Works Cited

Campbell, Keith *Body and Mind* 2d ed Notre Dame, Ind University of Notre Dame Press, 1984

Horgan, Terence, and John Tienson, eds *Connectionism and the Philosophy of Mind* Dordrecht Kluwer Academic Publishers, 1991

Jacquette, Dale *Philosophy of Mind* Englewood Cliffs, N J Prentice Hall, 1994

Lillegard, Norman "No Good News for Data " *Cross Currents* 44, no 1 (Spring 1994) 28–42

Phillips, Hollibert E *Vicissitudes of the I An Introduction to the Philosophy of Mind* Englewood Cliffs, N J Prentice Hall, 1995

Plantinga, Alvin *The Nature of Necessity* Oxford Clarendon Press, 1974

Searle, John "Minds, Brains, and Programs " *The Behavioral and Brain Sciences* 3 (1980) Reprinted in Douglas R Hofstadter and Daniel C Dennett, eds *The Mind's I Fantasies and Reflections on Self and Soul* Toronto Bantam Books, 1981 Pp 353–73 Citations are from the reprint

Von Eckardt, Barbara *What is Cognitive Science?* Cambridge, Mass MIT Press, 1993

Wesley J. Wildman

But Consciousness Isn't Everything

The Spring 1994 issue of *Cross Currents* included an article by Norman Lillegard entitled, "No Good News for DATA." The reference, of course, is to the android of *Star Trek: Next Generation* fame. I'm quite fond of Commander Data, myself. He moves me to struggle to be more human, just because becoming more human is his life project. Professor Lillegard, however, refers to Data as "it." In fact, I appear to be a member of the cultural hordes Lillegard thinks have sold their distinctively human birthright in a pathetic attempt to latch onto the latest and greatest in technology to explain themselves to themselves. We are like complex plumbing one century, clocks and steam engines in another, computers and androids in our own time. So we say to ourselves, anyway, and not without some pride and amazement — though our en-

WESLEY J. WILDMAN, assistant professor of theology at Boston University, writes on a wide range of topics in the specialization of science and religion.

thusiasm for each model is as short-lived as its stay on the cutting edge of technology.

Lillegard doesn't think this trend of superseded analogies is about to be brought to an end by the latest model for wondrous human being: an android. On the contrary, he wants to remind us that we humans are not so easy to model. Lillegard's title expresses this compactly: the good news of the Christian gospel does not apply to any of the technological self-images with which we develop passing infatuations.

Nice point. Lillegard's list of hopelessly inept, yet popular, models of human persons supports his gem of wisdom. He strikes a blow for human uniqueness, and debunks one more infatuation with a model seriously ill-suited for the service into which we press it. I can live with that. I'll be sad to look upon Data as even more of a misleading literary fiction than I had thought, but glad to be assured of my importance in the limited sense that I resist modeling.

More than this, however, Lillegard offers an argument that a whole class of models for human persons are too seriously flawed to be of much use. This amounts to a prediction that Data's future literary and scientific progeny will do no better as models than Data himself (itself). The models Lillegard is aiming at all espouse functionalist theories of mind. That is — and Lillegard might have been clearer about this — they tell a causal story about how mental states relate to sensory inputs, behavioral outputs, and other mental states, in such a way as to avoid having to worry about some of the trickier aspects of mental states, such as how they might *feel* to the being whose they are.

Lillegard argues that such functional stories are contaminated by a kind of residual dualism, since there is always the possibility of transferring all of those causal relationships from a brain to, say, a traditional computer, a neural net, a Hoyle-like black cloud, or something else. This is uncontroversial, and is why the view is sometimes described as "black box functionalism" — the physical substrate for the causal relations is needed, but not essential in its specifics. The essence of the human person, for this (Lillegard's) kind of functionalism, is the transferable pattern of causal relationships. The physical substrate on which it depends for its realization is contingent, because it can be exchanged in principle for some other physical substrate without loss of information or function.

Now why would Lillegard complain about such a view of the essence of the human person? Well, ask yourself, would you think of yourself as essentially the same person after your causal, functional relationships

had been implemented in another body or, worse from Lillegard's point of view, in Data's android substrate? Maybe you would. More pity you, then, according to Lillegard, because the uniqueness of you is expressed in the concrete, bodily realization of your potential, in the actual nexus of particular bodily threats, agonies, and aspirations that you are. All of this includes *your body,* as well as your mental states; both are essential, and probably finally indivisible. Lillegard's conviction in this regard is informed by the biblical perspective that human essence is always embodied soul, a perspective that he expounds with some care.

While Professor Lillegard makes some fine points in his essay, he also leaves himself vulnerable to criticism. Professor Sennett makes those criticisms precisely and judiciously in his enjoyable "Requiem for an Android: Response to Lillegard." It is just because of this contention that the editors of *Cross Currents* asked me to write these remarks. Now, because this note appears in the same issue as Sennett's article, there is no need to summarize his reply. Instead I will do what I hope is most useful: evaluate Sennett's criticisms.

Sennett is evidently a well-trained philosopher of religion of the analytical persuasion, and we should be appreciative of the clear criticisms that his competence engenders. But we should also remain wary of one of the potential weaknesses of detailed criticism, namely, that it can obscure on the large scale even as it illumines on the small scale. For example, it is crucial to ask whether Lillegard's position can be reexpressed so as to avoid the worst of Sennett's criticisms while remaining true to Lillegard's own central insights. If Lillegard's argument could not be reexpressed in this way, then I could perhaps enjoy Commander Data with less ambivalence, thanking Sennett for delivering me from the clutches of Lillegard's unjustifiable obsession with human uniqueness. But if Lillegard's argument can be reexpressed, then we might be forced to conclude that Sennett's incisive prosecution doesn't amount to much after all, being useful only to help Lillegard tidy up some loose ends and improve the focus of his case.

To my way of thinking, Sennett scores hits with most of his detailed criticisms. However, because his initial characterization of Lillegard's article half misses the backbone of the argument, it is not so difficult for Lillegard to apply a few bandages to a partially sound structure, make a few adjustments here and there to avoid cheap hits in the future, and then proclaim that he is ready to fight another day.

From my earlier exposition of Lillegard's case, it follows that I would schematize it differently than Sennett does:

1. The character Data conforms to a functionalist model of mind.
2. Functionalist models of mind assert that the essence of the human person is a set of causal, functional relationships requiring a — but not a *particular* — substrate; this is residual dualism.
3. Residual dualism is inconsistent with the biblical conception of the essence of the human person.

Therefore,

4. The character Data does not conform to the biblical conception of the essence of the human person.

So described and organized, Lillegard's case avoids a number of Sennett's criticisms without betraying its own character in the least. Particularly odd in Sennett's description of (3) is the reference to materialism; Lillegard only argues that the biblical view contradicts dualism of all kinds (including residual dualism), not that it affirms a materialist theory of mind. Otherwise, however, Lillegard does leave his argument prone to being summarized in the way that Sennett does. It is one of those cases where the critic might have done well to look with more generosity for the strongest case to criticize, rather than settling for the case most forthrightly (albeit imprecisely) stated.

The disagreement between Lillegard and Sennett over (1) stems from Sennett's narrower, and Lillegard's broader, definition of functionalist theories of mind. On the definition of functionalism Lillegard probably ought to have given, and seems most often to intend (the black box functionalism defined above), Sennett's claim that an analog connectionist system is not a functionalist system does not hold. Moreover, Sennett seems willing to grant (2), at least as stated here, so the remaining contention hovers around (3).

I failed to detect in Lillegard's article the ghost of Sartre or his ilk that, evidently, Sennett thought was haunting its pages, though I could appreciate how certain of Lillegard's phrases might suggest such a reading. I wonder, therefore, whether Sennett was hasty to suppose that Lillegard rejects the conception of a resurrected body. Lillegard seems far more open-ended on the question — somewhat agnostic, perhaps. At stake here, of course, is whether residual dualism is such a bugbear: if you experience continuity of identity between your current body and your resurrected body, isn't that a kind of transfer of your essential being from one substrate to another, à la residual dualism? Sennett says yes, and so denies (3). Lillegard seems to want to say no, though not

because of a rejection of the resurrection so much as because of an implicit conviction that this change would not really be a new substrate, but only a transformed one. But Lillegard's article, unfortunately, is not completely clear on this point.

Anyway, if Lillegard is right that the essence of a human person can never be transferred between different substrates, then he makes his case: Data would not conform to the biblical conception of the essence of the human person. That does not mean much with regard to Data, however, since the Bible's conception of the essence of the human person itself needs to be evaluated for its adequacy, an inquiry that Lillegard does not undertake in his article. Similarly, if Sennett is correct that Christian doctrine calls for moving the essence of a human person among various substrates, then Lillegard's inference to (4) fails. But (4) might still be true; it might simply need to be established differently.

Sennett's summary of Lillegard's argument does not recognize clearly enough that it is two-pronged. Lillegard's second (more compressed and more complex) line of argument is somewhat ethical in character. It hinges on "parochiality" — the concrete, personally contextualized "feel" of psychological predicates like fear and love — and can be summarized as follows:

1. The character Data conforms to a functionalist model of mind.
2. Functionalist models of mind are formally neutral to the question of the parochiality of psychological predicates.
3. Formal neutrality to the question of the parochiality of psychological predicates makes impossible valid human self-identification and self-description.
4. Failure to maintain the conditions for valid human self-identification and self-description is a betrayal of the moral essence of the human person.

Therefore,

5. The character Data conforms to a model of mind that, if applied to human beings, would be a betrayal of the moral essence of the human person.

It may be because Lillegard relies on this second argument that he feels no need to press the adequacy of the biblical conception of the essence of the human person, which was the missing last step of his first argument. Sennett does not adequately address this second argument, but it

is a much trickier case to make than the first, which Sennett discusses at length.

In concluding, note that an argument considered by neither protagonist throws the issues surrounding this debate into a slightly different — and, to my mind, a helpful — light. Bypassing Commander Data for now, the steps in the argument are as follows:

1. Being a person requires the capacity for moral choice (this is uncontroversial on any sound definition of person).
2. Moral choice requires consciousness (again, fairly unproblematic; in fact, self-consciousness may be required for moral deliberation, which is a stronger claim).
3. Consciousness requires a complex biological system, such as one with a central nervous system, and an intricately organized brain.

I think (1) and (2) are stable enough propositions that this argument helpfully draws attention to (3) as being close to the heart of the debate in which Lillegard and Sennett are engaged.

The upshot is, of course, that the conditions necessary for consciousness are thereby (*some* of the) requirements for being a human person. If consciousness is a property exclusively of biological systems, then no matter how good nonbiological androids get at simulating human behavior, they will never be conscious, nor persons in the ordinary moral sense. If consciousness is potentially also the property of other kinds of complex systems, such as positronic brains, then positronic brain androids may well have to be acknowledged one day as exercising moral choice, and perhaps being persons. Precisely what can be a substrate for consciousness is partly a question for cognitive science and biochemistry, though the question is infamous for being almost impossible even to frame in these spheres of research.

In an era in which some cognitive psychologists write books with "consciousness" in the title but little or nothing about consciousness between the covers, it might be politically incorrect to think about Data (or ourselves) in these terms — though people not familiar with cognitive psychology may wonder what the fuss is about. Consciousness, however, cannot be marginalized in these debates, even if it is difficult to handle. The price to be paid for such self-deception is the devastating consequence that Lillegard warns us against: losing track of something essential about ourselves.