**What You Are and Its Affects on Moral Status: Godman’s Epistemology and Morality of Human Kinds, Gunkel’s Roboto Rights, and Schneider on Artificial You.**

**Lantz Fleming Miller**

Books Considered:

The Epistemology and Morality of Human Kinds by Marion Goodman.

Oxon: Routledge, 2021.

Robot Rights by David J. Gunkel. Cambridge: MIT Press, 2018.

Artificial You: AI and the Future of Your Mind by Susan Schneider. Princeton: Princeton University Press. 2019.

Thanks to mounting discussion about projected technologies’ possibly altering the species mentally and physically, philosophical investigation of what human beings are proceeds robustly. Many thinkers contend that whatever we are has little to do with how we should behave. Yet, tampering with what the human being is may tread upon human rights to be whatever one is. Rights given in widely recognized documents such as the U.N. Declaration of the Rights of Indigenous Peoples assume what humans are and need depends upon ontological assumptions about human existence.

Godman’s monograph turns to human ontology—just what kind of beings humans are—from an angle distinct from the other two books and how these play into rights. She is concerned with how well the social sciences accord with what we are. Specifically, humans come in various kinds, due to social or physical factors, such as ethnicity, skin color; and a mix of these. Such factors are also central to individuals’ developing and maintaining identity. On the one hand are empirical facts about humans and how they are defined. Such facts allow us to say cogently “There are about one-point-six adherents of religion X in the world.” These typical social/human-science matters can be empirically corroborated with statistics or falsified.

On the other hand: How well can social sciences take into account what human agents individually adopt in developing their identities? They may resist external categorizations, which social sciences rely upon in practice. It seems that social sciences and individual self-fashioning may contradict each other. On the other hand: many social sciences proceed so as to improve the condition of their subjects, who may thereby maintain, in policy and daily practice, those individual identities. To protect identity against categorization, one must categorize one’s subjects. To ensure that the interest group’s members are protected against discrimination, one must classify those group members, despite certain group members’ personal identities. What may be hardest to accept is that some categorizations have been invented by parties who have little respect for those peoples. By acceding to such categorizations, one is effectively endorsing such discriminatory beliefs. Categorization “appears associated with practices of exclusion, stereotypes and misguided generalization…. Can this seemingly inescapable use of human categories be explained by something prior to their use?” (1)

The challenge here is brought out in her conceding that:

1. “How empirical knowledge is … possible in the humans sciences depends on existence of human kinds.

2. “The [human sciences’] quest for knowledge… often arises for… moral and political interests.

3. “One should [explicitly justify] scientific inquiries that are of a moral nature.”

4. “Many… scientific and philosophical projects regarding human kinds [may] be of emancipatory nature.” (18)

Thus, acknowledging the central role and study of human kinds is consistent with social sciences’ pursuits, even if not necessary to them.

Godman’s preferred approach to reconciling the social-science/personal identity dichotomy is that of historical: kinds. This approach proposes that the nature of human cultures is to reproduce themselves: “humans reproduce their genetic material but also a range of behavior and pre-existing culture.” (2) A crucial concept is that of the channeling of cultural practices from a “pre-exiting culture” to the current culture, thence to the culture’s daily practices. Hence, the ongoing historical process of kinds.

Godman assesses other philosophical accounts of human kinds: multiple realization not reducible to physical kinds, functional kinds, social ontology, homeostatic mechanisms, looping effects, and naturalist social construction. Godman finds these inadequate for grounding human kinds. She then justifies historical kinds in general, the salient feature being that “a kind is reproduced because there are chains of reproduction from certain individuals that stand as models for other individuals. That chain of reproduction to common models is also the common cause of kinds.” (55) Thus a literary work like Flaubert’s *Salammbo* is multiply realized in chains of reproduction, including different editions and translations, from the original model. “The common cause is a common lineage of reproduction.” (45) Godman holds that a “*reproductive lineage* among instances of a historical kind” is necessary (47, emphasis in original). Continuity in reproduction over time must be ensured, extending from past to present and presumably into future, establishing a lineage. She further defends the approach by describing how it copes well with social-science injunctions to explain and individuate. Importantly for the book’s theme is how general historical kinds can account for specifically *human* historical kinds.

Godman provides three cases vis-à-vis historical kinds: gender, cultural, and ethnicity/race kinds. The issue of gender-kinds involves the centrality of our highly social species: Humans do not just learn from others but have a strong need to belong, as via gender. It seems that “the intensive and extensive roles that gender plays in many areas of life… are likely to matter to the individual.” (61) Some cultures allow alternatives to binary gender, such as the Zuni We’Wha gender among Zuni. Godman observes it is “plausible that some real, or in the very least perceived, sex differences played an important role in the original emergence of gender models.” (63)

Godman applies the human historical-kinds approach to inquiring what is a culture and how is it realized. Although cultures have no physical existence, they do have physical manifestations, as in artefacts, which help signal what those cultures are and how these may be delineated by historical kinds. Moral issues come into play in ethnicity/race, as problems such as “labeling” influence natural kinds. “Labeling effects… contribute to the transformation of what it means to be a kind.” (87) Historical kinds can also bear on moral, and policy matters. as in indigenous peoples’ group rights.

Gunkel homes in on a more specific moral topic than Godman\s general concern with identity’s role in rights: automata moral status and rights. He enjoins us to inquire into what human rights are before we take on the new technologies’ challenge to it: May humanlike automata (“robots”) exhibit sufficiently humanlike traits to merit some of humans’ moral status? The inquiry has two components: What are rights? What are automata? Discussions of what (human) rights are date back at least to Cicero and Aristotle. Automata, too, have been around since then, as seen in the metallic Talus of Greek myth. In his *Politics*, Aristotle discussed the idea of an artificial servant that might circumvent its owner’s possible ethical qualms. Discussion of whether such devices should merit rights has been notably rare until now. What is new is coupling these two components—“rights” and “automata”—to form the study of automata rights.

“Automata” and “rights” are commonly defined via reference to human beings and their rights. An automaton is considered a human construction to help users solve problems. A (commonly human) right is considered to be a means to empower and privilege human agents by recognizing and protecting claims and immunities. Such rights are established in terms of the individual, in contrast with others or the government. If other kinds of entities merit rights, they and their presumed rights are defined in contrast with those of human beings.

An obstacle for granting nonhuman entities rights is that these ultimately depend upon rights concepts developed by humans, geared toward human needs and assuming what kinds of things humans are. Humans are autonomous bipedal social beings who in fact are so social they could hardly survive, much less thrive, without others. This fact only underscores their equal needs for autonomy in decision-making, thus the freedoms of life, liberty, property, and pursuit of happiness. Yet, these freedoms combined are evidently unique to humans. Other presumed candidates for rights must be viewed through a lens created by and for humans, thus conceptually subservient to them. Some investigators may find not a challenge here but a boon to understanding the rights due to nonhuman entities: Humans thereby recognize the degree of rights they say are due to them for their type of entity. Hence nonhumans merit a degree of rights—perhaps not the same as humans—which they must appropriate to their own ontology and bases for wellbeing.

Gunkel elegantly deals with how and whether we can understand human rights so as to apply them to nonhuman entities, particularly automata. First, consider his views on definitions of “rights” and “robots.” He finds their definitions elude definition. We rely on intuitions nurtured by cultural media, including science-fiction. As for rights, he takes the Hohfeld approach proposing four rights: privileges, claims, power and immunities. Gunkel provides two basic theories of who deserves rights: A “will” theory “requires that the subject of a right possess the authority and/or capacity to assert the privilege.” (31) Interest theories, by contrast, hold that the subject of right has interests that a right can preserve. Either theory tends to a different scope of application. Will-theory tends toward a restricted scope, perhaps applicable only to humans, whereas the interest scope tends to include about everything alive.

Gunkel asks whether automata are such that they can have rights; or, whatever they are, they should have rights, as Figure 1 illustrates.

|  |  |  |
| --- | --- | --- |
|  | *Should not* | *Should* |
| *Could not* | 1. Automata could not have rights and should not have rights (Chapter 2) | 2. Automata could not have rights but should have rights (Chapter 5) |
| *could* | 3. Automata could have rights but should not have rights (Chapter 4) | 4. Automata could have rights and should have rights (Chapter 3) |

Figure 1. Chart showing Gunkel’s four perspectives on whether automata could or should have rights.

Gunkel dedicates a chapter to each of these four perspectives, in the end finding all four wanting. Chapter 2’s option falters primarily because of the presumed emphasis on ontology preceding ethics. As Gunkel quotes Floridi, “‘What the entity is determines the degree of moral value it enjoys, if any.’” (140) Giving ontology precedence over ethics is the implicit violation of the Humean “is/ought gap.” Yet, it is unclear whether Hume says the gap is unbridgeable. He states moralists problematically too often do not provide enough justification for their leap, not that they shouldn’t leap. Furthermore, to give the ethical question traction, we must make weighty assumptions about the ontology of the moral agent or patient. We presume humans are rational, that gorillas are sentient, that a forest is an organism. Each of these ontologies feeds into pertinent practical philosophy. It is insufficient to dismiss the entirety of ontology’s role in morality without directly analyzing that role.

Chapter 3, assesses the fourth option. Which qualities give humans’ moral status?—sentience? intelligence? We end up presuming which traits of an entity warrant rights. However, why should such traits warrant rights? Because those traits that humans merit intersect with those of another kind of being? It is also unclear whether the entity’s traits which fall outside that intersection would make a difference in its rights desert. Future automata, for example, may approach the presumably cognition-originated behaviors of a mouse, so the device may merit murine moral status. One may extend such reasoning to automata that exhibit presumably humanlike cognitive-derived behaviors. Functionalism may say: “Precisely!” Yet, functionalism is far from a widely accepted faith.

Chapter 4’s option holds that automata could but should not merit rights. Automata should be our slaves even if they were sentient; human lives are the concern. Maintaining automata as mere tools should help ensure this relation. Gunkel sees this view as unrealistic measured against the “moral intuitions and practical experiences of users” and designers (130) who find deploying automata ethically sound. Yet, Gunkel does not justify why such intuitions are valid.

By Chapter 5’s Option 4, automata should have rights approaching humans’ whether or not automata merit such rights. This Kantian option holds that treating animals cruelly increases our chances of being immoral. Gunkel contends that the sentimentalist version of this argument saves it from the “is/ought” problem. But he objects that, like the other three options, this one still builds upon what is important for *us* rather than for the automata.

The four options bring up the issue of whether meriting rights depends upon what kinds of being *we* are. Perhaps one must go into the automaton’s world, say via Levinas’s work on “The Other.” While Levinas does not explicitly apply his views to technologies, Gunkel contends Levinasianism can illuminate this moral-status issue. The main Levinasian concept here is that ethics precedes ontology. Whatever kinds of being we are is subservient to ethics. Establish ethics, *then* see how it plays into human ontology. The Other is sensed, acknowledged, known only through the fact of differences. A thing is what it is because it is not what the other is—via a manifold of differences, not mere dualisms. There is an eventual state of “irreducible otherness.” (164) One is an Other among Others. Machines, in Gunkel’s view, can be such others. If ethics does indeed precede ontology, we need go no further in invoking ontology for determining ethical issues. The Other makes ontology possible. Exposure to the face of automata may be essential to formulating arguments about their moral status.

Some readers may find “ethics precedes ontology” obscure, along with the elusive notion of a being as Other. Striving to circumvent ontology to establish a being’s moral status in Leveniasian terms may encounter a tough practical challenge. Gunkel’s reason for doing so—to circumvent the human-centeredness of the four options—is well- appreciated. The question remains whether it is necessary to circumnavigate the human-centered, ontological approaches to solve such moral problems as automata rights.

Schneider’s *Artificial You*, while not explicitly about rights, poses a rights quandary. Certain individuals with reputable resources and practical backing explicitly seek to evolve the species as they see fit may. This matter may seem highly conjectural, since no currently available method may so alter the species. By what right can technicians justify changing what we are? Early on, Schneider proclaims herself a long-time “transhumanist” (p. 14), yet is highly skeptical about most proposed life-extending techniques, for practical and often moral reasons. Primarily, the monograph critically reviews ongoing research, notably digital-based technologies: consciousness engineering; machine consciousness; human merging with AI; mind uploading; superintelligence; and preserving the brain/mind indefinitely. She upholds bio/machine melding and biogenetic engineering as more viable than digital methods of life prolongation. Regrettably, she scarcely discusses these, much of her criticism concerning their technical shortcomings—the unlikelihood they would work. But she assesses such technologies ethically, especially privacy loss, slavery, and the machines’ threat to human existence. All of these are rights concerns.

One ethics worry is machine consciousness. Having an inner life may play into determining an entity’s moral status and rights. A device lacking sentience or awareness, not experiencing its existence through the senses, would not meet the qualifications for meriting basic human rights. A nonconscious entity such as a bicycle needs no liberty, self-expression, property protection, or assembling with whom it wants. Similarly for a laptop. Nonetheless, machine consciousness remains a possibility. Such a device would have a greater moral status and warrant for certain rights than the nonconscious machine. Schneider concurs that such a device, if sufficiently intelligent, autonomous, and powerful, could threaten human life. But she also anticipates that consumers may demand conscious machines without such deleterious traits: Conscious machines could be more empathetic or make better companions.

Because of potential dangers, she recommends the Precautionary Principle in developmental stages: Researchers strive to preempt and circumvent right-endangerment. The primary concern is that, without an irrefutable test for consciousness, we cannot know whether a consciousness-engineered device in development is indeed conscious or to what extent. Certainly, as in military development, competition may direct targeted-outcomes. But for the most part, Schneider recommends “the excluded middle” (68), whereby researchers produce only conscious machines whose moral status is clear. Another caution is that even the best tests for consciousness may leave that moral status unclear. A further matter remains: the *device’s* interests. Research-and-developing a sentient device without its informed consent may work against that device’s status.

Among the ethics problems, Schneider sees threats of privacy infringement and possible (machine) slavery, from more and more powerful AI. AI-driven algorithmic devices may allow companies to mine enormous amounts of information about individuals from widely available and legally obtained data. Even if such AI guesswork is not 100 percent accurate, it could be accurate enough to harm the moral patient. Potentially false information could be as detrimental to the individual, or more so, than accurate data.

Such playing with fire in peoples’ lives constitutes a use of humans as objects. Unfortunately, Schneider gives minimal attention to slavery as an AI moral violation. The prospect of freedom violations, by contrast with those of privacy, is a threat to potential rights of the machines rather than its users’. Engineering consciousness, while programming the machine to undertake tedious or dangerous tasks, may constitute slavery. Or would it be slavery, if the machine is perfectly at-home doing its labor? Such a scenario may reek of Southern U.S. plantation owners’ illusions of the happy slave. Dealing with hi-tech slavery needs thorough investigation into what kinds of entities we humans are, to merit full rights, and into what kind of beings the projected AI or automata are.

The book offers few new perspectives but, being accessible, its subject matter should attract a range of readers, professional and amateur. It is bold in taking positions contrary to the mainstream of techno-enthusiasm. The human being may not be stretched indefinitely as some enthusiasts maintain. There is something that humans are that precludes such an outcome. For Schneider that entity is basically biological.

These three books build upon a needed, and far from complete, understanding of what is a human being. Schneider’s program for industrialists, within some practical (and theoretical) limitation, allows reshaping the human form as clay. Still, it must have a viable notion of what is the thing being remodeled. Gunkel’s keen investigation into whether automata could have rights much like humans’ explicitly calls for well-understanding what are humans and what are rights. Godman, refraining from anticipating global autocratic change in the human sphere. is less ambitious than Schneider or Gunkel. She merely asks human sciences to inquire freely into what humans are, while allowing for improvement in human beings’ social environment.

If there is urgency among these works, it is that we need: a sharper idea of what we are, how we should act and be treated, and how our rights fare among potential human improvements. We may better assess whether and how to make such improvements, ideally in fairness to all.