Cognitive Assemblages: Technical Agency and Human Interactions

N. Katherine Hayles

In the work of Gilles Deleuze and Félix Guattari, *assemblage* (agencement) carries connotations of connection, event, transformation, and becoming.¹ For those who privilege desire, affect, and transversal energies over cognition, it would seem unlikely to define assemblages as cognitive; however, the broader definition of *cognition* I employ brings my argument somewhat closer to theirs (although significant differences remain). I want to define cognition as a process of interpreting information in contexts that connect it with meaning. This view foregrounds interpretation, choice, and decision and highlights the special properties that cognition bestows, expanding the traditional view of cognition as human thought to processes occurring at multiple levels and sites within biological life forms and technical systems. *Cognitive assemblage* emphasizes cognition as the common element among parts and as the functionality by which parts connect.²

My reason for choosing *assemblage* over *network* (the obvious alternative) is to allow for arrangements that scale up. Starting with cognitive

---


² Alison Williams suggests that *assemblage* is an appropriate way to think about drones; however, she does not foreground the role of cognition, as I do here; see Alison J. Williams, “Enabling Persistent Presence? Performing the Embodied Geopolitics of the Unmanned Aerial Vehicle Assemblage,” *Political Geography* 30 (Sept. 2011): 381–90.
processes occurring at a low threshold—using information to make choices within contexts—cognitive assemblages can progress to higher levels of cognition and consequently decisions affecting larger areas of concern. Other advantages include the notion of an arrangement not so tightly bound that it cannot lose or add parts, yet not so loosely connected that relations between parts cease to matter; indeed, they matter a great deal. A cognitive assemblage operates through contextual relations at multiple levels and sites, with boundaries fluctuating as conditions and contexts change. Further comparisons emerge through considering the kinds of materialities involved in networks versus those in assemblages. Networks consist of edges and nodes and are analyzed through graph theory, conveying a sense of a spare, clean materiality. Assemblages, by contrast, allow for contiguity in a fleshly sense—touching, incorporating, repelling, mutating. When analyzed as dynamic systems, networks are sites of exchange, transformation, and dissemination, but they lack the sense of these interactions occurring across complex three-dimensional surfaces, whereas assemblages include information transactions occurring across membranes, involuted and convoluted surfaces, and multiple volumetric entities interacting with many conspecifics simultaneously.

Because humans and technical systems in a cognitive assemblage are interconnected, the cognitive decisions of each affect the others, with interactions occurring across the full range of human cognition, including consciousness, the unconscious, the cognitive nonconscious, and the sensory/perceptual systems that send signals to the cortex. Moreover, human decisions and interpretations feed back into technical systems, sometimes decisively affecting the contexts in which those systems operate. As a whole, a cognitive assemblage performs the functions identified with cognition—flexibly attending to new situations, incorporating this knowledge into adaptive strategies, and evolving through experience to create new strategies and kinds of responses. Because the boundaries are fuzzy, where one draws


N. Katherine Hayles is professor of literature at Duke University. She is the author of ten books, including *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (1999) and, more recently, *How We Think: Digital Media and Contemporary Technogenesis* (2012).
the line often depends on the analytical perspective and the purpose of the analysis. Nevertheless, for a given situation, it is generally possible to specify the kinds of cognitions involved and consequently to trace their effects through an evolutionary trajectory.

The most transformative technologies of the later twentieth century have been cognitive assemblages; the internet is a prime example. While many modern technologies also had immense effects—the steam engine, railroads, antibiotics, jet airplanes, nuclear weapons and energy—cognitive assemblages are distinct because their transformative potentials are enabled, extended, and supported by interactions between human and technical cognizers. Hybrid by nature, they raise questions about how agency is distributed among cognizers, how and in what ways actors contribute to systemic dynamics, and consequently how responsibilities—technical, social, legal, ethical—should be apportioned. They invite ethical inquiries that recognize the importance of technical mediations, adopting systemic and relational perspectives rather than emphasizing (I would say overemphasizing) individual responsibility.

Developing the concept of a cognitive assemblage, I begin with the basic level of a technical cognitive system, a city’s infrastructure. Nigel Thrift has called infrastructures governing our everyday assumptions about how the world works the “technological unconscious”; it consists of predispositions that regulate our actions through routine anticipations, habitual responses, pattern recognitions, and other activities characteristic of the cognitive nonconscious. From there my analysis will move inward toward the body to discuss digital assistants that interact directly on a personal level. As these devices become smarter, more wired, and more capable of accessing information portals through the web, they bring about neurological changes in the mindbodies of users, forming flexible assemblages that constantly mutate as information is gathered, processed, communicated, stored, and used for additional learning. As the user’s responses and interactions reveal more about predispositions of which the user may not be aware, the possibility for surveillance grows progressively stronger, a trajectory analyzed here through the sociometer developed by Alex (Sandy) Pentland and epitomized by Frans van der Helm’s extreme proof of concept in the MeMachine.

Continuing the theme of surveillance, I turn next to an analysis of the implications of increasing technical autonomy evident in many research programs underway: for example, self-driving cars, face recognition systems, and autonomous weapons. My focus is on the transition from pilot-operated

drones to autonomous drone swarms. Amplifying technical autonomy requires an increase in the cognitive capabilities of technical devices, so that distributed agency is preceded by and dependent upon a redistribution of cognition. The tendency of technical devices to unsettle discursive formulations and shake up cultural practices is exacerbated in the case of military drones, where nothing less than life and death is at stake. International treaties delineating so-called laws of war illustrate such seismic disturbances; they assume that agency and consequently decisional power lie entirely with humans without considering the effects of technical mediations. The significant changes brought about when technical devices do have agency illustrate what happens to complex human social systems when they are interpenetrated by technical cognition. As I will show, the resulting cognitive assemblages transform the contexts and conditions under which human cognition operates, ultimately affecting what it means to be human in developed societies.

**Infrastructure and Technical Cognition**

Imagining the future of technical cognition, Pentland writes, “it seems that the human race suddenly has the beginnings of a working nervous system. Like some world-spanning living organism, public health systems, automobile traffic, and emergency and security networks are all becoming intelligent, reactive systems with . . . sensors serving as their eyes and ears.” The analogy has not been lost on neuroscientists, who adopt traffic metaphors to characterize information flowing through the body’s nervous system. As Laura Otis argues about the connections nineteenth-century scientists made between nerves and networks of telegraph lines, such analogies have real conceptual force: “metaphors do not ‘express’ scientists’ ideas; they *are* the ideas.”

A good site to explore interactions between a city’s “nervous system” and those of humans is the Automated Traffic Surveillance and Control (ATSAC) system in Los Angeles that controls traffic on seven thousand miles of surface streets. I made a site visit there in November 2014 and spoke with Edward Yu, ATSAC’s director. The computer system at

---

ATSAC’s heart, fed by information flowing from sensors and actuators throughout the city, is intelligent, adaptive, and evolutionary, capable of modifying its own operations. When combined with human operators who work with it, ATSAC illustrates the ways in which technical nonconscious cognition works with human intelligence to affect the lives of millions of urban inhabitants.

Located four levels below the street, the center is housed in an underground bunker originally designed to protect city officials from bomb attack (that it has now been turned over to ATSAC is perhaps an unintentional acknowledgement of how crucial traffic control is to Los Angeles). Information flowing into the center includes reports from eighteen thousand loop detectors, working by magnetic induction, that register traffic volume and speed every second in over four thousand intersections, while more than four hundred surveillance cameras monitor the most troublesome or important intersections. Analyzing these data streams, computer algorithms automatically adjust the signal lights to compensate for congested lanes. This requires changing the relevant signals in coordinated fashion so that side streets coming into a main thoroughfare, for example, are timed to work together with the main street’s lights. The system also monitors traffic in the dedicated bus lanes; if a bus is behind schedule, the algorithms adjust the signals to allow it to make up time. All the monitoring information is processed in real time. The entire system is hard-wired to prevent latency, with copper wire from the loop detectors running into hubs, where the information is carried by fiber optics into the center. ATSAC thus represents a considerable civic investment in infrastructure. Begun for the 1984 Olympics in Los Angeles, it was finally completed in 2012.

In addition to everyday traffic, engineers adapt the system for special events such as a presidential visit or a blockbuster premiere. Even the automatic cycles have special provisions. In neighborhoods with large Orthodox Jewish communities, for example, the push buttons operating walk signals are programmed to work automatically during the daylight hours of the Jewish Sabbath, during which times Orthodox Jews are prohibited from working machinery and thus cannot operate the buttons manually. Since daylight hours differ by seasons, the system has programmed into it the times for sunrise and sunset for the entire year.

Without the help of the system’s sensors, actuators, and algorithms, it would be impossible for humans to conduct such widespread traffic coordination and prohibitively expensive even to attempt it. According to studies, the system has resulted in 20–30 percent fewer stops at intersections, reduced travel time by 13 percent, cut fuel consumption by 12.5 percent
These statistics have real consequences for the lives of Angelenos. Having lived in Los Angeles for two decades, I can testify to how important traffic patterns become, often dictating work schedules, entertainment possibilities, and friendship networks. When Yu attends community meetings, he likes to ask audiences how many of their lives have been affected by a major crime. Typically two or three people out of a hundred raise their hands. Then he asks how many have had their lives affected by traffic; hands shoot up from virtually everyone.

Specifically, how do the technical cognitions instantiated in ATSAC interact with human cognitions? The algorithms are coordinated with a database in which the traffic information is stored for a week; this allows the system to extract patterns, and it uses these patterns to update the algorithms accordingly. Drivers also detect patterns, no doubt at first consciously and then nonconsciously as they travel the same routes over and over. When anomalies occur, they are quick to notice and often call the center to alert operators to problems at particular intersections. The operators also must internalize the patterns so they can make informed decisions. Yu reported that it typically takes about a year and a half to train new personnel before they have enough experience to distinguish the ordinary from extraordinary. For example, Santa Monica Boulevard feeds into the Santa Monica freeway; if the freeway entrance is backed up, there is no point in arranging the signals to permit traffic to move faster on the street because that would only exacerbate the problem.

When an intersection becomes congested, the screen graph displaying that information begins to flash red, and the operator can pull up the camera feed to identify the problem. The afternoon I visited, an intersection on Alameda Street in the downtown area became congested, and the camera image revealed that police had blocked off a portion of the street to prepare for a demonstration. Unfortunately they had not informed the center, and with rush hour approaching active intervention was necessary to prevent traffic from snarling throughout the downtown area. With a single command, an operator can change a whole network of signals, as was necessary in this instance.

ATSAC exemplifies productive collaboration between human conscious decisions, human nonconsciously recognition of patterns by both operators and drivers, and the technical cognitive nonconscious of the computer algorithms, processors, and database. As Ulrik Ekman notes in discussing the topology of intelligent cities, “design here must meet an ongoing and

exceedingly complex interactivity among environmental, technical, social and personal multiplicities of urban nodes on the move.” 12 Functioning within these complexities, ATSAC demonstrates how a cognitive assemblage works. At any point, components are in flux as some drivers leave the system and others enter. Although the underlying infrastructure is stable, the computer’s screenic interfaces constantly change as older trouble spots are cleared up and new ones emerge. Similarly, the basic cognitive structures of the algorithms are set, but they are also modified through the extraction of patterns; the patterns will be used to modify continuing operations as contexts change and new meanings are produced.

In a city where traffic is crucial to the quality of life, ATSAC contributes positively to the experiences of Angelenos; the downside, of course, is that by decreasing traffic congestion it allows the (in)famous LA dependence on cars to continue and even increase. Nevertheless, it can be argued that the net consequences are still positive overall. The system is the result of investment by the city over several decades and different political regimes that nevertheless managed to summon the political will to keep the system running and to enlarge it until the entire city was included. ATSAC thus shows the possibilities for constructive outcomes from the deep penetration of the technical cognitive nonconscious into complex human systems. It is worth noting that the system has no direct connection to market considerations of profit and loss.

**Digital Assistants and Information Portals**

From the mediations of an urban traffic infrastructure, we move up close and personal to consider the more intimate and arguably more neurologically consequential interactions with a digital assistant. VIV, a digital assistant being developed by VIV Labs in San Jose, evolves its capacities through web reading, geolocation, mobile interactions, and real-life queries. The program, soon to be marketed as the “next generation Siri,” combines GPS orientation with an open system that programs on the fly, parses sentences, and links to third-party sources. 13 Developers Dan Kittlaus, Adam Chever, and Chris Brigham say that VIV can parse relatively complex queries, such as one shown in *Wired* alongside a flow chart indicating VIV’s search techniques: “On the way to my brother’s house, I need

---


to pick up some cheap wine that goes well with lasagna.” The program first parses “brother” as a family relationship and looks for the appropriate entry in the user’s Google contacts. It then plots the route and, noticing a missing variable, asks the user how far she is willing to deviate to get the wine. With this information, the program searches the web for lasagna recipes, identifying it as a food item containing tomato, cheese, and pasta. Searching for wine-food matches yields appropriate varietals, and further inquiries yield price ranges from which a “cheap” spectrum is identified. The next step queries catalogues of the wine stores en route, with the final result being a list of wines at the appropriate stores.

VIV is designed to learn continuously, keeping track of its inferences in a growing database. Unlike language learning programs such as the Never Ending Language Learning (NELL) program developed by Tom Mitchell and his group at Carnegie Mellon University, VIV has the advantage of interacting with the real world, including locations, the user’s range of interests, and indicators of tastes and preferences. With a wider range of movement and knowledge, VIV operates not only through the formal manipulation of symbols but also by correlating trade-offs such as time versus money, quality versus price, and proximity versus distance. It interacts with the user’s cognitions in directing movements, interpreting sensory signals such as location indicators, and responding to queries. Moreover, it has the advantages of massive storage space in the cloud, fast processing speed, and computational intensity of data manipulation. VIV and the user, considered as parts of a cognitive assemblage, constitute stable communicating nodes around which hover a cloud of other functionalities associated with web queries and curating algorithms that correlate VIV’s data with other information about the user stored in myriad relational databases.

If the rollout of VIV is successful, we can anticipate that it will have considerable commercial value because it will integrate geolocation with specific product requests. Moreover, it will enable data collection on a scale that surpasses even that available with present smartphones and desktop searches. It will be able to correlate user movements in real time, location data about present trajectories as well as past trips, and the queries and purchases associated with these trips. Some of this functionality is already available in GPS devices—for example, the store locations displayed under search categories, information that corporations pay to have listed. When several routes are available, the GPS typically chooses the one that passes by the greatest number of listed stores, for example,

14. Ibid.
malls or shopping plazas. All this, and more, will be potentially available with VIV.

Having such a smart digital assistant will also affect how users integrate this technical cognition into their daily lives. We can expect that certain evolved cognitive abilities present in human brains—for example, the ability to navigate and judge where one is in the world—will receive less stimulation with this device because now the device navigates for the user and the synaptic networks involved with navigation will tend to shrink. We know from experiments evaluating web scanning versus print reading that human neurology can undergo changes after even minimal exposure to digital media, with long-lasting effects. The predictable result is that the more one uses a digital assistant such as VIV, the more one needs to use it, as one's natural abilities to navigate in the world decline, perhaps to the point of atrophy.

Moreover, the device will likely result in a certain homogenization of behavior; exploratory wandering will decrease and routinization of movement will increase. The same will be true of shopping practices; there will be less wandering around the store and more direct selection of desired items, where desire is itself manipulated by marketing. Overall, the interpolation of the user into corporate designs will be intensified and expanded. To the extent that augmented reality may also be part of VIV’s functionality, this intensification will occur on nonconscious as well as conscious cognitive levels. As with other digital affordances, VIV will follow what Bernard Stiegler has characterized as a pharmacological dynamic of poison and cure, offering the powerful advantages of convenience, satisfaction of desires, and enhanced navigation while increasing surveillance, directed marketing, and capitalist exploitation.

Are devices such as VIV the predecessors of fully conscious technical systems, as Spike Jonze’s film Her (2013) suggests? If the twentieth and twenty-first centuries have taught us anything, it is that only fools would rush in to proclaim “impossible!” The real temptation here, however, is to imagine that we are awaiting the arrival of technical cognition when it is already realized in myriad complex systems and computational devices. Humans living in societies with deep technological infrastructures are already enmeshed in cognitive assemblages shaped and interpenetrated by technical cognition, including language-learning systems. If “the language

“instinct” separates humans from other biological organisms in a rupture so extreme it marks human uniqueness, as Steven Pinker has controversially argued, the language gap is narrowing between humans and their digital interlocutors.\footnote{17. See Steven Pinker, \textit{The Language Instinct: How the Mind Creates Language} (New York, 1994).}

**Social Signaling and Somatic Surveillance**

Pentland, working with his graduate students at MIT Media Lab, has developed a device he calls the sociometer; it detects, measures, and displays physiological indicators of social signaling among groups (see \textit{HS}). The device, worn like a shoulder badge, detects who is talking to whom (via IR transceivers), for how long, and with what intensity (see \textit{HS}, pp. 99–111).\footnote{18. See Tanzeem Choudhury and Pentland, “The Sociometer: A Wearable Device for Understanding Human Networks,” Alumni.media.mit.edu/~ttanzeem/TR-554.pdf} It also analyzes nonlinguistic speech features such as consistency of emphasis, tracks body movements and infers from them the activities involved, measures proximity to other people, and, from these data, identifies the social context. Pentland calls the behaviors measured by the sociometer “honest signals” because they take place below the level of conscious awareness; he further argues that attempting to fake them would require so much effort that it is virtually impossible (see \textit{HS}, pp. 2–3).

The sociometer, like the human cognitive nonconscious, integrates incoming somatic markers and other chemical and electrical signals to create integrated representations of body states. As we have seen, the human cognitive nonconscious recognizes and interprets patterns of behavior, including social signals emanating from others.\footnote{19. See Hayles, “The Cognitive Nonconscious.”} The importance of this function becomes apparent when it is externalized in the sociometer, for the social signals it detects enable Pentland and his group to predict outcomes for various kinds of interactions, from salary negotiations to dating preferences. Even in slices of behavior as short as thirty seconds, sociometer data indicate accurately what decisions a group will reach. It is worth emphasizing that these predictions are derived solely from the sociometer’s analysis of social signals with no attention to verbal content or rational argument.

There are several implications to be drawn from these experiments. First, they indicate the usefulness of the sociometer as a feedback device to help a group improve its functioning. Pentland reports that his lab has developed “a computer algorithm that builds on the sociometer’s ability to read the group’s honest signaling. Using this technology we are
beginning to build real-time meeting management tools that help keep groups on track, by providing them with feedback to help avoid problems like groupthink and polarization” (HS, p. 49).

More fundamentally, the sociometer data indicate how essential social signals are to human sociality and, conversely, how much more limited rational discussion and conscious deliberation may be than traditionally supposed. In this respect, humans may have something in common with social insects such as bees and ants (as E. O. Wilson argues in another context).20 “Honest signals are used to communicate, control the discovery and integration of information, and make decisions,” Pentland writes (HS, p. 83). Extrapolating from this insight, he argues that “important parts of our intelligence exist as network properties, not individual properties, and important parts of our personal cognitive processes are guided by the network via unconscious and automatic processes such as signaling and imitation.” With classic understatement, he predicts that “we will come to realize that we bear little resemblance to the idealized, rational beings imagined by Enlightenment philosophers” (HS, p. 88).

In addition to this major conclusion, there are important corollaries. Because social signals take time to generate, transmit, receive, and recognize, they operate on a slower timeline than either the cognitive non-conscious (in the two-hundred-millisecond range) or consciousness (in the five-hundred-millisecond range). Pentland estimates they work in the thirty-second range, a temporality analogous to the time it takes to interpret complex verbal information such as narratives (see HS, pp. 107–11). This implies that the processing of verbal information and social signals occur with similar delays, opening the possibility of a feedback loop between them, such as happens when two people engaged in conversation begin to mimic one another’s gestures as they become more engaged with each other verbally, each form of communication reinforcing the other. Pentland references research showing that when this kind of mimicry occurs, interactors report that they like the other more, trust each other more, and reach more accommodating outcomes (see HS, pp. 10–11). Another corollary is that social signaling affects both participants, unlike verbal language, which can be directed at another without necessarily affecting the speaker. Pentland summarizes this effect: “When you engage in social signaling, you are often affected just as much as the other person. Signals are two-way, not one-way . . . so that pulling on one corner of the social fabric stretches all members of the network” (HS, p. 40).

Evolutionarily, it is likely that social signaling developed before language; many mammals use such signals to negotiate territory, communicate intentions, and coordinate group activity. Pentland references brain research indicating that “we all have networking hardware that gives us the ability to read and respond to other people,” specifically the mirror neurons mentioned in my previous essay (HS, p. 37). This signal-response channel seems to have evolved much earlier than the linguistic channel,” he continues, “with language building on top of the capabilities of this channel,” a trajectory analogous to nonconscious cognition developing first, with consciousness emerging later and being built on top (HS, p. 42).

The sociometer, then, may be regarded as an exteriorization of the cognitive nonconscious, collecting, interpreting, analyzing, and displaying information in ways that make it available for conscious consideration. Insofar as social signals are essential to effective group functioning, we may draw a somewhat different conclusion from those that concern Pentland. The sociometer’s exteriorization reveals that the work of the cognitive nonconscious is crucial to social networking, group decision making, and indeed human sociality. Functioning as part of a cognitive assemblage, the sociometer is enmeshed in a human-technical system with porous borders that depend on who is using the device for what purpose—whether to monitor one’s own responses, to surveille those of someone else, or to analyze group behavior from the outside, so to speak, analyzing its dynamics with or without the group’s permission.

The significance of this opening out suggests the need for a new category of analysis that I call somatic surveillance. While traditional surveillance technologies focus on exterior appearances, movements, clothing, and such, somatic surveillance registers the body’s interior markers and displays them in real time. The concept is not entirely new. Lie detector tests, for example, measure physiological responses such as heart rate and galvanic skin response and display them through a moving graph; hospital medical devices display heart rate on monitors. Although these technologies operate in real time, they lack two crucial features illustrated by the sociometer—mobility and miniaturization, properties necessary to make the technology wearable. The idea of using wearables for somatic


22. For an analysis from a medical viewpoint of the importance of mobility, see R. S. Epstein, “Mobile Medical Applications: Old Wine in New Bottles?” Clinical Pharmacology and Therapeutics 95 (May 2014): 476–78.
surveillance is still relatively new, and its implications remain largely unexamined.

This state of affairs catalyzed Frans van der Helm to create what he calls the MeMachine. Internationally renowned for his cutting-edge research in robotics and prosthetics, van der Helm and his lab have developed hardware devices and software programs to aid them in their research, including an animated display of an anatomical figure mirroring in real time the movements of a subject wearing multiple tracking sensors on his body. Because their work requires detailed knowledge of how muscles, tendons, and joints move, the figure is imaged sans skin to reveal more clearly the complex interactions that constitute human mobility. This figure, appearing somewhat like one of the flayed bodies in Gunther von Hagens’s Körperwelten (Body Worlds), is all the more dramatic because it moves in perfect synchrony with its real-life counterpart. The capability to create this image had already been developed for research purposes, so it required only a few tweaks to make it into the occasion for a spectacular demonstration of somatic surveillance: the MeMachine.

In 2012, van der Helm demonstrated the MeMachine in a videotaped presentation before a large audience. The video shows the preparations: an anonymous hairy chest being shaved so sensors can be attached; sensors being carefully placed on each finger; a headband with eye-tracking devices slipped into place; and a backpack with the computer equipment communicating wirelessly with the sensors hefted onto the shoulders. Preparations complete, a nattily suited van der Helm strides on stage wearing this gear, his face partially obscured by the headband, gloves concealing the finger sensors, clothing the ones on his body. On screen, however, everything is revealed; as the flayed figure mirrors his posture, gestures, and body language, data scrolls underneath showing electrocardiogram (ECG) readings of his heartbeat, electromyography (EMG) tracking of electrical impulses generated by his muscles, galvanic response indicating skin conductance, electroencephalogram (EEG) readouts showing his brainwaves, and an inset using the eye-tracking data to show where he is looking. Each data cluster is estimated in real time, so the audience knows his focus of attention, arousal level, emotional state, brain activity, and muscle tension. The kinds of information that the cognitive nonconscious processes are here opened out to an unprecedented degree. Moreover, in Stelarc-like fashion, he suggests in his videotaped lecture that in future instantiations of the MeMachine the audience may be invited to intervene in his somatic states; for example, they may set limits for physiological parameters that, if exceeded, “punish” him by removing some of his functionality. In this case, the boundaries of the MeMachine’s cognitive
assemblage would include any number of spectators as well as van der Helm and the technical devices monitoring his activities. The prospect suggests that the MeMachine’s exteriorization could create surveillance scenarios in which virtually nothing of the surveilled person’s thoughts, affects, and autonomic responses would remain private or capable of resisting exterior scrutiny.

When I spoke to van der Helm after his presentation at Delft Technical University, he framed the project as a provocation designed to catalyze discussions about the ethical, social, and cultural implications of somatic surveillance. He reported that ethicists, philosophers, and others who had seen the MeMachine demonstration were shocked by its surveillance potential, proposing that issues of privacy, public policy, and ethical guidelines should be explored before the technology was developed further. To this view, van der Helm was vehemently unsympathetic, asking why these people had not come to his lab to engage in dialogue and what gave them the right to impose constraints after the fact. Whatever one makes of these conflicting positions, the MeMachine demonstrates the extent to which the workings of the cognitive nonconscious can be externalized through technical mediation, creating situations in which the human cognitive nonconscious, technical cognition, and human consciousness interact in real time through multiple feedback loops and recursive circular causalities to create a cognitive assemblage of unprecedented surveillance potential.

**Distributed Agency and Technical Autonomy**

Cognitive technologies show a clear trajectory toward greater agency and autonomy. In some instances, they are performing actions outside the realm of human possibility, as when high-frequency trading algorithms conduct trades in five milliseconds or less, something no human could do. In other cases, the intent is to lessen the load on the most limited resource, human attention—for example, with self-driving cars. Perhaps the most controversial examples of technical autonomy are autonomous drones and robots with lethal capacity, now in development. In part because these technologies unsettle many traditional assumptions, they have been sites for intense debate, within both the military and civilian communities. They can therefore serve as test cases for the implications of distributed agency and, more broadly, for the ways in which cognitive

---

assemblages interact with complex human systems to create new kinds of possibilities, challenges, and dangers. To limit my inquiry, I will focus on autonomous drones, but many of the same problems attend the creation of robot warfighters, as well as nonmilitary technologies such as self-driving cars and quasi-military technologies such as face recognition systems.

The present moment is especially auspicious for analyzing technical autonomy because the necessary technical advances are clearly possible, but the technical infrastructures are not so deeply embedded in everyday life that other paths are locked out and made much more difficult to pursue. In short, now is the time of decision. Debates entered into and choices made now will have extensive implications for the kinds of cognitive assemblages we develop or resist and consequently for the kind of future we fashion for ourselves and other cognitive entities with whom we share the planet.

My focus will not be on drone assassinations carried out by the US in other countries without respect for national boundaries, including some American citizens killed without trial or jury, in clear violation of the Constitution and civil rights. This aspect is well covered by Medea Benjamin, who passionately opposes the drone program both for its unconstitutionality and more specifically for the horrific toll in civilian deaths (“collateral damage”), estimated to be as high as 30 percent. I also will not consider the myriad uses emerging for civilian unmanned aerial vehicles (UAV), including rangeland monitoring, search and rescue missions, emergency responders in the case of fire and other life-threatening events, and UAVs used as mobile gateways or “data mules,” collecting data from remote environmental sensors scattered over large territories. Rather, I will focus on piloted and autonomous UAVs, as well as multivehicle systems proceeding autonomously, with the swarm itself deciding which individual will play what role in an orchestrated attack. This range of


26. Some nomenclature clarification is necessary. The technical term for drone is UAV (unmanned aerial vehicle), which requires two pilots, one to guide the craft and the other to monitor sensory data. If more than one aircraft is involved, it becomes UCAVS, unmanned aerial vehicle system; if intended for combat, UCAVS, unmanned combat aerial vehicle system. Autonomous drones are also called UAVs, unmanned autonomous vehicles, aerial understood...
examples, showing different levels of sensing abilities, cognitions, and decisional powers, illustrates why greater technical cognition might be enticing and what kinds of problems it poses.

With the massive shift after 9/11 from state-on-state violence to what Norman Friedman, a military analyst, calls expeditionary warfare, targets are not associated with a geographically defined entity but with highly mobile and flexible insurgents and “terrorists.” Friedman points out that if surveillance can be carried out without the enemy’s ability to perceive it, then the enemy is forced to devote resources to hiding and concealing its positions, which not only drains its ability to carry out offensives but also makes it more difficult for it to organize and extend its reach.\(^{27}\) These factors, he argues, combine to make UAVs superior to manned aircraft for expeditionary warfare. A fighter jet can typically stay aloft for only two hours before it needs to refuel and the pilot, fatigued from high altitudes, needs to rest. In contrast, the UAV Global Hawk can stay aloft for days, refueling in the air; with no pilot aboard, pilot fatigue is not a problem.\(^{28}\) These factors have led to a significant redistribution of resources in the US Air Force, with more UAV pilots currently being trained than pilots for all types of manned aircraft combined. Spending about six billion dollars annually on drone development and purchase,\(^{29}\) the Air Force currently deploys about seven thousand drones, compared to ten thousand manned aerial vehicles.\(^{30}\)

The factors that have made UAVs the contemporary weapons of choice for the US Air Force required the coming together of many technological advances, including global satellite positioning, superior navigation tools, better aerodynamics for increased stability and fuel economy, increased computational power, and better sensors for visual reconnaissance, obstacle avoidance, and ground coordination. The weak link in this chain is the need to maintain communications between the UAV and the remote pilot. As long as the UAV’s performance requires this link, it is subject to disruption either internally, as when the remote pilot banks too suddenly and the link is lost, or because the link has been hijacked and control wrested away by another party, as happened when a Lockheed Martin RQ-170

Sentinel drone landed in Iranian territory in December 2007, likely because the UAV was fed false GPS coordinates by the Iranians.

This implies that the next wave of development will be unmanned autonomous aerial vehicles (UAAVs), unmanned vehicles that fly autonomously, and UAAVS, unmanned autonomous aerial multivehicle systems. Still at a nascent stage, UAAVs and UAAVSs are nevertheless developing rapidly. The Navy, for example, is developing the experimental X-47B Stealth UAV that can perform missions autonomously and land on an aircraft carrier without a remote pilot steering it. Moreover, the technical components necessary to make UAAVs and UAAVSs reliable and robust are coming together very quickly in transnational research projects, particularly in the US and China.

A recent article written in English by Chinese researchers illustrates the growing awareness of UAAVs of their internal states as well as the environment. The study discusses the development of software that allows a swarm to coordinate its individuals in cases where one or more vehicles are assigned to attack. The model uses an “auction” strategy, whereby each individual responds to a request for a bid by assessing what the authors call its “beliefs,” “desires,” and “intentions,” which are calculated with weighted formulae resulting in a quantitative number for the bid. The software enables the swarm to balance competing priorities in rapidly changing conditions, taking into account their position, velocity, and proximity to one another (“beliefs”), their assigned mission priorities (“intentions”), and the intensity with which they will execute the mission (“desires”), with the latter parameters tailored for specific mission objectives. The anthropomorphic language is not merely an idiosyncratic choice, for it indicates that as the sensory knowledge of external and internal states, autonomous agency, and cognitive capabilities of the swarm increase, so too does its ability to make decisions traditionally reserved for humans.

With autonomous drones and other autonomous weapons on the horizon, there has been increased attention to the ethical implications of their use.

Most of these discussions refer to the Geneva Conventions and similar protocols, which require that weapons must “distinguish between the civilian population and combatants” (quoted in “LH”). In addition, international humanitarian law prohibits disproportionate attacks, defined as ones that “may be expected to cause incidental loss of civilian life, injury

32. Ibid.
to civilians, damage to civilian objects, or a combination thereof, which
would be excessive in relation to the concrete and direct military advan-
tage anticipated” (quoted in “LH’’). Finally, one might take into account
the rather vague concept of “military necessity,” defined by Armin Krish-
nan as dictating that “military force should only be used against the enemy
to the extent as is necessary for winning the war,”33 and the even vaguer
Martens Clause, intended to cover instances not explicit in the Geneva
Conventions. It requires that weapons be consistent with the “principles
of humanity” and the “dictates of public conscience” (quoted in “LH”).

In assessing autonomous weapons in these terms, Human Rights Watch
and the International Human Rights Clinic (IHRC) at the Harvard Law
School argue that autonomous weapons, including autonomous drones,
cannot possibly make the required distinctions between combatants and
civilians, particularly in the context of insurgent tactics that deliberately
seek cover within civilian populations. Peter Singer, for example, offers
the case of a gunman who shot at a US Ranger with “an AK-47 that was
propped between the legs of two kneeling women, while four children
sat on the shooter’s back.”34 They also argue that “proportionality” and
“military necessity” are violated by drones, although “necessity” clearly is
itself a moving target, given that what constitutes it is obviously context-
dependent and heavily influenced by the kinds of weaponry available (see
“LH”).

The Geneva Conventions were, of course, forged in the aftermath of World
War II, characterized by massive state-on-state violence, firebombing of cit-
ies, gratuitous destruction of cultural monuments, and the nuclear holo-
casts wreaked by the US upon Nagasaki and Hiroshima. With the move to
expeditionary warfare, rise of insurgent attacks, and continuing increases in
US drone attacks, these protocols seem badly outdated, even inappropriate.
Why keep coming back to them? On an even deeper level, why should we
care about ethics in wars where the intent is precisely to kill and maim? Why,
as Singer puts it, try to determine what is right when so much is wrong, a
question that drives straight to the oxymoron implicit in the phrase “laws of
war” (“E,” p. 309). His defense is that the Geneva Conventions, obsolete as
they may be, are the only international accords on the conduct of warfare we
have and that we are likely to have, short of another world war with even more
devastating violence. He believes there is an important distinction between

33. Armin Krishnan, Killer Robots: Legality and Ethicality of Autonomous Weapons (Farnham,
UK, 2009), p. 91.
34. Peter Singer, “The Ethics of Killer Applications: Why Is It so Hard to Talk about Morality
When It Comes to New Military Technology?” Journal of Military Ethics 9, no. 4 (2010): 303;
hereafter abbreviated “E.”
those who practice restraint in warfare and “barbarians” who are willing to go to any extreme, however savage and brutal (“E,” p. 309). To argue thus is to plunge into a definitional abyss; what counts as restraint and barbarism are as contextually and culturally dependent as the distinctions they propose to clarify.

A better approach, I argue, is to evaluate the ethical questions surrounding piloted and autonomous drones from the relational and processual perspectives implicit in the idea of a cognitive assemblage. Mark Coeckelbergh, one of the few philosophers urging a relational perspective on robotics, observes that most ethical theories carry over to robotics the assumption characteristic of liberal individualism, taking as “the main object of ethical concern the individual robot”; in contrast, he argues that “both humans and robots must be understood as related to their larger techno-social environment.”

Regarding ethical issues through the perspective of a cognitive assemblage foregrounds the interpretations, choices, and decisions that technical and human components make as information flows from the UAV’s sensors, through choices performed by the UAV software, to interpretations that the sensor and vehicle pilots give to the transmitted data, on to the decision about whether to launch a missile. This decision involves the pilots, their tactical commander, associated lawyers, on up to presidential advisors and staff—a situation vividly displayed in the recent film Eye in the Sky (dir. Gavin Hood, 2015). Autonomous drones and drone swarms would operate with different distributions of choices, interpretations, and decisions, but they too participate in a complex assemblage involving human and technical cognizers.

The choice, then, is not between human decision versus technical implementation, which is sometimes how the situation is parsed by commentators who prefer a simplistic picture to the more realistic complexities inherent in the situation. As Bruno Latour argues, changing the means of technical affordances always already affects the terms in which the means are envisioned; ends and means mutually coconstitute each other in cycles of continuous circular causality. Understanding the situation as a cognitive assemblage highlights this reality and foregrounds human and technical cognitions as crucial components in whatever actions are finally taken.

Although the cognitive assemblage approach can provide useful perspectives on ethical issues, it does not by itself answer the urgent question

---

of whether autonomous drones and drone swarms should be developed by the US military and, if developed, under what circumstances they should be deployed. Arguing in the affirmative, Ronald Arkin, a roboticist at Georgia Institute of Technology, envisions an “ethical governor” that would be built into the weapon’s software, requiring the weapon first to determine whether the presumed target is a combatant and then to assess whether the proportionality criteria is met. This proposal strikes me as naïve in the extreme not only because of the ambiguities involved in these determinations but more fundamentally because of the presumption that the weapon’s designers would agree to these criteria. Even if the US military decided to do so, when autonomous weapons designed by other states and nonstate entities fail to incorporate these restraints, would not “military necessity” dictate that the US do likewise?

The issue, then, cannot be resolved through technical fixes but requires considered debate and reflection. A chorus of voices argues that fully autonomous weapons should not be developed and certainly not deployed. Human Rights Watch and the IHRC, in their thoughtful white paper considering the issue, conclude, “the development of autonomous technology should be halted before it reaches the point where humans fall completely out of the loop” (“LH”). The summary emerging from the Stockdale Center’s year-long program on “Ethics and Emerging Military Technologies,” which culminated in the tenth annual McCain Conference on Military Ethics and Leadership, reaches a similar conclusion: “extreme caution should govern the actual deployment of autonomous strike systems.” Even before deployment, however, they write that “we strongly advise against incorporating ‘strong artificial intelligence’ in such systems, which would render them capable of learning and even choosing ends, inasmuch as strong artificial intelligence is highly likely to introduce unpredictability and/or mitigate human responsibility.” Noel Sharkey is more blunt: “Don’t make them fully autonomous. That will proliferate just as quickly and then you are going to be sunk.”

The issue of proliferation is real; already fifty-five countries have the capacity to manufacture or build arsenals of UAVs. Friedman’s appendix

listing UAVs larger than fifty kilograms in use around the world runs to a massive 220 pages. Matthew Bolton puts the issue of lethal force deployed autonomously through UAAs eloquently: “Growing autonomy in weapons poses a grave threat to humanitarian and human rights law, as well as to international peace and security. . . . Death by algorithm represents a violation of a person’s inherent right to life, dignity and due process.”

As these analyses recognize, the possibility of developing autonomous weapons signals a tectonic shift in how warfare is conducted. The masses of humans that required nations for mobilization, as in World Wars I and II, can potentially be replaced by masses of combat aerial vehicles, all operating autonomously and capable of delivering lethal force. UAVs can now be built for as little as five hundred to a thousand dollars, making it possible to field a two thousand vehicle swarm for as little as a million dollars, a sum that emphatically does not require the deep pockets of a national treasury. This enables an entirely different kind of warfare than that carried out by single UAVs, and it brings into view the possibility that UCAAs could be used to launch massive attacks almost anywhere by almost anyone.

This nightmare scenario is worked out in fictional form in Daniel Suarez’s Kill Decision. With a fair bid to become Tom Clancy’s successor, Suarez offers similarly detailed descriptions of military equipment and operations, but he is considerably more critical of US policies and more sympathetic to utopian impulses. Kill Decision’s plot revolves around Linda McKinney, an entomologist studying the warlike weaver ants in Africa, targeted for assassination by unknown persons and saved at the last moment by a black ops force led by the enigmatic Odin. It seems that the shadowy antagonists have marked for assassination anyone with in-depth knowledge of the technologies they are using to create UCAAs, including McKinney because the swarms recognize colony mates and coordinate attacks using the same kind of pheromone signals she found in the ant colonies.

As Odin and his crew (including two ravens) track the UCAAs, it becomes apparent that the drones represent a new kind of warfare. The clarion call comes first in targeted drone assassinations within the US (an eventuality almost certain to arrive some day in real life). At first their nature is concealed by the government, which blames the explosions on

41. See Friedman, Unmanned Combat Air Systems, pp. 69–248.
42. Quoted in Tucker, “Inside the Navy’s Secret Swarm Robot Experiment.”
43. See Daniel Suarez’s Kill Decision (New York, 2013).
terrorist bombs that have targeted seemingly random victims. But a larger plot is afoot, although the perpetrators remain in the shadows. All we know for sure is that the timing is coordinated with an US appropriation bill to create a large fleet of autonomous drones.

The question the text poses, then, is whether autonomous drone warfare on a massive scale is inevitable, given the advantages of UCAAVs over manned aircraft and piloted UAVs. Unlike these, autonomous drones have no limit on how far they can be scaled up and thus are able to mass in the hundreds or thousands. In the novel, the UCAAVS are a colony of ship-attacking drones inhabiting a huge commercial freighter, each autonomous but responsive to the colony’s chemical signals. The freighter is on course to intercept the US fleet as it carries out military exercises in the South China Sea. The predictable result would be the annihilation of US ships and an international incident blaming the Chinese for the attack. It begins to look as if the drone legislation might be approved, plunging the US and subsequently the world into a new era of automated drone warfare. Odin observes that “with autonomous drones, you don’t need the consent of citizens to use force—you just need money. And there might be no knowing who’s behind that money either.”

As the novel suggests, Suarez believes that autonomous weapons must be constrained by international treaty before we are plunged into a Terminator-like scenario in which we are no longer able to control the proliferation of these weapons. He also implies that the only political circumstance in which this is likely to happen is if the US, having exploited its technological advantage in being the first to develop and use drone technology, reaches the stage where drone proliferation by other state and nonstate entities becomes inevitable. In a report issued by the Council on Foreign Relations, Micah Zenko writes, “the current trajectory of U.S. drone strike policies is unsustainable. Without reform from within drones risk becoming an unregulated unaccountable vehicle for states to deploy lethal force with impunity.”

Ironically, the threat of unlimited drone warfare may be the strongest motivation for the US to reform its own drone policies first in order to argue for international accords prohibiting further proliferation. The situation is analogous to the US being the first to develop—and use—nuclear weapons but then being a leader in arguing for international controls when other states acquired nuclear capability.

44. Ibid., p. 261.
However cynically conceived, this strategy did rescue the world from all-out nuclear war. Nuclear weapons, of course, require massive resources to develop and build, which largely limits them to state enterprises. Autonomous drones are much cheaper. Whether the same strategy would work with them remains to be seen.

**Human Emotion and Technical Cognition**

So far my argument has emphasized the ways in which human and technical cognitions interact, but their cognitive capacities nevertheless have distinctive differences. On the technical side are speed, computational intensity, and rapid data processing; on the human side are emotion, an encompassing world horizon, and empathic abilities to understand other minds. Arkin tends to present human emotion as a liability in a warfighter, clouding judgment and leading to poor decisions. However, emotion and empathy also have positive sides; considered as part of a cognitive assemblage, they can make important contributions.

Grégoire Chamayou refers to suicide bombers as “those who have nothing but their bodies.” Applied to groups such as Al-Qaeda and the Islamic State, this is obviously incorrect; they also have AK-47s, rocket grenades, suicide bombs, improvised explosive devices (IED), and a host of other weaponry. There have, however, been instances of resistance by those who indeed have nothing but their bodies: the lone student confronting a tank in Tiananmen Square, the hundreds of satyagrahis (resisters) who followed Mahatma Gandhi to the Dharsana Salt Works in India and were beaten by British soldiers. Intentionally making oneself vulnerable to harm for principled reasons has the capacity to evoke powerful emotions in those who witness it, as world outrage over the episode at the Dharsana Salt Works demonstrated. Vulnerability, whether intentional or not, can also evoke strong emotions in those who perpetrate violence, in some instances leading them to reject violence as a viable solution.

Such is the case of Brandon Bryant, who performed as a drone sensor pilot for the US Air Force for almost six years until he refused to go on, turning down a $109,000 bonus to reenlist. When he finally sought therapy, he was diagnosed with post-traumatic stress disorder. The diagnosis represents, as Matthew Power notes, “a shift from a focusing on the violence that has been done to a person in wartime toward his feelings...”

about what he has done to others—or what he’s failed to do for them.”

Chamayou sees this shift as a cynical tactic by the US military to claim that drone pilots suffer also, but this interpretation fails to do justice to Brandon’s feeling that he has been through a “soul-crushing experience.” Granted that drone pilots suffer far less harm than those they kill or maim, the fact that some of them experience real “moral injury” can be understood as one of the contributions human emotions make to cognitive assemblages—something unique to biological life forms that has no real equivalence in technical systems.

Language, human sociality, somatic responses, and technological adaptations, along with emotion, are crucial to the formation of modern humans. Whether warfare should be added to the list may be controversial, but the twentieth and twenty-first centuries suggest that it will persist, albeit in modified forms. As the informational networks and feedback loops connecting us and our devices proliferate and deepen, we can no longer afford the illusion that consciousness alone steers our ships. How should we reimagine contemporary cognitive ecologies so that they become life enhancing rather than aimed toward dysfunctionality and death for humans and nonhumans alike? Recognizing the role played by nonconscious cognitions in human/technical hybrids and conceptualizing them as cognitive assemblages is of course not a complete answer, but it is a necessary component. We need to recognize that when we design, implement, and extend technical cognitive systems, we are partially designing ourselves. We must take care accordingly. More accurate and encompassing views of how our cognitions enmesh with technical systems and those of other life forms will enable better designs, humbler perceptions of human roles in planetary cognitive ecologies, and more life-affirming practices as we move toward a future in which technical agency and autonomy become increasingly intrinsic to human complex systems.

49. Quoted in ibid.