sensitive material reacting to the light reflected from the spatiotemporal field exposed before the lens.

The supposed loss of photographic indexicality in a putative “postphotographic” era of digital image simulation was loudly debated in the 1990s (Mitchell, 1992; Robins, 1996). Recent revisions question the simple analogue-digital binary, with claims that the substitution of photoelectronic and computational processes for photochemical and darkroom ones need not have eroded—though it may subtly have altered—photography’s indexical quality (Soderman, 2007).

Given this intellectual commotion, what can we learn from the selfie about photographic indexicality that has not already been said? Two things: first, that the selfie as an index is less the trace of a reality imprinted on the photograph than of an action enacted by a photographer; second, that the selfie exploits indexicality in favor of connective performance rather than semantic reference. These two things are intertwined, and they reconfigure photography in relation to distinct dimensions of indexicality that are often fused, which Doane (2007) calls “index as trace” and “index as deixis” (see also Niessen, 2011). The former dominates traditional photography theory, casting the photograph as equivalent to a footprint in the sand: A material trace or imprint created by “contact” with its object, the photograph foregrounds the temporal relation of pastness with its original event. Yet Doane (2007) claims that index as deixis is equally important. Deixis is the pointing finger, directing attention onto a present object: as Barthes says, “the Photograph is never anything but an antiphon of ‘Look’, ‘See’, ‘Here it is’” (2000, p. 5). The connection to deictic language—“this,” “that,” “here,” and “now” and personal pronouns “I,” “you,” etc.—is especially revealing. These terms acquire sense in reference to a present context of discourse that constantly changes: Where “here” is alters from one execution to another. The index as deixis operates in the temporality of the mutable present rather than of the salvaged past.

Like much everyday digital photography, the selfie tips the balance between these forms of indexicality. The advent of photography as a “live” medium, using digital networks to connect interlocutors in space rather than in time, brings it closer to a conversational practice that draws images and their referents into the immediate moment of discursive interaction (which applications like Whatsapp and especially Snapchat both promote and exploit). It also turns the temporal oscillation of the photograph as trace (Durand, 1993)—between the “now” of viewing and the “then” of the depicted scene—into a spatial oscillation between a proximal “here” and a distal “there.” The selfie is a form of relational positioning between the bodies of the viewed and viewers in a culture of individualized mobility, where one’s “here” and another’s “there” are mutually connected but perpetually shifting (Garcia-Montes, Caballero-Munoz, & Perez-Alvarez, 2006). It continually remolds an elastic, mediated spatial envelope for corporeal sociability.
But the selfie does more than this: It deploys both the index as trace and as deixis to foreground the relationship between the image and its producer because its producer and referent are identical. It says not only “see this, here, now,” but also “see me showing you me.” It points to the performance of a communicative action rather than to an object, and is a trace of that performance.

This performance is embodied. As Jerry Saltz observes:

selfies are nearly always taken from within an arm’s length of the subject. For this reason the cropping and composition of selfies are very different from those of all preceding self-portraiture. There is the near-constant visual presence of one of the photographer’s arms, typically the one holding the camera. (2014, p. 4)

These arms assume the role of the pointing finger: They implicitly designate the absent hands and their held devices as the site of pictorial production.

*Figure 1. Deictic arms—one of Justin Bieber’s many selfies.*

Source: http://bobritzema.wordpress.com/2013/12/12/21st-century-selves-v-the-selfie
The selfie, then, is the culmination and also the incarnation of a gesture of mediation. It is an observable "sensory-inscription" (Farman, 2012) of the body in and through technological means. The body is inscribed in part into an already existing order of interpersonal signification—gestures have meanings in face-to-face interactions—but it is also inscribed as a figure for mediation itself: It is simultaneously mediating (the outstretched arm executes the taking of the selfie) and mediated (the outstretched arm becomes a legible and iterable sign within selfies of, among other things, the selfieness of the image). To understand more about how the selfie communicates this sensory inscription, another term needs to be brought into play.

Composition

Visual composition usually refers to the arrangement of elements within the space of a picture and their orientation to the position of the viewer (Kress & van Leeuwen, 2004, p.181). One key feature of conventional photographic composition that has remained relatively unchanged across the analog-digital divide is the spatial separation between photographed objects and the photographer's body. The depicted scene is produced from a position behind the camera, a position almost always occupied by the photographer and subsequently adopted by the viewer. Although there is a venerable history of photographic self-portraiture (Lingwood, 1986), literally putting oneself in the picture (Spence, 1986) relies on technological work-arounds like timers or remote-control devices, the use of reflective surfaces, or a human proxy. Taking a conventional photograph means, as a rule, not being in it.

This backstage of image-production generates a linear gaze through the apparatus of the camera toward those being photographed. It also encourages a directorial performance of spatial evacuation: photographers shooing unwanted objects off frame as potential interferences. Traditional camera design and use—of both analogue and digital devices—means that the camera is not just a machine for making pictures; it is also a barrier between visible photographed spaces and undepicted locations of photographing and viewing. Composition, the integration of elements "into a meaningful whole" (Kress & van Leeuwen, 2004, p.181), is thus based on a foundational cleavage between seeing and being seen, directing others and being (com)posed. Such compositional separation tends to underpin asymmetrical power relations between viewer and viewed (Beloff, 1983; Frosh, 2001; but see Rose, 2010, on the complexity of these relations), drawing its scripts from broader disciplinary "scopic" regimes shaping social relations (Jay, 1988; Sekula, 1989; Tagg, 1988).

Three features of smartphone design enable the selfie to challenge this spatio-representational segregation: They can be held and operated relatively easily by one hand, they display an image of the pre-photographic scene large enough to be viewed at arm's length, and they include front- and back-
facing cameras. The first consequence of this challenge is that the photographing self is easily integrated into the depiction. The space of photographic production or enunciation is effortlessly unified with the space of the picture itself, and not photographing oneself as part of an event or scene becomes an aesthetic, social, political, and moral choice (Andén-Papadopoulos, 2014; Becker, 2013) rather than a sine qua non of the photographic act. Group selfies are particularly striking examples of this transformation, where the photographer is usually at the forefront of a mass of faces and bodies, visibly participating in the process of composing the image as it is taken.


Additionally, the unified space of production and depiction becomes a field of embodied inhabitation. The camera becomes literally incorporated, part of a hand-camera assemblage whose possibilities and limitations are mutually determined by technical photographic parameters (available light, field of view, angle, etc.) and the physical potential and constraints of the human body. The most important embodied constellation consists of (1) moving one’s outstretched arm holding the smartphone or tablet at a calculated angle before the face or body, (2) the sensorimotor coadjustment of those body parts that are to be photographed (frequently the face and neck), and (3) the visual and spatial coordination of these two in composing the image to be taken via the device’s screen. The very term “composition” is reconfigured through this constellation. To “com-pose” acquires a hyphen. No longer does it refer to the arrangement of elements in a representation whose origin it hides; now it refers to the act of posing together, mutually emplacing the photographing body and the depicted figure. The dominant figuration of the body shifts from the still, invisibly directed pose of others in traditional everyday
photography to the dynamic, visible, self-animated gestural action of limbs and faces in selfies.\textsuperscript{4}

That the body is both the platform and the limitation of this new kind of self-depiction is evident from the deliberately extreme examples of photographic pyrotechnics assembled by projects such as the Selfie Olympics (Figure 3) and in reflexive images of oneself taking a selfie (Figure 4):

\textbf{Figure 3. The body as limit and platform: 2014 Selfie Olympics image.}

\textsuperscript{4} For a discussion of Ellen DeGeneres’ Oscar selfie (Figure 2) in terms of sociability rather than (only) celebrity and commercial promotion, see the post “The Ur-Selfie” at The Carceral Net: http://thecarcerainet.wordpress.com/2014/05/16/the-ur-selfie.
These athletic examples remind us that taking selfies is not natural to the body: It is an acquired skill and requires practice, the attainment of limbic and manual dexterity (activating the right button or icon to take the picture while often holding the device at extreme angles to maximize headspace), and the calibration of the body to technical affordances and desirable representational outcomes. The selfie is both expressive and disciplinary: This is the duality of most kinds of sensory inscription. Just as the moving body is the platform for the smartphone, so the device is the picturing agency that motivates, justifies, and disciplines the body’s performance. These two faces of embodied
technicity are inherent, and frequently explicit, in ordinary (i.e., nonathletic) selfies, combined in the visible arm gesture or in its implied presence.

Yet that gesture not only composes technicity and embodiment in the moment of image production; it also constitutes a deictic movement of the body that draws attention to the immediate context of image viewing and to the activity of a viewer. The suggestive power and versatility of this deictic movement of the photographing gesture toward the viewer is vividly demonstrated in "Around the World in 360°: 3 Year Epic Selfie" by Alex Chacon, a three-minute video stitching together the photographer’s video selfies taken as he traversed “36 countries in 600 days using 5 motorcycles” (Chacon, 2014). A key feature of the film is the use of the outstretched arm, especially the prosthetic limb of a stick camera mount, to create a visual-corporeal lexicon of direct kinetic relations between Alex’s body and the viewer.

Tug of war:
Walking together:

![Figure 6. Still from Alex Chacon “Around the World in 360°: 3 Year Epic Selfie.”](https://www.youtube.com/user/chaiku232)

Hand on shoulder:

![Figure 7. Still from Alex Chacon “Around the World in 360°: 3 Year Epic Selfie.”](https://www.youtube.com/user/chaiku232)
These (and other) gestures invite the viewer to infer and adopt a physical position in relation to the photographer. Manifested in the suggestion of bodily contact, the gestures propose a particular kind of sociable interaction: the act of accompanying and the subject position of companionship.

Murad Osmann’s Instagram series “Follow Me” provides a converse yet parallel example of how a deictic gesture can generate kinetic relations between viewer and viewed.
Figure 9. Murad Osmann #followmeto Brick Lane London.

Source: http://instagram.com/muradosmann
In this project, each image is based on an identical formal template: the intrusion of the photographer’s arm as he (and, by implication, the viewer) is led by the hand by a young woman, always seen from the back in a different location. Unlike the broad lexicon of deictic and kinetic visual figures in Chacon’s film, Osmann’s series demonstrates the power of a single gesture of implied physical movement and sociable companionship sustained across multiple iterations. As with Chacon’s project, however, the outstretched arm (or prosthetic stick mount) doesn’t just show the photographer depicting himself. It also draws the viewer in as a gesture of inclusion, inviting you to look, be with, and act.⁵

⁵ Purists may question whether Osmann’s images are really selfies. The photographer’s arm is a perpetual visual synecdoche for his presence in his own images; hence, if they are not true selfies, they are sufficiently close to be of relevance. Thanks to Gefen Frosh for help on this point.
Reflection

Oliver Wendell Holmes famously observed that photography is a “mirror with a memory” (1859/1980, p. 74).

This may have been metaphorically apposite to photography in general, but as we have seen, it was not literally applicable to most cameras. The popular iconography of the selfie literalizes Holmes’ trope. Rather than forming a barrier between photographer and viewed, the smartphone camera produces a reflective image for beholding oneself, resembling nothing as much as a pocket makeup mirror.
It is all too easy, then, to conceptualize the selfie’s gestural invitation to look through a voyeurism-narcissism model of mediated performance (Abercrombie & Longhurst, 1998). As others have noted, the accusation of narcissism is one of the most common themes in public discourse about selfies. While there is an enticing self-evidence to the accusation, it is often unnecessarily reductive. With important exceptions such as Mendelson and Papacharissi’s (2010) analysis of Facebook self-portraits, it tends to block further thought regarding both selfies and narcissism, and it frequently ignores its own gendered assumptions linking young women with fickle self-obsession.

Other interpretive extensions of the visual trope of reflection provide food for thought. The first, and most evident, is that the selfie is a reflexive image. Reflexive texts are usually understood to direct attention to the conditions and context of their own presence, activity, or efficacy: They are expressly self-referential (Stam, 1992). With regard to the selfie, this observation itself has two applications. One is that the selfie is self-referential as an image. It makes visible its own construction as an act and a product of mediation. Losh (2014) calls this “transparent mediation”; a parallel term is “hypermediacy” (Bolter & Grusin, 1999). Not every selfie is reflexive in this sense, though the popular subgenre employing mirrors and screens (see Figure 4) suggests the achievement of playful, generic self-consciousness.

More prosaically, and perhaps more significantly, selfies are a genre of personal reflexivity. This is true of all selfies by definition: They show a self, enacting itself. Selfies extend the photographic grammar of everyday communication: They are an instantly recognizable visual correlate to the linguistic self-enactment routinely performed by reflexive verbs. Indeed, their ability to combine transparent mediation and personal reflexivity reveals the very instability of the term “self” as a deictic shifter, fluctuating between the self as an image and as a body, as a constructed effect of representation and as an object and agent of representation.

Tied to technocultural conditions of synchronous connectivity with others, the self-enactment elicited and shown by selfies is simultaneously mediated, gestural, and sociable. These three features combine in the invitation to the viewer implicitly made by the outstretched arm. They also converge through a further twist in the trope of reflection: the centrality of imitation and mirroring to human cognition, emotion, and communication. Psychological work on imaginative projection and mental simulation (Currie & Ravenscroft, 2002), neurological research on mirror neurons (Gallese, 2005), and aesthetic scholarship on make-believe as the basis for mimesis (Walton, 1990), despite their important differences, connect human responses to representations with sensory and mental processes that imitate depicted states, especially, though not exclusively, the motor activity of the body. Put very crudely, responses to representations are built upon embodied simulation of what is shown: neurological or unconscious mental processes that perform bodily and sensory imitations, as it were, offline.

These ideas, though contested within their various fields, are extremely fertile for thinking about the selfie as a gestural invitation to distant others. They enable us to conceptualize the selfie as a sensorimotor (rather than merely sensory) inscription of a bodily gesture into a still image that summons

---

6 For a critique of the popular (and scholarly) charge that the selfie is narcissistic, see The Carceral Net: http://thecarceralnet.wordpress.com.
us to do more than look. The selfie invites viewers, in turn, to make conspicuously communicative, gestural responses. Sometimes, viewers respond to selfies in kind, taking reactive selfies that themselves summon further response. Here, sensorimotor mirroring is almost literally achieved. In most cases, however, the action is displaced into other physical movements that execute operations—“like,” “retweet,” “comment”—via the social media platforms on which the selfies are seen. Like the selfie, such operations are also performed through sensorimotor actions that are semiconscious yet habitual to the degree that we might even call them “reflex”: fingers swiping and tapping apps on touchscreens or scrolling, moving, and clicking a mouse attached to a desktop computer. In Osmann’s “Follow Me” series, for example, viewers cannot literally follow the woman’s outstretched arm into the image, but the kinetic power of the gesture redirects this sensorimotor potential to a different operation of the hand and a substitute performance of “Follow” (circled in red on Figure 12):

![Figure 12. Screen shot of Murad Osmann #followmeto the Camp Nou Stadium, Barcelona. Source: http://instagram.com/p/lcymRFmsyE/](image)

As a gestural image, then, the selfie inscribes one’s own body into new forms of mediated, expressive sociability with distant others. These are incarnated in a gestural economy of affection as the reflex bodily responses by which we interact with our devices and their interfaces: the routinely dexterous movements of our hands and eyes.
Conclusion: The Selfie, the Phatic Body, and Kinesthetic Sociability

We have moved far from the primarily visual terms of traditional photography theory and the conventional uses of indexicality, composition, and reflection. These terms have been steered toward recent conceptualizations of photography as an embodied performance and a material practice (Edwards, 2012; Larsen, 2005), intersecting with accounts of the corporeality of digital culture and its organization of affection (Hansen, 2004). The motivation behind this rethinking of core terms is not mere theoretical whimsy. Rubbing these concepts against the grain of the selfie reveals their enduring productiveness for thinking about photography. It also reanimates photography theory as a truly aesthetic project in that it is concerned not just with visual semiotic modes but also with broader somatic and sensory dimensions of cultural experience and practice.

That practice is sociable. “The impulse to sociability distils, as it were, out of the realities of social life the pure essence of association, of the associative process as a value and a satisfaction” (Simmel, 1910/1997, p. 125). The selfie—deictically indexical, inclusively composed, reflexive and reflex—alters and deepens the relationship between photographic mediation and the impulse to sociability.

Two final observations will clarify this claim. The first is that the selfie is linked to phatic communion, “a type of speech in which ties of union are created by a mere exchange of words” (Malinowski, 1923, p. 315), whose primary purpose is the production, expression, and maintenance of sociability. Miller (2008) argues that contemporary networked culture accentuates a phatic culture of nonsubstantive communication mediated across distances. His examples are primarily drawn from the banal verbal messages on Twitter and other social network services. The selfie represents a parallel process to this mainly verbal phenomenon: the production of the mediated phatic body as a visible vehicle for sociable communication with distant others who are expected to respond.

Response is crucial. Phatic exchanges stage sociability as a binding affective energy transferred between individuals in interpersonal settings, and response is an embodied social reflex—it is hard not to perform it. Highly ritualized yet profoundly routine, phatic utterances demand to be requited (Coupland, Coupland, & Robinson, 1992). Failure to acknowledge the nod of a passing acquaintance or her casual “How are you?” is easily perceived as an expression of nonrecognition and social exclusion. Yet how are mediated exchanges of responsive phatic energy made possible through the representational form of a still photograph?

This is where the second observation becomes important. The selfie is a preeminent conductor of embodied social energy because it is a kinesthetic image: it is a product of kinetic bodily movement; it gives aesthetic, visible form to that movement in images; and it is inscribed in the circulation of kinetic and responsive social energy among users of movement-based digital technologies. As a kinesthetic image, then, the selfie makes visible a broader kinesthetic domain of digital culture that is relatively overlooked as an object of analysis. This is the limbic, gestural register, overtly apparent in games systems such as Wii and Kinect, that creates a circuit for mediating social and corporeal affective energy by intersecting with two other registers of embodied technicity: the mobile, the mediated mobility of whole bodies in physical and augmented space provided by locative technologies (the body as a single
moving data point), and the operative, the habitually nimble coordination of hands and eyes used for navigating the virtual space of interfaces (body parts as media operators). The selfie is thus a new phatic agent in the energy flows between bodily movements, sociable interactions, and media technologies that have become fundamental to our everyday, routine experience of digital activities: part of what Hjorth and Pink (2014), following Ingold’s work, call digital “wayfaring.” It is a sign not only of digital photography’s agency as a form of popular depiction, but also of the further transformations of everyday figural representation as an instrument of mediated, embodied sociability.

References


