Critical Humanism: Theory, Methodology, and *Battlestar Galactica*
Patrick Thaddeus Jackson, American University

It seems relatively uncontroversial to assert that when social scientists set out to conceptualize their objects of study these days, one of their first stops when looking for models and metaphors involves the natural sciences. Thus we get "punctuated equilibrium," "complex learning," and other notions derived from evolutionary theory; we get notions of "balancing" that clearly display their mechanistic origins; and we get models of individual decision-making that rely on cognitive psychology and even neurology.[1] Our conceptual stock seems to have been, for the most part, seized in raids on the fertile fields of the natural sciences. Indeed, it is almost as if we expect that the intellectual parentage of a natural-scientific notion will magically imbue our theories with explanatory power.

One way to justify this practice is to embrace physicalist reductionism: the notion that what is established in a more "basic" science like physics or biology or psychology will necessarily provide a good starting-point for those of us operating in the "special" sciences like IR. Unless we want to go all the way with physicalist reductionism, however—and few of us seem to want to do so[2]—there is precisely *no* reason why the fact that a notion like "punctuated equilibrium" or "balancing" serves an explanatory

---

[1] The use of these latter notions in IR, though, might also be warranted by a critical realist methodology of using a laboratory to adduce causal powers that are then presumed to continue to operate in the open system of the real world. If used in that sense, the issue is not whether the notion transfers its explanatory powers across domains; it is whether the laboratory effectively discloses real-but-undetectable dispositional properties. See Jackson 2011, Chapter 4.

[2] Some of Wendt's most recent work on "quantum consciousness" does, in fact, pursue this reductionist line, and maintains that consciousness—and therefore, the social relations built on top of consciousness—must ultimately be consistent with and reduced to physics.

function in one domain should have *any* implications for what that notion does in another domain. Our privileging in IR and the other social sciences of models and metaphors drawn from the natural sciences seems, then, a mere prejudice, perhaps fueled by the dream of borrowing some prestige from a more culturally esteemed domain.

After all, we do this when we are thinking about research design and execution, so why not do it for substantive notions too? So many of our pieces of methodological advice are based, more or less explicitly, on parallels with what supposedly happens in the natural sciences, that a neutral observer might be easily forgiven for diagnosing a kind of field-wide inferiority complex vis-à-vis physics and biology, and even vis-à-vis paleontology (as in Van Belle 2006). The manifest evidence that these efforts to obtain reflected glory from the natural sciences rely on seriously misleading accounts of what natural scientists actually *do* when conducting research does not serve to shake the boundless enthusiasm of those IR scholars seeking to ground their knowledge-production efforts in one or another universalist—one might even say *fundamentalist*—set of commitments that are supposed to ensure that finally, this time, we're going to get it right and be able to place IR on a proper intellectual footing. Here again, the prestige of the natural sciences serves as the often-implicit justification for conducting research in one way rather than another: natural science, actual or imagined, is a privileged domain from which we social scientists more or less have to borrow in order to make any real progress.

Needless to say, I am not persuaded of the value of these practices. My issue is not that models of research-practices and conceptual metaphors drawn from the natural sciences might not prove fertile when applied to the study of world politics; indeed, they might, but if they did, that would have precisely *nothing* to do with their origins in physics or biology or wherever. Instead, they would have to prove their explanatory worth in *our* field, according to whatever standards we find appropriate for cashing out "explanatory worth."[3] Karl Popper was famously indifferent to the origin of scientific notions, and in this (although in little else) I would consider myself something of a Popperian: scholars should use notions derived from religion, from poetry, or even from physics and biology, but should subject them all to standards of evaluation proper to the field into which they have been imported. Why restrict our theoretical vocabulary to the natural sciences, when there are wider vistas to be explored?

Analytical practice

In fact, when it comes to the study of social and political phenomena, I would submit that novelists and other creators of fictional products might even be a better source of models and metaphors than the natural sciences are. The act of creating a fictional world is a complicated balancing-act, in which the audience has to be induced to suspend their disbelief in the untrue aspects of the fictional world while clinging to

---

[3] The fact that there are multiple such standards in IR—that there is a plurality of logics of inquiry, each of which is irreducible to the others—complicates the story without affecting the central point at issue.

those aspects that command acknowledgement or recognition. This "political economy of production and consumption" (Jackson and Nexon 2003) applies perhaps even more vigorously to serial fiction, because as the series goes on, the audience may become ever more vigilant about policing the internal consistency of the fictional world, and they will have more opportunities to evaluate emplotted events against their own experiences and sensibilities. Keeping these various aspects of the production in sync is a challenge, and it stands to reason that those creators and producers who successfully create compelling fictional worlds have the kind of insight into their audiences that our mechanical models and adventures in large-n data-mining can rarely if ever achieve.

Indeed, the position of the creator of a fictional world is not unlike the position that Max Weber described as characteristic of the scientific investigator.[4] Scientists, Weber maintained, operate with ideal-types rather than with purely faithful representations of an externally-existing world. Although most of us have seen this before, it is worth quoting at some length Weber's definition of an ideal-type as given in his essay establishing an editorial policy for the journal he and two colleagues had just assumed control over (Ringer 2004, 78-79, 107-108). Rather than "a 'presuppositionless' copy of 'objective' facts," ideal-types are
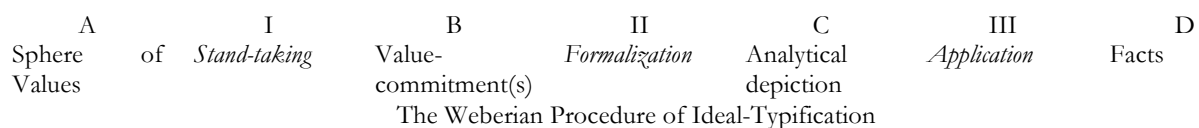
> formed through a one-sided *accentuation* of *one* or *more* points of view and through bringing together a great many diffuse and discrete, more or less present and occasionally absent *concrete individual* events, which are arranged according to these emphatically one-sided points of view in order to construct a unified *analytical construct* [*Gedanken*]. In its conceptual purity, this

---

[4] The next few paragraphs remix material from Jackson 2010, Chapters 1 and 5.

analytical construct [*Gedankenbild*] is found nowhere in empirical reality; it is a utopia (Weber 1999, 191).

Weber goes on to point out that "whoever accepts the standpoint that knowledge of historical reality should or could be a 'presuppositionless' copy of 'objective' facts will deny any value to ideal-types" (*ibid.*, 192-193). He himself clearly does not accept the notion of a presuppositionless copy, instead placing the human "capacity and the will to deliberately take up a *stance* towards the world and to lend it a *meaning*" at the center of his reflections (*ibid.*, 180). "The quality of a process as a 'socio-economic' event is not something that inheres 'objectively' in the process as such," Weber argues; this is a rather surprising statement to make in the editorial introduction of a journal devoted to what we would now probably call the analysis of political economy. "It is far more conditioned by the direction of our knowledge *interest* as it arises from the specific cultural significance that we attribute pertaining to the process in an individual case" (*ibid.*, 161). In this way, the social sciences are *productive* of the world of facts, beholden not to some externally existing set of objects or their essential dispositional properties but rather to the cultural values that define and orient the investigation from the outset.

Indeed, the whole procedure of ideal-typical analysis is all about the transmutation of cultural values into useful analytical tools. Weber's overall procedure can be summarized in the following diagram, in which three distinct intellectual moves link four elements:

A	I	B	II	C	III	D
Sphere of Values	*Stand-taking*	Value-commitment(s)	*Formalization*	Analytical depiction	*Application*	Facts

The Weberian Procedure of Ideal-Typification

Reading from left to right, the procedure begins in (A) the concrete sphere of values and purposes in which the researcher is concretely located. Ideal-typification begins with that researcher (I) taking a value-laden ethical stand, by which she or he locates him- or herself with respect to the values and norms in circulation in her or his social context. The (B) value-commitment(s) contained in this stance are then (II) formalized and idealized, in part by blending them with empirical observations in order to create limiting-case representations, thus producing an analytical depiction consisting of one or several ideal-types (C). Then that analytic is (III) consistently applied to specific empirical cases in order to produce (D) facts: what David Easton (1953, 53) refers to as "particular ordering[s] of reality in terms of a theoretical interest." These facts are both dependent on and distinct from the value-commitments that ground the ideal-typical analytic in the first place: "dependent on" because particular factual knowledge-claims are generated by a conceptual apparatus with its roots in specific value-commitments, but "distinct from" because it is possible to *scientifically* evaluate a given piece of research and the facts it produces by focusing our attention on intellectual moves II and III, and essentially ignoring intellectual move I and the specific contents of element B.

Several things are noteworthy in this description of the process. First, ideal-types are nothing like pictorial representations of objects or processes; they are more like deliberate caricatures or partial sketches, or perhaps specialized conceptual filters that focus our scholarly attention on particular aspects of actually existing things to the detriment of other aspects of those same things (Drysdale 2007, 43-44; Gunnell 2007, 67-68). The value-commitments at the core of an ideal-type ensure that an ideal-type,

whether "charismatic authority" or "liberal democracy" or even "public opinion,"
necessarily functions as a way of expressing values even as it calls attention to specific
features of the actual world and gathers them together under one conceptual heading.
Second, this deliberate slant is less due to any putatively dispositional characteristics of
the object under study, and more to the "emphatic points of view"—what I have labeled
"value-commitment(s)" (B) in the diagram above—which, in a sense, *direct* us to focus
on particular aspects and not others.

From this it follows, third, that a different researcher, formalizing different value-
commitments into a different analytical depiction—or even the same value-
commitments into a different analytical depiction—might well focus on different
aspects of the same entity or object, and they would not in any simple sense be "wrong"
for doing so (Weber 1999, 192). Finally, since ideal-types cannot be falsified as one
would falsify a hypothesis—comparing an ideal-type to the actual existence of the
object the ideal-type was derived from would *invariably* "prove" that the ideal-type was
descriptively deficient in some respect[5]—the only meaningful way to evaluate whether
an ideal-type is a good one or not is *pragmatically*: i.e., to examine whether, once
applied, the ideal-type is efficacious in revealing intriguing and useful things about the
objects to which it is applied. This last observation might be thought of as the Weberian
equivalent of what John Dewey called "the pragmatic rule"—"in order to discover the

---

[5] Along these lines, note that every theorist of bureaucracy who begins their analysis with a ritualistic pronouncement
that actual bureaucracies don't function like Weberian ideal-typical bureaucracies is, quite bluntly, *missing the point*. Of
course real bureaucracies don't look and function precisely like their ideal-typical conceptual limits. It would be quite
surprising if they did, and it would most likely also mean that the supposed "ideal-type" was in actuality no such thing.

meaning of the idea ask for its consequences" (1920, 163)—translated into the sphere of scientific inquiry. A good ideal-type is a useful ideal-type, an appropriate means to the analytical end that animates the scholar's scientific activity. Hence it is quite literally nonsensical to speak of an ideal-type itself as being "valid" or "invalid," because in the sphere of scientific analysis these terms cannot be applied to analytical constructs—only to applications, and then only in a technical sense.

In this way, ideal-types can be regarded as provisional idealizations of our value-commitments; they are not simply pure types, but remain always and inextricably wrapped up with cultural and moral commitments (Duvall 2001). But it does not inevitably follow that all researchers inhabiting the same culture or tradition of scientific inquiry will generate or be compelled to use the same set of ideal-typical instruments; our actual cultures and living traditions are ambiguous and flexible enough to be idealized and formalized in different ways, unless and until homogeneity is *imposed* on them, either from within or from without (Shotter 1993a, 156-159; Joas 1997, 162-163). Rather, it is the act of selecting and formalizing one's value-commitments in the process of forging one's ideal-typical tools that affords the construction of valid scientific knowledge, which means that it is always scientifically appropriate both to inquire after the value-commitments encoded into any given ideal-type and to question the way in which a scholar has idealized her declared value-commitments—to ask about the content of element B and the dynamics of move II in the diagram above. But it is not scientifically appropriate to question the specific values that a researcher holds, although it is of course ethically and even politically

appropriate to do so. Recognizing and foregrounding value-commitments, and critiquing the way that they are idealized into pieces of conceptual equipment, are logically distinct activities from the enterprise of criticizing someone's values.[6]

So what does any of this have to do with *Battlestar Galactica*? I want to suggest that serial fiction operates in a way that is eerily similar to the way that Weber proposed that social science be conducted.[7] This is particularly true of serial *science* fiction, because of its peculiar suspension between actuality (it is arranged around a scientific core, even if that core itself contains unrealistic or even by-current-science impossible notions like faster-than-light travel) and unreality (it is, after all, fiction). The "science" in science fiction means something like what Weber referred to as the disenchantment of the world: we may not know how particular machines operate, but we know that we could find out pretty easily, and that the answer wouldn't involve magical incantations or "mysterious, unpredictable forces" but would instead involve comprehensible and calculable processes of control (Weber 2004, 12-13). In this way, science fiction

[6] The precise character of the transmutation of value-commitments into ideal-typical analytical depictions also underpins Weber's logical claim that science cannot possibly resolve value-questions. Because there can't be definitively correct values or value-commitments, the only things that science can say about any given value-commitment are that a) idealizing it in some particular way is flawed or weak or internally incoherent, and b) an ideal-type stemming from this value-commitment is not useful for comprehending some particular case or set of cases. For Weber, there simply is no super-experiential standard for either the conduct or the scientific evaluation of scholarship, and as such it is simply inappropriate to scientifically reject a study on the grounds that one does not agree with its operative values—so long as those values have been properly transmuted into ideal-types, so that the study does not simply devolve into an ideological restatement of its premises in the guise of conclusions. Of course, one might *ethically* or *politically* disagree with such a study, but that would take us outside of the proper sphere of science.

[7] Is this a fifth use of popular culture for IR scholarship, supplementing the four famously delineated in Neumann and Nexon 2006? I'm not sure. It might be a fifth category, or it might be a twist on their "mirror" category that would move away from popular culture as illustration and towards popular culture as expression—in particular, as an expression that could be treated in parallel to the expressions at the heart of other ideal-typical models. The suggestion here is not that the popular cultural product itself constitutes cultural values, but that the process of formalizing value-commitments and systematically applying them to specific situations might contribute to the (re)construction of those values.

demands[8] an particular kind of internal consistency, because the craft of building a compelling science-fictional world involves providing a non-mystical explanation for everything and relating it to an overarching system of rational deductions from initial premises—precisely what a Weberian social scientist does when crafting and applying ideal-typical conceptual instruments to empirical cases. Done well, the result in either genre is an admittedly made-up world that still feels plausible, in part because it demands no leaps of faith to accept.[9]

Further, in both realms, the author/scholar/creator starts off in a sphere of cultural values, and articulates some kind of formalized version of a set of those values. Although in the diagram above I have depicted this as a constitutively separate phase, and implied that formalized value-commitments have to appear as an explicit model (Clarke and Primo 2007), in practice things are rarely so simple: readers of solid analytical social science, like fans of *BSG*, have to extract the value-commitments informing the presentation by a careful and close reading of what happens in the narrative.[10] But the *function* of those formalized value-commitments is the same across

---

[8] Alternatively, one might say: the community of science fiction fans demand. Thanks to Ken MacLeod for emphasizing to me the paramount importance of fan reaction (and even *anticipated* fan reaction) when it comes to the demand for internal consistency.

[9] Note that I am implicitly distinguishing between a leap of faith (by which one commits to the truth of some transcendent claim or value) and the kind of suspension of disbelief required by any reader of fiction. In science fiction, as in social science, explanations are *technical*, and can be evaluated on that basis.

[10] Explicit modeling exercises are rare in IR, and when they occur they usually only present the substantive part of the model, and not its associated and informing value-commitments (what exactly *are* Waltz's value-commitments, anyway?). An exception might be Robert Cox's corpus of critical scholarship, but in order to make that reading stick one would really have to read Cox against his own reflexive and dialectical intentions. Explicit modeling and declarations of value-commitments are similarly uncommon within the imagined worlds of science fiction (as opposed to, say, in the prefaces and afterwards to particular novels, or on the director's DVD commentaries), with *Star Trek*'s explicit and repeated in-universe statements of its own liberal humanist values standing as a notable exception. (Robert Heinlein's novels also feature extended explicit philosophical disquisitions, as do Ayn Rand's—the important difference between the two authors being their relative degree of literary skill in integrating the expositions into the flow of the narrative.)

the two domains: the concrete result—a fictional narrative or a social-scientific, case-specific one—is generated by the operation of bringing value-commitments to bear on particular situations. It is more or less operationally irrelevant whether those situations take place in the imagined world or in ours; while this matters for the question of whether the result is social science or (science) fiction, it does not matter to the practical relationship of value-commitments and outcomes, or to the "practical-moral" (Bernstein 1983) value of the resulting narratives. And in both realms, the author has to bring the audience along step by step, by simultaneously not losing touch with their cultural sensibilities and not losing the internal consistency and coherence of the product. *BSG*, like a good piece of analytical social science, invites the audience into a kind of systematic dialogue about what value-commitments do in practice; as such, it helps to enable the kind of pragmatic deliberation that might allow us to construct a more humane social order.

Battlestar Galactica's value-commitments

So the central methodological lesson that IR scholars can learn from *BSG* is that one mode of good social science[11] can proceed by constructing case-specific narratives that show the concrete and practical implications of a particular set of value-commitments. As such, *BSG* is a better model for how to construct "analytic narratives" than the essays collected in the volume *Analytical Narratives* (Bates et. al. 1998)—Ron

---

[11] There are others, of course. See Jackson 2011 for my effort to delineate four different ones.

Moore and his associates have crafted eminently more watchable stories than those constructed by scholars mainly concerned to validate their models by re-producing stylized histories that unfold in terms of those models themselves. Such tautological reasoning[12] is precisely why Weber so strenuously objected to the idea that any empirical research could serve to justify value-commitments. But to their credit, the producers of *BSG* don't seek to validate their values through their empirical narrative; instead, the goal is to show what a particular set of value-commitments gets us when applied to a specific—and in the case of *BSG*, quite extreme—situation. All that this can do is to contribute to an ongoing conversation about value-commitments—it cannot, even in principle, put the matter to rest.[13]

But what *are* the value-commitments that inform *BSG*? What theory of social and political life does the show advance? It should, but probably does not, go without saying that in order to answer this question we have to look at the overall plot of the series, and not focus overmuch on the statements made by individual characters (Brooks 1992). *What happens* is the important thing; the various construals that characters put on what happens might or might not be insightful, but the value-commitments of *BSG* are not necessarily to be found in the mystical visions of Laura

---

[12] The tautological character of such claims is why, contra Büthe (2002), a theoretically-informed historical narrative cannot possibly serve as any kind of a test of the theory that was used to construct the narrative in the first place.

[13] The fact that the creators of *BSG* aren't bound by the same kind of "reality constraint" as social scientists—in that the authors of fiction are free to invent events as required, while social scientists are bound to a set of intersubjective rules and procedures that limit their flexibility in generating empirical data—is a separate factor contributing to the inability of the creator of a fictional world to validate their value-commitments by illustrating them in narrative form, because a skeptic can always respond "well, that's only a story." But I am suggesting that there is also a logical problem with using an empirical narrative as a source of evidence in favor of certain values, a problem which obtains irrespective of whether the narrative takes place in our world or in another one.

Roslin or the gruff declarations of Bill Adama. *BSG* often uses the literary device of having a pair of characters spell out opposing sides of a complicated issue, but no single character functions consistently as the voice of the series' overall narrator. Rather, the series' values have to be extracted from the complete sequence of events, including the outcome of various initiatives and strategies.[14]

When we do this, I suggest, one thing that quickly becomes clear is that *BSG* is not founded on the values of liberal individualism so firmly ensconced in contemporary US and Western European societies. Such liberal individualism—founded on the values of individual liberty, equality, and reason (Jackson and Heilman 2008)—finds perhaps its clearest expression in rational choice theory, which posits constitutively autonomous individuals who are all endowed with identical abilities to make choices based on their particular utility-functions; social order, in such a worldview, is nothing other than a set of instrumental arrangements designed to enhance the individual benefits of the people participating in it. The prophet of such an idealization was John Locke, for whom the only justification for leaving the state of nature and entering into society was to remedy the "inconveniences" of the state of nature; after all, Locke's state of nature already possesses property, currency, and natural law, so there is nowhere near as pressing a need to band together as there is in other political theories. Indeed, in this kind of liberal theory—and this applies whether we are talking about theories about individuals,

---

[14] There's a parallel here to reading Thucydides' *History of the Peloponnesian War* as a tragedy, instead of doing the usual IR thing and extracting bits of Athenian dialogue and re-presenting it as though it represented Thucydides' enduring wisdom (Garst 1989; Lebow 2001).

theories about states, or whatever—social arrangements are always *chosen*, and chosen for essentially private reasons.

BSG does not exemplify a liberal world. Indeed, choices that follow a rational-choice logic of maximizing individual returns are consistently shown to lead to problematic consequences, as when the series shows us examples of black marketeering interfering with essential supplies, or when illustrating the machinations of the unscrupulous Tom Zarek. The initial set-up of *BSG* involves the curtailing of individually rational actions for the sake of a more compelling purpose: the survival of the human species, after its near-extermination in a devastating Cylon attack. Unlike liberal fictional worlds such as that depicted in *Star Trek*, the Colonial characters in *BSG* always labor under the burden of necessity, and the military hierarchy is again and again explicitly validated as essential to maintaining what remains of human society. And reason does not serve as the dominant standard for decision-making, especially for the most important decisions: Laura Roslin successfully leads the fleet along what can only be described as a vision quest, and prophecy and scripture stand in as equal, and sometimes more significant, inputs to deliberative processes.[15]

But if *BSG* does not depict not a liberal universe—as evidenced by the fact that these non-liberal strategies and commitments *succeed* on a regular basis—what does it depict? I would suggest that the series is founded on a commitment to a *critical*

---

[15] Note that the question of how to square these seemingly supernatural messages with the genre demand for disenchanted explanations is one of the ongoing challenges in *BSG*; other authors in this volume discuss that aspect of the show in more detail, but for the moment, suffice to say that the *BSG* writers are generally quite carefully to give both a supernatural and a naturalist account of events—at least, until the end of the last episode of the series.

humanism, a set of values decidedly alien to liberal individualism and its valorization of rational decision-making logics. By "humanism" I clearly do not mean anything as provincial as *homo sapiens*-centrism; it would be awkward indeed to pose such speciesism as the central value-commitment for a series whose conclusion involves the interbreeding of Colonials and Cylons, and even more awkward for a series whose "humans" are not *homo sapiens* (although they are part of our ancestry). Rather, I mean something that can perhaps best be expressed in the terminology provided by another science fiction author—Orson Scott Card—in his delineation of four different modalities of foreignness or otherness:

> The first is the otherlander, or *utlänning*, the stranger that we recognize as being a human of our world, but of another city or country. The second is the framling…This is the stranger that we recognize as human, but of another world. The third is the raman, the stranger that we recognize as human, but of another species. The fourth is the true alien, the varelse, which includes all the animals, for with them no conversation is possible. They live, but we cannot guess what purposes or causes make them act. They might be intelligent, they might be self-aware, but we cannot know it (Card 1991, 34).

Adopting Card's terminology, the kind of "human" that I have in mind here is the raman: a kind of human-ness that transcends species, perhaps as a more general kind of sentience or person-hood but still possessing, in some sense, *moral* character and value.[16] Hence, "humanism" means a commitment to appreciate and learn from the widest possible range of human experiences, and not to place arbitrary inviolable boundaries between humans that would prevent the sharing of such experiences and

---

[16] The references in the closing minutes of the *Battlestar Galactica* spin-off series *Caprica* to the Cylons as "otherwise sentient" perfectly captures the spirit of this expanded notion of humanity.

the wisdom derived from them.[17] A certain skeptical tolerance of difference follows (as in Toulmin 1992): since no human grouping has a monopoly on valid knowledge, the only viable option is the kind of pluralism that allows multiple traditions to flower.

The "critical" part of critical humanism presses this skeptical spirit even further, and challenges any and every specific definition of the boundaries of the human. There are doctrinaire humanisms that declaratively establish the essential characteristics of human beings and then more or less imperialistically dictate a set of social arrangements that *must* be erected in order to preserve or protect those characteristics. It is irrelevant, for the purposes of this classification, whether those essential characteristics are rooted in theology, in an ethical system, in a physiology or psychology or anthropology, or in a body of legal or policy precedents. Doctrinaire humanisms of whatever flavor proceed on the assumption that humanity as a whole can be *defined*, and goals can be set for humanity that will trump the goals set by individual human beings; in the discourse of any doctrinaire humanism, 'humanity' functions as the ultimate polity within which all other human actors are nested, and as such, it can neither be contravened nor negotiated with. Much as Locke is the prophet of liberal individualism, Kant is the prophet of doctrinaire humanism, seeking to specify the characteristics of the rational (human)[18] being *as such* in advance—indeed, *a priori*—in a way that brooks no compromise.

---

[17] "Humanism" in this sense is very much like what Jens Bartelson (2010) calls the "world community" tradition in political thought. "World" in this tradition means *cosmos* rather than "planet," so these might even be the same tradition.
[18] Slight ambiguity in Kant's terminology here, since Kant clearly prefers to talk about rational beings rather that human beings. But if we apply Card's terminology, Kant is talking about all ramen, which means "all humans" by the expanded

Critical humanism, by contrast, challenges or suspends all such definitions as unnecessarily limiting. Instead, 'human' is something of a weakly-shared commonplace, a deliberately vague notion that needs to be filled in through its practical application in concrete circumstances. What Huntington claims about civilizations—that they are "the biggest 'we' within which we feel culturally at home as distinguished from all the other 'thems' out there" (1996: 42)—is even more true of the 'human' in critical humanism, both because the notion is even bigger (conceptually speaking) than any more provincial human community, *and* because it depends not on a self/other differentiation but on something like sentience/non-sentience. States, civilizations, and other categorically bounded communities rely on a particular form of otherness[19] that Carl Schmitt refers to as the friend/enemy distinction, a distinction that both defines existential threats to the community and enables a kind of nobility of struggle between the community and others who threaten it: one fights to repel the enemy, not to exterminate them. Humanity, Schmitt observes, cannot serve as the basis of a genuine friend/enemy distinction of this sort, precisely because "the enemy does not cease to be a human being" (2007: 54); as such, the notion of fighting wars in the name of humanity strikes Schmitt as an absurdity, because to characterize one's human opponents as

definition of "human" in terms of which there can be humans of other species. (Recall that that in Card's typology, the distinguishing characteristic of the varelse is that no communication with them is possible; this strongly suggests that ramen "humans of another species" would have to be Kantian "rational beings," since we would have no way of communicating with them if they were not.)

[19] "It *is* just as if we had to conform ourselves to an objective reality existing independently of any of the individuals involved: but we have to conform ourselves to it, not because of its material shape, but because we all require each other *morally* to conform to the 'situations' emerging into existence between us. They exist as third entities, between us and the others around us. Thus, to us as individuals, such situations may seem like one or another kind of 'external' world, as something lying at the other end of the person-world dimension of interaction…However, such situations are not external to 'us' as a social group. As neither 'mine' or 'yours,' they constitute an Otherness that is 'ours,' our own peculiar form of Otherness" (Shotter 1993b: 8-9).

inhuman both justifies unfettered brutality and violates the expansiveness and inclusivity of the very concept of humanity in the first place.

But where Schmitt traces this potential for the abuse of the appeal to humanity to a set of liberal notions about the goodness of man and a concomitant refusal to embrace the political reality of existential threats to one's own community, the critical humanist would trace the problem to the promulgation of doctrinaire specifications of what 'humanity' means. Doctrinaire humanism of a liberal flavor might indeed be particularly prone to such abuses of the appeal to 'humanity' because of its self-valorization as uniquely and universally warranted by reason, a self-valorization that renders it especially incapable of regarding non-liberal Others as anything but deficient versions of itself. But the critical humanist would see this potential for abuse in *any* advance or *a priori* account of humanity or the human, since such a categorical notion would be able to function as the ultimate warrant for ruling out some course of action or some subset of the potentially human population. Critical humanism does not provide a definition of "the human," but it does demand that the question "what are the boundaries of humanity?" be raised whenever important deliberations take place. The specific answer is not especially important, since any such answer is provisional at best; what matters instead is that the question be asked, and asked repeatedly, because the

asking and provisional answering of the question allows the interlocutors to construct an ever more expansive conception.[20]

Obviously, it is not the case that most of the characters in *BSG* would accept this kind of critical humanism, at least not initially; almost all of the Colonials spend most of the series fighting the Cylons, and most of the Cylons spend the series trying to exterminate the Colonials. Indeed, if one were to ask about the boundaries of humanity in the early days of the series—if one were to conduct a focus-group for the crew of the battlestar *Galactica*, say—one would be almost certain to find "the Cylons" referenced as the crystal-clear Other against which the Colonials (who think of themselves as "humans") was opposed. But as the series unfolds, Cylon/Colonial border transgressions multiply, yielding one viable hybrid child and at least two relatively stable Cylon/Colonial relationships. Again and again the series challenges the Colonials' definition of humanity, perhaps most obviously in revealing several trusted members of the Colonial crew to be Cylons, but also whenever the Colonials are faced with the possibility of violating their own sense of decency to achieve a victory against the Cylons—whether the issue involves torturing a prisoner, engaging in suicide bombings, or violating the established rules of warfare.

One of the most striking examples of the latter comes in the third season episode "A Measure of Salvation," when the Colonials discover a way to infect the entire Cylon race with a deadly virus; although Laura Roslin as President orders the military

---

[20] Whether this ever-more-expansive definition of humanity ends up fusing humanity with the divine, or even allowing for such an evolution to take place, is the sort of theological question that the background Mormon conceptual material invoked in the series has an answer to. On this point, see Iver Neumann's chapter in this volume.

establishment to use this biological weapon to destroy the entire Cylon race, Helo—the Colonial officer who has fathered a child with one of the humanoid Cylon models—argues that doing this would be a "crime against humanity." The way in which he makes this pronouncement is important; he doesn't just declare it, but instead, he stops himself, considers, and then stakes his claim—and immediately qualifies it by saying that committing genocide against the Cylons would mean that the Colonials would lose a piece of their souls. After he is overruled, Helo secretly uses his trusted officer position to sabotage the Colonial plan, suspending his sworn duty to obey orders for the sake of a broader notion of "humanity"; after the incident is concluded, Adama quotes Helo's ultimate justification for opposing the plan—"we would lose a piece of our souls"—back to Roslin when he decides not to investigate the mater further. Throughout the episode, we see a notion of Colonial *identity*—who they are, what kind of human beings they are—trumping what might have otherwise been considered the rational course of action. As Roslin points out, the Colonials are fleeing from a Cylon race that seems intent on destroying them, and if there were ever a justification for an extreme measure like genocide, one might think that this would be it—particularly, she claims, since it is unclear whether future generations of Colonials will regard the destruction of the Cylons as "genocide." But in the end, it is the opposing position that is vindicated by events: not only do the Colonials survive the episode without committing genocide, but later plot developments illustrate that the Cylons are essential to the Colonials' own continued survival (including the Cylons' repair of the *Galactica* itself).

This moment of crisis also exemplifies a further dimension of critical humanism: the inability of any formal ethical system to serve as the final expression of humanity. Instead, we repeatedly see characters suspending ethical codes for a higher purpose, but that purpose is not (*contra* what Schmitt might expect) the physical survival of the political community. Instead, it is the cultural or moral survival of the humanity of the Colonials that is at stake in such moments, and this imperative overshadows everything else: rational strategic calculation, personal ambition, or the desire for a vengeful justice. The latter is most prominently on display during Gaius Baltar's trial at the end of the third season, during which a verdict of "not guilty" is returned not because Baltar was innocent, but because of an awareness of *everyone's* guilt and the need to move past blame to focus on the future. The law and the notion of compensatory justice are suspended not because they conflict with some higher law, but because the Colonials partially enter a realm where simply putting Baltar to death would mean sacrificing Baltar to salve their own consciences. This might be the rational response, and it might be the expedient response, but it is not the critical human response, because the critical human response cannot involve cutting a human being out of humanity—that would mean erecting a boundary around the human, rather than pursuing the most expansive definition possible.[21] And here again, the narrative proves this decision prescient, since in the end Gaius is important to the eventual fate of both the Colonials and the Cylons with his role in the rescue of the Cylon/Colonial hybrid child Hera.

---

[21] There is a parallel here to Nietzsche's discussion of the pale criminal in *Thus Spoke Zarathustra*, but space prevents me from working out the parallel in any detail.

"Not cutting anyone out of humanity" is similarly on display in Helo's second major moment serving as the conscience of the Colonial fleet. In the episode "The Woman King," Helo starts to suspect that a Colonial physician (Dr. Roberts) is deliberately killing refugees from Sagittaron in order to preserve medical resources for other Colonials; Helo persists in this belief despite the admonitions of his superior officers and the mocking derision of his friends (Sagittarons are not well-liked), trusting instead the evidence he finds that does not quite sit right with his expansive sense of humanity and the responsibility that serving humanity implies. Helo is proven right in the end, and the moral indictment of Dr. Roberts does not simply critique his racist attitude towards Sagittarons—instead, the whole idea of anyone taking it upon themselves to make "the tough decisions" about who lives and who dies is shown to be morally bankrupt and riddled with the potential for abuse. A clear definition of humanity might allow such decisions to be made with a clear conscience, but that's precisely the point: because there is no clear definition, any decision like this is groundless and arbitrary. But unlike in Schmitt, the moment of sovereign decision is not celebrated in *BSG*, but is instead reviled in favor of a concerted effort to keep the boundaries of the human as wide as possible.

There is also a Cylon side to this critical humanism. All we know of the Cylons and their master plan (besides a weekly invocation of the Cylons' plan in the opening of virtually every episode) in the early part of the series as originally aired is that the extermination of the human race seems to be an important part of it. The "skinjob" humanoid Cylons are fascinated by human beings, but seemingly in the manner of the

way a lab technician might be fascinated by a rat: tests and obstacles are set up to see how the humans will react, and this set of actions is easily reconcilable with an overall program of extermination inasmuch as greater knowledge can show the Cylons exploitable weaknesses in their human quarry. The exceptions to this rule are the three humanoid Cylons who profess to have fallen in love with individual humans (Caprica 6 and Boomer, and the Sharon model who returns to the Colonial fleet with Helo and eventually becomes known as Athena). Sharon-Athena could be written off as an aberration, even though she becomes pregnant with a Cylon-human hybrid apparently foretold in a series of visions shared by other characters both human and Cylon; this might easily have been a story about *one* Cylon becoming in some sense human, while the others remained implacably opposed. Such a storyline would have maintained the Colonial/Cylon boundary more or less intact, and established a definitive barrier to the boundaries of the human.[22]

But the dramatic turning-point experienced by the other two humanoid Cylons professing love for individual humans[23] in the episode "Downloaded" drives the series in a quite different direction. Caprica 6 and Boomer realize that they are no longer in complete agreement with the original plan to wipe out the human race. Caprica 6's "inner Gaius"—an apparition that, like Baltar's "inner 6," might simply be a

---

[22] Similarly, the fact that "Simon"—the number 4 model seen living among the Colonial fleet in the TV movie *The Plan*—is married to and seemingly in love with a human woman might also have been an aberration, since all of the other Cylons traveling with the fleet end up carrying out the program of frustrating the humans' efforts to carry on. And the brutal murder of a child by a number 1 model underscores the inhuman cruelty of at least some of the Cylons.

[23] The dramatic irony of the fact that Boomer's love interest Chief Tyrol is actually a Cylon, albeit one of the Final Five rather than a production-line humanoid Cylon, complicates things later on, but does not affect the overall point since Boomer *thinks* Tyrol is a human at this point.

manifestation of subconscious guilt prompting a reconsideration of the categorical boundary between Colonial and Cylon[24]—taunts her with accusations of "inhumanity," and the revelation that Gaius himself is still alive on the Colonial fleet prompts her to openly question the official narrative of the Cylon campaign. Boomer and Caprica 6 thus initiate the discussion among the Cylons that leads, in time, to a civil war between Cylon factions trying to articulate their relationship to humanity. And in the course of that civil war, we see that the true initiator of the extermination campaign— John/Cavil/One—is inspired by nothing more than a loathing of his own human-like qualities (particularly his physical and cognitive qualities) and a desire to categorically bound the human so that he can separate himself from it.

All of this comes to a head in the fourth season episode "No Exit," in which we simultaneously learn the identity of the fifth hidden Cylon (Ellen Tigh), the back-story of the Final Five's flight from the original Earth and their role in stopping the first Cylon war by creating the humanoid Cylon models, *and* the fact that the *Galactica* itself is falling apart. The interweaving of these revelations make it abundantly clear that the show sides with efforts to expand the boundaries of the human and not with efforts to keep the boundaries clear; this is perhaps clearest in the contrast between Cavil's extended harangue about being trapped in a humanoid body with humanoid senses, and Adama's growing realization that the *Galactica* simply cannot be repaired without Cylon assistance. While Cavil's maniacal quest to purify himself and the Cylon race of

---

[24] At least until the very last episode of the series, when the inner manifestations are revealed to be something like angels. But by this point—roughly the last 45 minutes of the series finale—the show has veered off into pure myth and consequently abandoned the commitment to disenchanted explanations that characterized the series up until then.

any traces of humanity leads to increasingly cruel and brutal behavior, Adama's grudging acceptance of the fact that his ship and crew can only be saved by widening the boundaries of humanity so as to incorporate Cylons as equal members of the community leads to a measure of salvation. Adama does not reach this point through love, as Caprica 6 and Helo did, but instead comes to widen his definition of humanity through hard practical experience. The process is different, but the result is the same: time and again the show leads the viewer to appreciate the value of refusing to define the limits of the human in advance, but instead allowing the widest possible range of experiences of human diversity.

The Great Debate

But what are we to *make* of *Battlestar Galactica's* depiction of a critical humanism in action, even and perhaps especially during an extreme situation where survival is on the line almost daily? The kind of systematic development of the implications of such a value-commitment that one finds in a science-fictional literary production is of course different in some ways from what one might expect in a piece of professional social-scientific research; the style of argumentation in professional social science is more explicit and formal, and there is an assumption of either a greater degree of homogeneity between author and audience on what constitutes a superior explanation or a need to lay out the criteria for a good explanation in a more or less deliberate fashion. As such, professional social-scientific research has at least the possibility of

coming to consensus on matters of fact—at least a *technical* consensus on the way that applying a value-commitment to an empirical situation yields findings, and at most a *substantive* consensus that can persist as long as the value-commitments underlying the research exercise are widely accepted in the scholarly community.

Matters are somewhat different in the realm of science fiction, where the idea that public criticism of an argument can lead to an overall improvement in knowledge-claims is somewhat less of a guiding constitutive principle for the activity of composition. Certainly authors do write stories that critique other stories, but rival treatments of the same scenario (the terraforming of Mars, for example, or the construction of self-aware AI) can and do co-exist in the same literary space without readers having to unequivocally choose one over the other—or synthesizing them to create a single superior account. *BSG* inhabits a literary terrain populated by many other accounts of human/non-human encounter, many of which take sharply different perspectives on the boundary of the human community. But how might we say that, for example, the eliminationist approach to life-forms other than *homo sapiens* that we find in films like *Starship Troopers* or *Aliens*—or serial science-fiction television shows like the reimagined series *V*, which whole-heartedly embraces the attitude that *BSG* calls into question—is disproven or refuted by *Battlestar Galactica*? How, even, might we argue that *BSG*'s critical humanism is superior to those alternatives? Clearly we can't do it by referring to empirical evidence the way we would have to if this were a social-scientific dispute; rather, the entire discussion would have to take place in the subjunctive voice, and spend a lot of time pondering "what if" questions.

Which is precisely the point. Science fiction, as Ursula LeGuin (1976: Introduction) once noted, is descriptive rather than predictive; it tells us not about a world beyond itself, but about the world that we (authors and readers alike) inhabit. By working through various scenarios, works of science fiction contribute to an ongoing cultural dialogue about some of the most profound issues confronting human society today—in *BSG*'s case, the ever-present issue of who gets to claim membership in the human community, and what approach we ought to take to encounters with actions and beings that at first appear to fall beyond the pale. Working through the extended case study that *BSG* presents us with, in dialogue with other treatments of similar cases, may help sensitize us to such concerns the next time we concretely encounter strangers of various sorts. The payoff, as it were, is not in a concrete prescriptive lesson that one walks away from the story with, as much as it is in a certain informed cultural disposition which one carries into the next in-some-ways-similar-but-never-exactly-the-same situation.

And this may be the final lesson that social science can learn from science fiction, because at its best *it too* provides extended case studies that, when worked through, can contribute to the formation of a particular constellation of cultural tools and techniques on which actors draw in making sense of their situations (Joas 1997: 157-162). As John Dewey noted, in refining their conceptual equipment and using it to analyze empirical situations both contemporary and historical, scholars systematize the diffuse, tacit, value-laden intellectual heritage of their present situation, forging conceptual instruments (like "the social network" or "the process of legitimation") that express a

particular, contemporary sensibility about the world—in the case of social scientists, about the *social* world of global politics. In so doing, the analyst both preserves something of value from that sensibility and liberates it to make it "better fitted…to deal with any one of the indefinite variety of things that may later present themselves" (Dewey 1920: 150; see also Shotter 1993a: 203-207). Like science fiction, social science can provides materials that might inform case-specific deliberations about appropriate courses of action, and in that way make its proper contribution to social and political life more broadly. Perhaps if social scientists kept their eye on that contribution, and gave up the futile dream of successfully subsuming the infinite variety of concrete situations under simple covering-laws that would allow both prediction and control, they might one day produce contributions as compelling as *Battlestar Galactica*'s saga of an ever-expanding human community.

References

Bartelson, Jens. 2010. *Visions of World Community*, Cambridge: Cambridge University Press.

Bates, Robert H., Avner Greif, Margaret Levi, Jean-Laurent Rosenthal, and Barry R Weingast. 1998. *Analytic Narratives*. Princeton: Princeton University Press.

Bernstein, Richard J. 1983. *Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis*. Philadelphia: University of Pennsylvania Press.

Brooks, Peter. 1992. *Reading for the Plot: Design and Intention in Narrative*. Harvard University Press.

Büthe, Tim. 2002. "Taking Temporality Seriously: Modeling History and the Use of Narratives as Evidence." American Political Science Review 96 (3): 481-493.

Card, Orson Scott. 1991. *Speaker for the Dead*. New York: Tor.

Clarke, Kevin A., and David M. Primo. 2007. "Modernizing Political Science: A Model-Based Approach." Perspectives on Politics 5 (4): 741-753.

Easton, David. 1953. *The Political System: An Inquiry Into the State of Political Science*. New York: Alfred A. Knopf.

Dewey, John. 1920. *Reconstruction In Philosophy*. New York: Kessinger Publishing, LLC.

Drysdale, John. 2007. "Weber on Objectivity: Advocate or Critic?" In *Max Weber's "Objectivity" Reconsidered*, ed. Laurence H. McFalls, 31-57. Toronto: University of Toronto Press.

Duvall, Raymond. 2001. "Ideal Types." In *Routledge Encyclopedia of International Political Economy*, ed. R. J. Barry Jones, 703-704. London: Routledge.

Garst, Daniel. 1989. "Thucydides and Neorealism." *International Studies Quarterly*, 33 (1): 3-27.

Gunnell, John G. 2007. "The Paradox of Social Science: Weber, Winch, and Wittgenstein." In *Max Weber's "Objectivity" Reconsidered*, ed. Laurence H. McFalls, 58-88. Toronto: University of Toronto Press.

Huntington, Samuel P. 1996. *The Clash of Civilizations and the Remaking of World Order*. New York: Simon & Schuster.

Jackson, Patrick Thaddeus. 2011. *The Conduct of Inquiry in International Relations*. London: Routledge.

Jackson, Patrick Thaddeus, and James Heilman. 2008. "Outside Context Problems: Liberalism and the Other in the Work of Iain M. Banks." In *New Boundaries in Political Science Fiction*, ed. Donald M. Hassler and Clyde Wilcox, 235-258. Columbia, SC: University of South Carolina Press.

Jackson, Patrick Thaddeus, and Daniel H. Nexon. 2003. "Representation is Futile? American Anti-Collectivism and the Borg." In *To Seek Out New Worlds: Science Fiction and World Politics*, ed. Jutta Weldes, 143-167. New York: Palgrave MacMillan.

Joas, Hans. 1997. *The Creativity of Action*. Chicago: University Of Chicago Press.

Lebow, Richard Ned. 2001. "Thucydides the Constructivist." *American Political Science Review* 95: 547-560.

LeGuin, Ursula. 1976. *The Left Hand of Darkness*. New York: Ace.

Neumann, Iver B., and Daniel H. Nexon. 2006. *Harry Potter and International Relations*. Lanham, MD: Rowman & Littlefield Publishers, Inc., May 25.

Ringer, Fritz K. 2004. *Max Weber: An Intellectual Biography*. Chicago: University of Chicago Press.

Schmitt, Carl. 2007. *The Concept of the Political*. Chicago: University Of Chicago Press.

Shotter, John. 1993a. *Cultural Politics of Everyday Life*. Toronto: University of Toronto Press.

Shotter, John. 1993b. *Conversational Realities: Constructing Life Through Language*. Thousand Oaks, CA: Sage.

Toulmin, Stephen. 1992. *Cosmopolis: The Hidden Agenda of Modernity*. Chicago: University of Chicago Press.

Van Belle, Douglas. 2006. "Dinosaurs and the Democratic Peace: Paleontological Lessons for Avoiding the Extinction of Theory in Political Science." *International Studies Perspectives* 7: 287-306.

Weber, Max. 1999. Die "Objektivität" Sozialwissenschaftlicher und Sozialpolitischer Erkenntnis. In *Gesammelte Aufsätze zur Wissenschaftslehre*, ed. Elizabeth Flitner, 146-214. Potsdam: Internet-Ausgabe, http://www.uni-potsdam.de/u/paed/Flitner/Flitner/Weber/.

Weber, Max. 2004. *The Vocation Lectures*. Indianapolis: Hackett Press.