Morgan's Canon Revisited*

Sean Allen-Hermanson†‡

The famous ethological maxim known as 'Morgan's Canon' continues to be an object of interpretive controversy. I reconsider Morgan's Canon in light of two questions: First, what did Morgan intend? Second, is this, or perhaps some reinterpretation of the Canon, useful within cognitive ethology? As for the first issue, Morgan's distinction between higher and lower faculties is suggestive of an early supervenience concept. As for the second issue, the Canon, both in its original form and in various recent rereadings, offers nothing useful to cognitive ethology.

In *An Introduction to Comparative Psychology*, C. Lloyd Morgan offers the following 'basal' principle, or canon, as a way to avoid anthropomorphic error when explaining the behavior of animals:

> In no case may we interpret an action as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale. (Morgan 1894, 53)

Morgan's Canon is "perhaps the most quoted statement in all of psychology," writes primatologist Frans de Waal (2001, 67); and quite "possibly the most important single sentence in the history of the study of animal behavior," according to another commentator (Galef 1996, 9, as reported in Radick 2000, 3). Though Morgan's Canon (hereafter, 'the Canon') has enjoyed a wide influence among ethologists, it has also been the focus of interpretive controversy, at least since the rise of behavioristic psychology (Newbury 1954).

The Canon is somewhat less familiar to philosophers. The general intent is to help solve the problem of other minds with respect to nonhumans.

*Received March 2005; revised October 2005.

†To contact the author, please write to: Sean Allen-Hermanson, Department of Philosophy, Florida International University, 11200 SW 8th Street, Miami, Florida 33199; e-mail: hermanso@fiu.edu.

‡I would like to thank Elliot Sober, Collin Allen, and the anonymous referees at *Philosophy of Science* for helpful comments.

Philosophy of Science, 72 (October 2005) pp. 608–631. 0031-8248/2005/7204-0005$10.00
Copyright 2005 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

The Canon is supposed to guide our mental state ascriptions by directing us towards the hypothesis positing the 'lowest' mental processes that suffice to account for a given behavior. The central interpretive issue concerns the distinction between higher and lower mental faculties. What is meant by this, and the 'psychological scale' Morgan speaks of? Morgan's actual position has become more, rather than less, obscured by recent scholarship, and as a result, the Canon has come to be used rather like a mirror, reflecting and reinforcing the preconceptions brought before it. On the one hand, it has been championed by behaviorists who treat it as a cognate for methodological parsimony,[1] and, concomitantly, a "refutation of the Darwin-Romanes view of the continuity of mentation . . . and consequently, as a denial of consciousness to animals" (Rollin 1989, 75). This sort of depiction is easily found, as in this introduction to comparative psychology:

> Consider the case of a butterfly's orientation toward a source of light. The parsimony principle [i.e., Morgan's Canon] would favour an explanation in terms of the action of a preprogrammed mechanism (called "tropism") that attracts the insect towards lighted spots, rather than an interpretation invoking, for example, the butterfly's curiosity. (Vauclair 1996, 3)

However, many cognitive ethologists and philosophers also defend the Canon, but on the quite different grounds that it is *not* inimical to animal mentality, and in fact serves as "a welcome prophylactic against sentimental and unwarranted attributions of finer sensibilities to our favorite animals" (Wilder 1996, 35).

Over the years, commentators—including Newbury (1954), Radner and Radner (1989), Thomas (1998), and Kimler (2000)—have repeatedly pointed out that psychologists are in error when they interpret the Canon as a call for behaviorism. Skinner (1938), Boring (1950), Vauclair (1996), and many others[2] have incorrectly interpreted the Canon as a call for explanatory simplicity, parsimony, and/or Ockham's razor. Other recent authors have argued that considerations of parsimony actually lead *towards*, not away from, mentalism.[3] Morgan would have agreed that the simplest hypothesis is that other organisms are just like us. But it is precisely for this reason that his Canon is not intended as a form of

1. I.e., find the simplest theory that accounts for all the evidence (Sober 1988).

2. Thomas (2001) is a systematic compilation of the various and pervasive misrepresentations of the canon within psychology.

3. Regan (1983), Rollin (1989), Bekoff and Allen (1997), and Sober (2001) have independently argued that considerations of parsimony lead towards mentalism, not behaviorism.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

parsimony: The simplest hypothesis is just naïve anthropomorphism. Special attention is given in the second edition of *An Introduction to Comparative Psychology* (1903) to emphasize that his Canon in no way "excludes the interpretation of a particular activity in terms of higher processes, if we already have independent evidence of the occurrence of these higher processes in the animal under observation" (Morgan 1903, 59). Morgan also stresses that "the simplicity of an explanation is no necessary criterion of its truth" (Morgan 1894, 54). Morgan says this in order to *reproach* those who would impute complex psychologies to animals on the grounds that it is the 'simplest'. Though "[i]t is the apparent simplicity of the explanation that leads many people naively to adopt it" (Morgan 1894, 54), this is a mistaken impression just as it would be a mistake to endorse creationism just because it offered a simpler explanation than evolution by natural selection, he argues. While the tendency towards blithe antimentalistic readings have been answered by recent scholarship, misconceptions and errors persist in other reappraisals of the Canon. Most importantly, there is nothing near a consensus about what the call for a preference of explanations that posit 'lower' faculties really amounts to.

In this paper I will revisit the question as to who should be considered the proper heirs of Morgan's Canon, and whether it is worth keeping. My discussion will be structured around the following two questions that need to be considered separately: (1) What did Morgan really intend by his Canon? (2) Is this, or perhaps some other reconstruction, a plausible guide for our ascriptions of mental states to animals? Here is a rough sketch of the answers I will defend: As for the first issue, though Morgan was certainly not a skeptic about animal minds, comprehension of the Canon has been hindered because of confusion about what he meant by the terms 'higher' and 'lower'. The situation is further complicated because of Morgan's occasional restatements of the Canon. However, a close reading of his writings strongly suggests continuity of thought. The higher-lower distinction must be understood in light of Morgan's tentative and inchoate idea that higher phenomena, such as mental processes, must 'emerge from', that is, 'supervene on', lower 'basal' conditions.

Morgan's doctrine of emergence, for which he sometimes uses 'supervenes' as a cognate, covers both ontogenetic and phylogenetic forms of development. Mind emerges in the individual from lower neurophysiological conditions, and likewise, mind as a general phenomenon gradually appears on the evolutionary stage. For Morgan, the emergence, or supervenience, of mind also encompasses a hierarchical ordering of three mental faculties, to wit: reason (conceptual thought), which is higher than intelligence (stimulus-response conditioning), which is higher than mere instinct.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

The role of emergence in contextualizing the Canon is perhaps hinted at in the "scale of psychological evolution and development" spoken of in a 1906 (59) revision. A later revision, offered in 1932, offers this:

> The evolutionary canon that we should not interpret an earlier and lower stage of mental development in terms applicable only to the interpretation of the higher and later stage (Morgan 1932, 262).

If this, too, seems merely suggestive, the version he offers in 1925 leaves no doubt about the important connection between the Canon and the emergence doctrine:

> In no instance should we interpret events in terms of concepts appropriate to a higher level of emergence if they can adequately be interpreted in terms of concepts appropriate to a lower level of emergence (Morgan 1925, 61).

I contend that Morgan's Canon will continue to be misunderstood if we fail to acknowledge that it presupposes a form of mind-body emergence or 'supervenience'. I will also show that Morgan's supervenience doctrine also appears much earlier than has been previously assumed—in short, his commitment to the supervenient/emergent dependence of higher phenomena on lower phenomena actually precedes his introduction of the Canon, and is an essential part of its context. Oddly, these connections are all but invisible to contemporary discussions of the Canon.[4]

Turning now to the second issue mentioned above, various rereadings of the Canon, though historically inaccurate, still ought to be assessed in light of possible contributions to our continuing efforts to shield the study of animal mentality from unbridled romanticism. As Sober says, "It is worth trying to understand the Canon in its historical context. It is also worth seeing whether the Canon can be defended in ways that its originator never imagined" (Sober 1997, 228). On this score, I will draw the pessimistic conclusion that recent interpretive innovations do not pass muster. Sober's (1997) reworking, which reads higher and lower faculties in terms of behavioral entailments, is one notable and instructive example that I will address in detail. But what of the Canon itself? Does it help or hinder scientific investigation of animal minds? Read properly, that is,

4. Newbury (1954) is unusual in acknowledging that Morgan's doctrine of evolutionary emergence contributed to the context of the canon. More recent discussions, such as Allen and Bekoff (1997), Fodor (1999), Sober (1997), Radner and Radner (1989), and Rollin (1989) do not explore this connection. Sober mentions that "Philosophy of mind is, and for a long while has been, 99% metaphysics and 1% epistemology" (2001, 225). Morgan's Canon is, of course, an epistemological principle, and maybe this is why its connection to emergence is neglected.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

with a protosupervenience thesis in the background, the Canon seems more like a truism than a useful methodological principle. I will argue that it asserts the conditional claim that if something can already be explained reductively, then it cannot also be explained in nonreductive terms.

To begin, on the assumption that crude anthropomorphism is contrary to a scientific approach, we need a principled way to tell when our mentalistic attributions to animals are unwarranted. Morgan's Canon is supposed to help us here by "guiding inferences towards hypotheses that are 'conservative' in the psychological mechanisms they assign" (Sober 1997, 228). Ersatz psychological ascriptions are those that can be replaced by 'deflationary' explanations, as above where the imputation of curiosity to a butterfly is replaced by the idea of a stereotyped behavioral routine. However, the utility of the Canon is contested by others who consider it to be a behaviorist relic.

Costal, for example, writes that although the "Canon is usually represented as a brave step towards mechanistic behaviorism" (1993, 113), in fact this myth neglects the important fact that it "was not meant to prohibit explanations of animal behaviour framed in intentionalist language" (1993, 117). Rollin concurs, adding that the Canon was "directed to over-zealous Darwinians who are too quick to attribute reason . . . to animals . . . . Morgan is not responsible for the use to which his ideas were put by people like Watson and other behaviourists" (1989, 78). Rollin is right to point out that Morgan's formulation of the Canon is only concerned with the evaluation of rival *mental* hypotheses. Read literally, it does not urge the elimination of the mental at all. Moreover it is clear that Morgan did not mean to suggest that animals were unconscious automatons. This makes sense, Rollin observes, given the importance he placed on evolutionary continuity and his commitment to Spinozistic panpsychism. ("Universal correlation is also part of my creed assuredly beyond proof. And here my cry is: Back to Spinoza" [Morgan 1923, 52].) Rollin also criticizes misguided behaviorist interpretations of Morgan, insisting that the Canon "in fact presupposes [animal consciousness], in order even to make sense" (Rollin 1989, 75). Rollin goes so far as to claim that the Canon is intelligible *only* in light of Morgan's panpsychism. According to his interpretation, which I will call the 'Panpsychism Canon', the deflationary method can only be used to adjudicate between rival *psychological* explanations. It is for this reason, Rollins says, it offers nothing to the behaviorists.

We need to tread delicately here. Of course, if Morgan intends both higher and lower faculties to be understood as *mental* faculties, then strictly speaking it is a mistake to think otherwise. But we can also ask to what extent Morgan's metaphysical commitments can be distinguished

https://doi.org/10.1086/505187 Published online by Cambridge University Press

from the Canon itself. By my lights, the Canon still seems intelligible when plugged into a variety of metaphysical doctrines about the mind-body relationship. The Canon makes no explicit mention of either panpsychism, or animal consciousness, and Rollin does not show that *we* must adopt these assumptions, in order to coherently prefer lower faculties in our explanatory practices (i.e., even if lower faculties are mental). For that matter, I can see nothing incoherent about extending the Canon to include showing a preference for *nonmental* faculties over ones that are psychological. If I am right that Morgan thinks mental phenomena emerge from nonmental phenomena, it seems prima facie reasonable to extend the distinction between higher and lower so as to accommodate this.

Let me now return to the basic issue propelling the interpretive controversy: What do the predications 'lower' and 'higher' mean for Morgan? As Sober (1997) points out, disagreement and confusion about the Canon will persist until we elucidate this distinction.

Next I will consider three alternatives to the Panpsychism Canon which offer differing interpretations of the central distinction. Two of these are already extant in the contemporary literature, while the third is novel. I shall call these the 'Metarepresentational Canon', 'Sober's Canon', and finally the 'Supervenience Canon'. The rest of my argument will proceed as follows. The first and second are not plausible elucidations of Morgan's own view. As regards issue two, the former interpretation differs little from a call for parsimony, and so seems redundant. On the other hand, the latter fails to capture plausible intuitions about the relationship between mind and behavior. Finally, the third view is a conjecture about what Morgan probably had in mind, though his Canon's usefulness to contemporary ethology is also doubtful.

I turn now to the Metarepresentational Canon. Here Morgan's distinction between higher and lower faculties is read as indicating higher and lower *order* intentional or representational states, and thus exploits the relationship between metarepresentational levels as in, "the clergy know that I know that they know that they do not know" (Ingersoll, quoted in Burghardt 1991, 76), where there are four representational layers. Those following Dennett (1976) and especially his (1987) exposition of the intentional stance, as well as R.W. Mitchell (1986) and his framework for conceptualizing deception, might find this reading of 'higher' and 'lower' compelling. Here the Canon directs us towards the hypothesis making the lowest-order representational ascription. Consider next how this version of the story handles Issue 1 and Issue 2 raised in my introduction.

Morgan's Canon also mentions a psychological scale. Here he is referring to a hierarchical ordering of faculties in which instinct is placed lower than sensory intelligence, which is lower still than reason. I will

https://doi.org/10.1086/505187 Published online by Cambridge University Press

discuss this 'psychic ladder' in more detail further on, but for now, allow me to point out that it hardly entails the metarepresentationalist reading of the Canon. For that to be true, it would have to be the case that the ascending levels of the psychic ladder imply further layers of representational complexity, and there is no reason to agree that this is what Morgan intended.

Well, then, what about Issue (2)? Does the metarepresentationalist reading at least have the virtue of plausibility, regardless of what Morgan had in mind? Consider a version of this view that I will call 'Dennett's Canon', which is found in his article, "Conditions of Personhood." "One should attribute to an organism as little intelligence or consciousness or rationality or mind as will suffice to account for its behavior" (Dennett 1976, 182). There, Dennett claims that this in no way demands we adopt "radical [eliminative] behaviourism" (Dennett 1976, 182), for it only requires explanatory conservativism *within* the practice of mentalistic ascriptions. It requires only "when we adopt the intentional stance toward a thing to ascribe the simplest, least sophisticated, lowest-order beliefs, desires, and so on, that will account for the behavior" (Dennett 1976, 182). This version of the Metarepresentational Canon would require us to minimize metarepresentational complexity, but only when it is already *given* that some intentional-stance explanation is warranted. For example, Dennett suggests we should perhaps say of a "begging" dog, Fido, not that he *wants* his master to *believe* that it is supper time, but only that Fido *wants* the "supper-machine" to do its thing.

However the Metarepresentational Canon is not satisfying for two reasons: first, the caveat restricting the Canon to judgments about intentional hypotheses seems artificial; second, as Dennett (1987, 246–247) himself suggests later on, the emphasis should not be on finding the least sophisticated intentional explanation that will account for the behavior, but simply finding the *best* one, and this may well require an inflation of representational levels.

Taking up the first of these points: why *must* we use the Canon 'as directed' by those such as Morgan, Rollin, and now (early) Dennett, who would have us apply it *only* when adjudicating rival mental hypotheses? For example, it is not clear why we should not take things a step further, and say of the dog that it has merely been conditioned so as to adopt such-and-such a posture whenever such-and-such a stimulus situation is in play. Dennett's reasoning as given above is underwhelming, for he simply asserts that yucky radical behaviorism *can* be avoided if the Canon is used as directed. But this just begs the question presently at issue, for we have been given no reason why we *should* consider the Canon to be inapplicable in nonmental explanatory contexts, other than to point out that this allows us to dodge an undesirable eliminativism. Is there a better

https://doi.org/10.1086/505187 Published online by Cambridge University Press

reason? I think there is, but it is not in virtue of a restricting the Canon to evaluations of only mentalistic explanations. Why should it matter if the lower faculty is not psychological? A nonrepresentational posit is, after all, just another step 'lower' than a first-order intentional state. Examples abound; one need only revisit Vauclair's butterfly in my introduction for a reminder (Sober (1997, 225), by the way, concurs). Behavioristically minded ethologists recognized this, and, as a result, enthusiastically pursued nonmentalistic explanations as a safeguard against anthropomorphism.

One possible worry arising from these considerations is that the Canon might be drawing us into an implausible antimentalism. I do not actually think that this is a real danger, but this concern should be briefly addressed. What I mean is that if by 'higher' we mean 'mental', and by 'lower' we mean 'nonmental', then we might worry that the success of reductive explanations of animal (and human) behavior will ultimately lead towards an elimination of mind from science. The spirit of the Canon, after all, is "deflationary" in the sense that we begin by assuming that the animal is less clever than it seems (Wilder 1996, 33). Clearly the Canon treats higher and lower explanations as, in some sense, in *competition* with each other. We could sum up this attitude with the slogan: *Don't go high if you can go low!* On the assumption that the lowest we can go, in principle, is nonmental, perhaps the Canon calls for the elimination of the mental as soon as a 'deeper' (e.g., physiological) explanation becomes available. Of course, many would insist that epistemological reductionism is a pipe dream, and that the mental will always survive as an autonomous explanatory realm (whereas otherwise heterogeneous classes of behaviors exhibit certain patterns picked out by irreducibly mental concepts). But this is just slightly beside the point. The worrying claim at issue does not depend on the fact of reductionism, but only its counterfactual possibility: What if reductionism were true? Would the Canon require that we eliminate the higher (mental) concepts, and replace them with explanations pitched at lower (nonmental) levels? The answer to this last question partly depends on what is meant by 'higher' and 'lower'. *If* the distinction simply stands for the mental-nonmental distinction, then the Canon might be presupposing that reduction implies elimination. (Perhaps the worry can be soothed by assuming that the Canon does not apply to explanations that are compatible with each other, say as when there are physical realizers of mental faculties.) In any case, on my reading, 'higher' means, not 'mental' per se, but 'emergent'—while 'lower' means 'basal' or 'subvenient'. Further on I will show that since an emergent is a property that cannot be reductively explained, the truth of Morgan's Canon follows practically as a truism. The Canon is correctly, but perhaps trivially, saying that we cannot explain a behavior in terms of what is irreducible, if we

https://doi.org/10.1086/505187 Published online by Cambridge University Press

can already fully explain it in terms of what is reducible. So, the worry that the Canon makes the mistake of assuming reduction-of-the-mental-implies-elimination-of-the-mental is misplaced.

Dennett has since distanced himself from what I am presently calling 'Dennett's Canon'. In *The Intentional Stance*, Dennett does not revive his erstwhile enthusiasm for a Canon-esque methodological principle. Instead, Morgan's "Canon of parsimony" is associated with behaviorism and overindulgent "killjoy" hypotheses (Dennett 1987, 246). Though behaviorists are right that a purely physiological story could always be arrived at, in principle, this is not interesting, he argues. What is important about the intentional stance is that it allows us to see regularities in behavior that would otherwise go unnoticed. Insofar as folk psychology allows us to discover these 'real patterns' in certain systems, intentional ascription yields gains in perspicacity, predictive power, and generalization, and is therefore justifiable. For Dennett no 'stance', whether physical, design, or intentional, is epistemically privileged, apart from the interests of the stance-taker. In the case of intentional psychology, we trade away fine-grained detail in order to purchase access to the high-level patterns in otherwise unrelated and seemingly miscellaneous classes of behavior. Instead of the Canon (or a close cousin), Dennett suggests that we should allow ourselves to both "demote" and "promote" (1987, 269) creatures along the "scale of intelligence" (i.e., grades of metarepresentational capacity) as the evidence leads us (1987, 242). Wilder echoes Dennett's more recent attitude, agreeing that "[i]nflationary accounts—flying in the face of Lloyd Morgan's conservative counsel—might be fecund, yielding insights into patterns of animal behavior otherwise invisible" (Wilder 1996, 37).

Before moving on I should acknowledge that Sober (2001) also discusses what I am calling the 'Metarepresentational Canon' (where there is an a priori preference towards minimizing the number of representational levels we attribute towards an organism). Sober offers an objection similar to Dennett's. Sober says the Metarepresentational Canon (what he calls the "principle of conservatism") can be distinguished from both Morgan's Canon (2001, 226, note 1) and parsimony (2001, 229, 231), though it is a 'descendent' of the latter. He then argues that this conservatism principle is structurally analogous to certain methodological rules offered in evolutionary biology. Sober rejects these biologically based principles, because they try to establish a priori constraints on the evaluation of causal hypotheses. Sober thinks that only empirical tests can adjudicate between competing claims about causes (2001, 232), and that the principle of conservatism is also suspect on the same grounds (2001, 232). When we are considering models of animal behavior of differing representational complexity (e.g., a first-order versus a second-order model), we should

https://doi.org/10.1086/505187 Published online by Cambridge University Press

not start by assuming that either is a priori more likely. As both make claims about the underlying causes of observed behavior, they are on an "epistemic par." What is needed to decide between them is "observational test, not the invocation of principles" (2001, 232). In the absence of this, we should remain agnostic.

Dennett and Sober appear to endorse similar criticisms of the Metarepresentational Canon. First, it seems that this version of the Canon is not simply a version of parsimony, since it is asking to minimize representational levels *even if this means we need to inflate something else*, such as the number of system states (Sober 1997, 230–231). But then, we can ask, *But why should representational levels be given priority when we are thinking of applying parsimony?* Sometimes this might be appropriate, but other times it might be more appropriate to posit a higher degree of representational complexity, in return for fewer system states. In this latter sort of case, it seems, an a priori methodological principle gets in the way of the evidence, and for that reason the Metarepresentational Canon seems to have little to offer.

I turn now to Sober and his (1997) analysis of Morgan's Canon. Here Sober thinks that the Canon is neither a version of the law of parsimony, nor a covert vehicle for the elimination of the mental. If his reconstruction is correct, then "the Canon has been widely misunderstood" (Sober 1997, 239). Sober's view is that the Canon countenances realism about mental states, while acknowledging more fundamental physical explanations of behavior. Higher and lower forms of explanation are taken to be compatible. Sober's assessment crucially turns on his understanding of the terms 'higher' and 'lower', however, I contend that his interpretation is not what Morgan had in mind, and in any case, has undesirable behavioristic implications.

Sober thinks that Morgan's Canon rests on an 'evolutionary' justification that ultimately fails, but Sober is open to the possibility that the Canon might have a nonevolutionary justification. Morgan conceived of three possible 'methods' or distribution patterns among higher and lower faculties (see fig. 1). What is important here is Morgan's remark that, given the Method of Variation, we are "logically bound not to assume the existence of these higher faculties until good reasons shall have been shown for such existence" (Morgan 1894, 9). Sober thinks that this represents a "slippage in Morgan's logic" (Sober 1997, 233). The apparent trouble is that this Method seems to logically allow for an organism to possess a higher faculty in the absence of a lower one—hence the Canon, which explicitly denies this arrangement, could not be a corollary of the Method.

Sober proposes a different justification that can resolve this difficulty.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

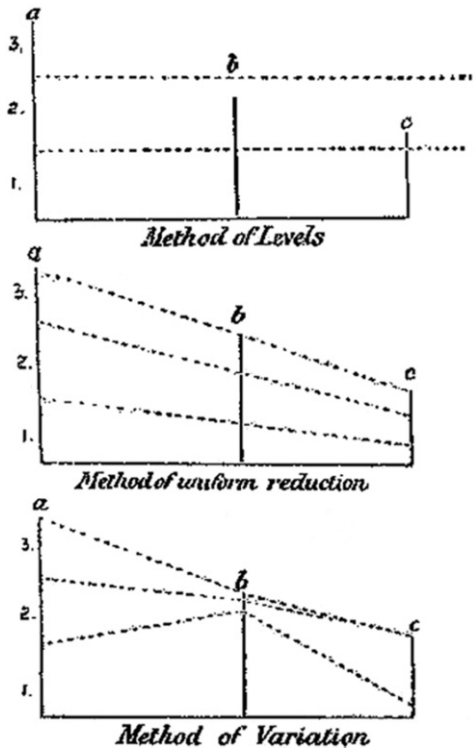


Figure 1. Morgan's depiction of the three Methods (see Morgan 1903, 56).

Sober's idea is to start with the following definition of the relationship between higher and lower traits:

> One internal mechanism is higher than another if and only if the behavioral capacities entailed by the former properly include the behavioral capacities entailed by the latter.[5] (1997, 236)

It follows from this that the probability an organism has a higher faculty must, by necessity, be less than the probability that it has a correspondingly lower faculty. An important consequence of this definition is that it allows for the possibility that an organism has *both* a higher and lower trait for the same behavioral capacity. As Sober says, it is still an open question whether a given organism has only the lower faculty, or whether

5. Sober means that higher faculties entail lower faculties, *but not conversely*. I take it Sober intends "properly include" to be understood in the sense of set theory. If A properly includes B, then A entails B, but not conversely.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

it has both the higher and the lower. This is a significant result because it avoids the implausible presumption that the higher and the lower are mutually exclusive, and likewise the claim that the Canon forces the elimination of explanations positing higher faculties, given the availability of explanations that postulate only lower faculties. By Sober's reckoning, this explains why the Canon has no antimentalist tendency. Sober then argues that this revised Canon (call this 'Sober's Canon') can help us tell whether the probability an animal has *both* higher and lower behavioral faculties is greater than the probability it has only the lower faculty without the higher. This is provided we also make the assumption that "what an organism *can* do, it *will* do, given the right circumstances" (Sober 1997, 237). He then demonstrates how the preference for the lower plays itself out in Sober's Canon.

Sober says suppose we have two psychological faculties *H* and *L* that stand for 'higher' and 'lower', respectively, and two classes of behavioral capacities B1 and B2. Sober asks us to suppose that having *H* allows an organism to have the behavioral capacities B1 and B2, but having *L* only allows an organism to have the set of behaviors in B1. If careful observation reveals that a given organism is incapable of performing B2, then (by the assumption made above—the "can-do, will-do" principle) "we have observed a nonevent" (Sober 1997, 237). We have seen the organism "not perform B2" (1997, 237). It follows, by necessity, that the creature possesses *L,* but not *H* (the inference is $H \supset$ B2 and ~B2, therefore ~*H*). Note that this construal of Morgan's Canon does not involve a nondeductive principle of parsimony: "Deductive logic suffices" (1997, 237). If an organism had *H,* we would observe it perform B2, as well as B1. On the assumption that it can only perform B1, we are left with the conclusion that it only has faculty *L*.

If Sober is right, the Canon is a deductive principle of inference and not a version of parsimony at all. Nor, if he is right, does it "license a wholesale preference for purely physiological explanations over explanations that attribute representational content" (1997, 241, note 11). To see why, consider that different kinds of behavioral explanations (say psychological and physiological) might have *the same* behavioral entailments. Such would be the case when, say, the physiological level is thought to capture the *implementation* of the psychological.

At this point, perhaps it would be useful to say something about why this analysis squares with my earlier discussion of Sober's (2001) negative assessment of the principle of conservatism (the 'Metarepresentational Canon'). They are entirely compatible. In the earlier work, Sober (1997) accepted his reinterpretation of the Canon insofar as it expressed a point of *deductive logic*. In his later work, he (2001) rejects a quite different 'descendent' of the Canon—i.e., the Metarepresentational Canon—on the

https://doi.org/10.1086/505187 Published online by Cambridge University Press

grounds that we cannot draw conclusions about causal mechanisms in cases where the only observations we have to work with are equivocal.

Though Sober thinks Morgan's original evolutionary justification of the Canon fails (1997, 240), when read as a principle of deductive inference, we obtain both the result that lower faculties are explanatorily privileged, and the requirement that higher faculties logically entail the presence of lower faculties. He concludes that "[c]onstrued in this way, Morgan's Canon makes sense, but neither evolution nor parsimony is the reason why" (Sober 1997, 240). Despite this, in my view, Sober's (1997) reconstruction does not capture Morgan's intentions. Moreover, even if we push Morgan aside, "Sober's Canon" does not supply an adequate replacement. I shall take up these last two claims in reverse order.

What then of Sober's Canon? Sober's claim that 'higher' and 'lower' should be understood in terms of logical relations between behavioral capacities is not complementary to the intuition that higher faculties are mental. It should not be possible for the higher of two faculties to be nonmental, and the lower to be mental, for otherwise what is the point in offering the Canon as a principle of antianthropomorphic conservatism? And yet this result is compatible with Sober's reconstruction. Consider the following counterexample: Suppose, as before, we have two faculties H and L that allow for behavioral capacities B1 and B2 in the former instance, but only B1 in the latter. Since the relation between H and L is merely logical, notice that it is undefined as to whether either of these faculties are mentalistic or not. But that means it is consistent with Sober's Canon that *the higher, but not the lower, could stand for a nonpsychological mechanism*. Consider, for example, a creature that is the biological equivalent of a fancy stimulus-response 'jukebox'. What I am considering here is a being possessing a set of input-output contingencies so extensive, and auspiciously contrived, that it is prepared for virtually any environmental exigency (cf. the Blockhead in Block 1981). Now compare this seemingly clever 'behaviorist monster' with another being in possession of a genuine, though rudimentary, doxastic system. What is important here is that you assume this second (simpleminded) organism has fewer behavioral capacities in virtue of its limited cognitive and sensory capabilities.

Here is the point: If Sober is right, then we would have to say that, according to Morgan's Canon, the behaviorist monster described in the first case would possess 'higher' faculties than the humble thinker in case two. This is clearly evident given that the monster's nonpsychological faculties H could allow for behavioral capacities B1 and B2 while the creature with the genuine (though meager) mental capacities L is restricted to behaviors falling under the extension of B1 alone. This is a perverse result.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

I suppose that Sober might reply: "So be it. The Canon should allow for mental faculties to be counted as lower in some cases." But this will not do. Firstly, the result is at odds with Morgan's conception of higher and lower in the sense of a hierarchy of mental faculties. The method of variation explicitly allows for greater behavioral capacities in lower faculties (consider a dog's nose and its superior discriminatory powers as compared to me, or even the automatic control centres in spiders that facilitate web-building). But while these enhanced faculties would make more behavioral capabilities available, they would nevertheless be counted as lower by Morgan. Likewise, Morgan would think it ridiculous to insist that the behaviorist monster is the one with the higher faculties. Morgan's scale of psychological faculties is completely neglected by Sober's analysis.

Returning for now to the second issue I raised in my introduction, let me ask again, *Does Sober's Canon at least offer a reasonable update?* In fact, Sober's version is highly implausible. As my earlier example suggests, this amounts to a *behaviorism* that most would find untenable. It would be quite odd if it were a logical truth that there must be a strict patterning between an organism's higher faculties and its behavioral capabilities. To see this clearly, consider a further twist on my counterexample. Suppose we are studying an organism, and after counting, we decide the number of behavioral capacities for one of its mental faculties is *thirty-five*. In order to undermine Sober's proposal, we only need to think of some ridiculously crude 'lower' behavior-inducing mechanism that is capable of producing at least thirty-*six* responses. For example, suppose we somehow amplify its range of possible behaviors by integrating a tiny *roulette wheel* within its nervous system. Following Sober's Canon, we would have to say that this "roulette wheel head" possesses the higher faculty as compared to any being with fewer behavioral capacities. This is so, regardless of *what internal processes are responsible for producing behavior*. Indeed, any genuine mental faculty with capacities n could always be trumped by a randomizing faculty with a range of possibilities equaling $n + 1$. Of course, a roulette wheel would make for a terrible guidance system. But then, that is the point: Sober's Canon doesn't tell us to count faculties that produce *smart* behaviors as high. It defines 'higher' only in terms of 'more'. For this reason, the connection between possessing a higher faculty and enjoying greater mentality completely breaks down. Given that psychological faculties can fail to correlate with greater behavioral capacities, Sober's reformulation mistakenly permits the low to be raised high, and for the high to be laid low.[6]

6. Cartmill also notices that Sober forgets unconscious neural processes might permit a set of behaviors that properly include the behaviors permitted by mental processes. Cartmill concludes that Sober's Canon "fails to justify the methodological assumption

https://doi.org/10.1086/505187 Published online by Cambridge University Press

Sober's Canon does not capture Morgan's intentions either. Admittedly, Sober was not trying to capture the Canon in its original form, but only to find inspiration in it. But I should briefly emphasize that Morgan does not suggest anywhere in *An Introduction to Comparative Psychology* that these are best thought of as behavioral capacities. Clearly Morgan used these terms in more than one way. Sober notes that 'higher' and 'lower' are used to express Morgan's adherence to the fallacious Spencerian doctrine that evolution marches inexorably towards ever greater complexity (Sober 1997, 225). This is correct, but there is more to Morgan's use of 'higher' and 'lower'. I turn now to my own reading of Morgan.[7]

Morgan often writes as if higher and lower faculties refer to psychological (not behavioral) processes (1903, 59), such as motives (53). Elsewhere Morgan uses 'higher' and 'lower' to describe levels of psychical and evolutionary development (40–41), to compare cortical regions (32), to explain the concept of complexity itself (55), and in referring to contrasts between mental faculties such as "automatic centres" and voluntary "control centres" (41). Similarly, the "vague and contradictory beliefs" of "uncivilised people" and children are ranked "lower" than the "higher" intellectual capacities of "Englishmen" (and women), such as reason and introspection (41–43).

But there is also an explicit scale of particular faculties in which "reason or intellectual thought" is included among "higher acts" and contrasted with "mere intelligence" or "practical sense experience" (1903, 54). Morgan's interpretation of 'higher' and 'lower' is situated within a hierarchically ordered "ladder" of ascending sophistication. Instinct occupies the lowest rung; rudimentary inferential processes, such as associative learning, occupy the middle position; while reason and self-awareness are placed at the very top. This aspect of Morgan's thought was also firmly entrenched prior to his unveiling of the Canon. In "The Limits of Animal

---

that nonhuman behavior should be attributed to unconscious or nonmental causes whenever possible" (2000, 843).

7. Sober also offers that the canon can be contrasted with a different principle he calls "cladistic parsimony" (1997, 226), i.e., don't assume a trait has multiple origins if the evidence is compatible with there being only a single origin. Here Sober considers the idea that the higher-lower distinction corresponds to the distinction between derived and ancestral traits. In a lineage, ancestral traits are present earlier, while derived traits evolve later. But Sober then argues that cladistic parsimony delivers problematic advice, for example, about whether protorothyrids possessed egg sacs. His point is that this revision of Morgan's Canon gives the wrong answer, since there is no "general preference for 'lower' over 'higher'." The evidence might be neutral, or even favor attribution of the 'higher' trait. While I agree that this reading is not representative of Morgan's view, I note that his 1932 formulation of the canon, in which higher and lower are associated with later and earlier stages of development (1932, 262) does at least resemble it.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

Intelligence," Morgan advocates "a threefold division [of mental activity] into instinct, intelligence, and reason" (1893, 225).

The upshot is that Sober's Canon fails to meet our expectations concerning both Issue 1 and Issue 2. I am lead to suspect that it was not behavioral capacities, but rather a very special, though inchoate, sense of 'higher' and 'lower' that Morgan actually had in mind.

I acknowledge that others may still disagree. Dewsbury finds Morgan's distinction vague, and at best, metaphorical (Dewsbury 2000). Cartmill (2000, 841) contends that 'lower' means 'historically prior', in reference to the story of human evolution, and 'shared with other species'. For Cartmill, the Canon demands that we avoid interpreting animals as humanlike, and though Sober would not follow Cartmill in claiming 'higher' means 'distinctively human', Cartmill's reading is somewhat close to the reinterpretation of the Canon as a claim about cladistic parsimony (see my note 7). The analysis I wish to offer differs from these views. A careful examination of Morgan's works shows that his higher-lower distinction is ineluctably enmeshed in his doctrine of the psychic ladder. The final step in my analysis is to show how Morgan's ideas about emergence/supervience provide context for his Canon.

As discussed by Kim (1992, 1993), Morgan's later work, *Emergent Evolution* (1923), includes one of the earliest discussions of nonreductive psychophysical supervenience. There the terms 'higher' and 'lower' express a certain kind of dependence relation between two faculties. Morgan explicitly calls this special kind of dependence 'emergence' or '*supervenience*'. Mind, which is higher in terms of complexity and evolutionary development emerges or supervenes when certain 'basal' conditions of life, that is, certain biochemical processes, are present. In addition, Morgan claims that higher mental faculties emerge from, or supervene on, lower mental faculties.

It is commonplace today to say that supervenience is supposed to express some kind of dependency of the mental on the physical. As Kim puts it, "The core idea of supervenience is [that] things which are alike physically must be alike mentally" (Kim 1992, 132). I will follow Kim (1993, 140; 1995, 576) by taking the following three desiderata to be characteristic of a supervenience relation. Kim sets these out as: (1) *property covariation*, which is to say that indiscernibility with respect to the basal properties requires indiscernibility with respect to the supervenient properties; (2) supervenient properties are asymmetrically dependant on, or determined by, the subvenient set; (3) finally, the supervenient properties are, in some sense, 'irreducible'. What does this last condition involve? Kim (1992, 128ff.) points out that it depends on what sense of reduction one takes for granted. For 'anomalous monists', such as Davidson, mental predicates cannot enter into strict laws of any form (i.e., there are neither

https://doi.org/10.1086/505187 Published online by Cambridge University Press

laws that are purely psychological, nor are there psychophysical laws). Other antireductionists are motivated by the 'multiple realizability' of mental states, usually meaning they are only committed to the weaker claim that there cannot be Nagelian 'bridge laws' connecting the mental and the physical. Kim (1992, 126) says that Emergentists, in contrast, probably would have accepted both the reality of exceptionless mental-mental laws, as well as domain-crossing bridge laws, making comparison with contemporary antireductionists uncertain. But Emergentists at least share the idea that some aspect of the mental eludes capture in the physicalist net, and, despite this, the mental is dependent on the physical. Even if we possessed nomological statements connecting mental and physical phenomena, there would still be a gap in our understanding of the mental. We would still not understand *why* those mental aspects are related to those physical ones. In this sense, Emergentists believed that "relationships between emergents and their bases are essentially inexplicable, something we must accept, 'with natural piety'" (Kim 1992, 126).

Kim (1993, 133, 136) has also wondered who first articulated the concept of supervenience in a form close to the three desiderata. Consider the standard version of the story, in this case related by Bickle (1998, 8): "Most advocates of this notion . . . blithely (and correctly) trace it back to G. E. Moore's ethical writings." Bickle then adds in a footnote (1998, 213): "Citing Moore in this context goes back to Davidson (1970). Davidson is universally acknowledged as having introduced supervenience into the philosophy of mind." However Kim upsets this account by crediting Morgan with introducing it to the philosophy of mind. With the exception of Leibniz, Kim cannot find "any other [philosophical] occurrence of the term . . . until we come well into the present century" (1993, 136) and reach Morgan's *Emergent Evolution* published in 1923.

Morgan and other early twentieth century "Evolutionary Emergentists" were distinguished by their rejection of both mechanistic and vitalistic accounts of life and other 'higher' phenomena. Kim (1992, 122ff.) summarizes their position as involving three subdoctrines. These are (1) Ultimate Physical Ontology, (2) Property Emergence, and (3) the Irreducibility of Emergents. Taking these in order, these say that (1) There are fundamental nonemergent properties that are physical. (2) Basal level conditions alone suffice for emergents to appear. As Kim says, "all aspects of a given thing, or even of the whole world, are fixed once its total physical character is fixed," and in this sense Emergentistism holds that physical indiscernibility entails mental indiscernibility, and hence "entails a *supervenience doctrine*" (1996, 124). (3) Emergents are "genuinely novel" and "irreducible" with respect to their basal conditions (Kim 1993, 134). Even if we know, after the fact, that a certain emergent inevitably results from certain basal conditions, there can be no explanation for this rela-

https://doi.org/10.1086/505187 Published online by Cambridge University Press

tionship; it is "essentially inexplicable" (Kim 1992, 126). Kim (1992, 127, note 15) sees this doctrine as akin to the sense in which contemporary Mysterians speak of an "explanatory gap" between qualia and objective modes of understanding conscious experience. The Mysterians maintain that "we may know all about the neurobiological and physicochemical properties of a given process and yet have no idea what sensory quality, if any, to expect" (1992, 128).

According to Kim (1993, 134; 1995, 576) there is a "surprising" and "striking" similarity between Morgan's supervenience and the contemporary sense, as defined in terms of the three desiderata. There are patterns of property covariation between emergents and their basal conditions; there is an asymmetric dependence of the emergents on the lower levels; and the emergents themselves are "novel" in the sense that they are not (initially) predictable from the basal conditions, and in this they give rise to an explanatory gap. I follow Kim in saying that the central Emergentist doctrines rather neatly overlap with the core sense of supervenience defined in terms of property covariation, asymmetric dependence, and irreducibility. Kim concludes that "emergentism can be regarded as the first systematic formulation of non-reductive physicalism" (1995, 576).

Next, I wish to offer two further refinements to Kim's revised account of the early history of supervenience. First, Morgan's articulation of the modern supervenience concept can actually be traced back to his work in the late nineteenth century. Second, I propose that there is an important linkage between Morgan's use of the concept of supervenience and his methodological Canon of ethology. What I mean is that the terms 'higher' and 'lower' in the Canon ought to be read as something close to what we would today call 'supervenient' and 'subvenient'. In short, it would appear that a proper understanding of Morgan's Canon requires an acknowledgement of his commitment to psychophysical supervenience. This interpretation is reinforced by an examination of Morgan's work both before and after the Canon was formally introduced. A summary of this evidence includes the following: (1) In his 1890 *Animal Life and Intelligence*, Morgan interweaves a discussion of higher and lower mental faculties with ideas about psychical emergence (and on at least one occasion there he even uses 'supervenes' as a stylistic variant for 'emerges'). There he also discusses an unnamed Canon-like principle which serves to adjudicate between competing explanations of animal behavior, when these are cast in terms of higher versus lower psychical faculties. (2) I mentioned earlier that Morgan explicitly expresses the Canon in terms of emergence (1925), and even places it in the context of mind-body supervenience (1932). (3) Finally, the 'Supervenience Canon' reading makes sense of Morgan's otherwise puzzling claim that we are *logically bound* not to

https://doi.org/10.1086/505187 Published online by Cambridge University Press

assume the existence of higher faculties until good reasons have been shown for such existence.

Morgan's 1890 publication offers perhaps the earliest use of 'supervenience' with reference to the mind-body problem. Interestingly, this work only barely preceded his formal introduction of the Canon in 1892.[8] The connections between his protosupervenience and the Canon seem to be present from the beginning.

Morgan is also committed to a "scale of life" in which organisms and their characteristics may each be classified as either higher or lower (1890, 333, 338, 357). Morgan speaks again and again of higher faculties "arising out of" (319) or being "built up" from lower faculties, that is, they "emerge" from these basal conditions (322). Though higher faculties, such as abstract thought, judgment (330), conceptual inference, reason (331), and "emergent consciousness" (338), arise out of (relatively) lower faculties, such as "analysis of phenomena" (331) and "constructive experience" (322), or just "intelligence" (361), "exactly when [mental processes] do emerge in the ascending branches of the great tree of animal life it is exceedingly difficult, if not quite impossible, to determine" (337). At the high end of the scale, "through the aid of language, the higher faculties [conceptualization and perception] have been developed and evolved from the lower faculties [e.g., "unconscious inference" or mere physiological discrimination]" (362). Moreover, an emergent "constitutes in the possessor a new mental departure, which we may describe and constituting, not merely a specific, but a generic difference from lower mental activities" (350). At last, reason (what he calls 'conceptual analysis'), another "new departure in psychological evolution" is said to depend on language (373ff.), and is "grafted" or "supervenes" on lower faculties, such as perception and (associative) intelligence (373). A fair summary of this portion of *Animal Life and Intelligence* would be to say that various mental functions or faculties depend on, but are not reducible to, that is, they emerge from or supervene on, their subvenient precursors. Kimler also notes (2000, 856) that as early as 1890, Morgan accepted that mind is identical to brain, though somehow not reducible to physical processes. Morgan's use of the terms 'higher' and 'lower' already seems to express at least a "proto"supervenient conception of mental relations before the Canon is even introduced.

It is clear that Morgan (1890) also thinks that given this supervenient or emergent metaphysic of mind, we are best to be guarded in our assessments of animal intelligence. For instance, just because an African grey parrot stands upright whenever its guardian utters the phrase "live

8. According to Radick (2000, 3, note 1), Morgan introduced the Canon in the summer of 1892 at the Second International Congress of Experimental Philosophy in London.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

for your master," it does not follow that it understands the words, since the bird could have been taught the same trick with any arbitrary sound, such as "Hocus Pocus" (1890, 353–354). Morgan then makes a pregnant suggestion. As a general rule: "The principle I adopt is to assume that the inferences are perceptual, unless there seem to be well observed facts which necessitate the analysis of the phenomena, the formation of isolates and therefore the employment of reason" (1890, 363). Here he seems to be close to saying that if it is possible to account for an animal's behavior in terms of lower processes, such as stimulus-response conditioning, then we are not entitled to infer that it is capable of higher processes, such as reasoning.

Later he provides the formulation cast in terms of a scale of emergence (1925, 61), and still later he again asserts his belief in a supervenient dependence of higher properties on lower properties (1932, 249). The emergence of self-awareness—what Morgan calls "reflective organization" (1932, 258) from consciousness is accepted, and here Morgan even declares that it is "*[s]upervenient on perceptive organization*" (1932, 258). With this background in place he then invokes still another version his famous Canon:

> [O]ne may distinguish at least three salient stages in mental evolution—provisionally named percipient, perceptive, and reflective (subconscious, conscious, and self conscious) . . . since this scheme is professedly genetic and evolutionary, one should not on this basis interpret a lower form of behaviour as implying a higher level of mentality than the evidence demands. In evolutionary development regard should proceed from below upwards. (Morgan 1932, 260–261).

This is not to say that there are no differences between Morgan's supervenience and the contemporary conception. Kim and others typically conceive of supervenience as a synchronic relation, that is, where the *timing* of mental and physical events is simultaneous. But many of the passages above suggest that, for Morgan, supervenience is a diachronic relation in two senses. Later stages supervene on earlier evolutionary stages, and faculties also diachronically supervene during the course of an individual organism's development. This is just to apply the distinction between phylogeny and ontogeny to derive two versions of supervenient dependence, and might even account for the persuasiveness Sober's cladistic suggestion about the Canon. Perhaps the relationship between derived and ancestral traits is captured by the phylogenic variety of supervenience.

Morgan's view in *An Introduction to Comparative Psychology* also shares the contemporary conception's commitment to a synchronic form of dependence. Recall that Morgan's Method of Variation allows for the

https://doi.org/10.1086/505187 Published online by Cambridge University Press

presence of lower mental faculties in the absence of higher mental faculties. It is *not* assumed that the germ of a higher type must be present in lower types (Morgan 1903, 58). There is, however, a synchronic dependence of the higher on the lower. Hence "sense-experience . . . forms the foundation of our psychic life" (157) since it is "that upon which the superstructure of *higher* psychical faculty is based" (159, emphasis in original). This serves as foundation not simply in an evolutionary, but also physiological, sense:

> [N]o consciousness accompanies the transmission of impulses along the nerves; but . . . when the molecular thrill reaches the brain, or some part of it (probably the cerebral hemispheres), there consciousness emerges. One of the conditions of consciousness, therefore, seems to be the occurrence of certain molecular vibrations or disturbances in certain specialised nerve-centres within the cerebral hemispheres. (Morgan 1903, 28)

Consciousness depends on the "material structure of the brain" as the "vehicle of its manifestation" (Morgan 1903, 329).

Another difference between Morgan and more recent expositions of supervenience deserving emphasis is that he maintained psychical faculties supervened on *each other* as well as on purely physiological basal conditions. Psychological mechanisms come in relatively higher or lower grades and stand in synchronically emergent relationships, as in the relation between lower 'impressions' and higher 'percepts'. So, the "suffusion of all our impressions with more or less of a relational tone is what raises them to the level of percepts—the percept being an impression set in a relational background" (Morgan 1903, 267). Similarly, "sense-experience affords the raw material out of which all our higher conceptual thought is elaborated" (284–285).

Once we acknowledge that psychic faculties are also supposed to supervene on each other, we can clarify an otherwise puzzled response to my interpretation. For the link between the Canon and emergence to be useful, it should help us decide between, say, reason versus instinct. Yet given that these are both emergents, and so, 'higher', the Canon might seem to offer no guidance at all. But, I maintain, since Morgan seems to be claiming that faculties can be *relatively* higher or lower than each other, it is perfectly understandable how the Canon could lead us to favor one of these hypotheses over the other.

There is one loose end I want to address before moving on. Sober argued that Morgan made a "slippage in logic" when he claimed (1997, 232) that the Canon is entailed by the so-called Method of Variation. Sober is right: this Method allows for an organism to have a higher faculty in the absence of a lower faculty. This is why Sober (1997, 233) says that

https://doi.org/10.1086/505187 Published online by Cambridge University Press

Morgan's evolutionary justification for the Canon fails. And, indeed, Sober's Canon dispenses with this problem. If higher entails lower (as with behavioral capacities), then there is no logical possibility of a higher faculty in the absence of the lower. Sober's problematic example would also be incompatible with my Supervenience Canon reading, I should add.

Yet I do not think Sober is correct in assuming that the sole justification for the Canon rested with the Method of Variation (MV). Rather, this Method itself takes its justification from Morgan's ideas about emergence. Morgan thinks the MV is what we expect, given the emergent relationships between higher and lower faculties on Morgan's psychic ladder. We expect a priori that evolution by natural selection will require the widest latitude in conferring faculties to organisms in differing degrees in order to meet the multifarious ecological problems faced by cognitive beings. But we also *empirically* confirm this armchair presumption: "Hence it is the third method . . . which we should expect *a priori* to accord most nearly with the observed facts. And so far as we can judge from objective observation . . . this would appear to be the case" (Morgan 1903, 57–58). Kimler also notes that Morgan is well aware that we observe "differences in senses and behaviours. Animals have *different* minds because they are different structurally, and adaptively evolved to solve different ecological problems" (Kimler 2000, 857). I suggest that the MV is intended to be *constrained* by this Emergentist framework, and that is why he is quite right to say that *given* an organism "may be at a stage where certain higher faculties have not yet been evolved from their lower precursors," theorists are "logically bound not to assume the existence of these higher faculties until good reasons shall have been shown for such existence" (Morgan 1903, 59).[9] Thus, *given* an emergence-supervenience thesis, one is logically bound not to assume the existence of mind in the absence of the subvenient conditions.

With this, I conclude my investigation into Morgan's intentions. His use of the terms 'higher' and 'lower' in his Canon of ethology have indeed been subject to serious misunderstandings. What he meant by this distinction closely approximates the contemporary antireductionist's use of the terms 'supervenient' and 'subvenient'. A higher faculty, in the most general sense, is one that depends on a lower faculty, yet which cannot be reduced to the lower. The seeds of Morgan's later ideas about super-

9. Of course, one might argue that the other interpretations of Morgan's Canon can also accommodate this remark. For instance, perhaps we could be logically bound not to ascribe a higher-level representational capacity (e.g., the ability to token the belief that one believes that *P*) unless we also ascribe certain lower level representational capacities (e.g., the ability to token the belief that *P*).

https://doi.org/10.1086/505187 Published online by Cambridge University Press

venient dependence would seem to be already implicit in the formulation of his famous Canon.

Finally, *is Morgan's Canon, as I have explained it, useful to cognitive ethology?* In order to wield the Canon, we need to know which faculties are emergent, and then we need to know what would count as an adequate explanation in terms of emergent versus nonemergent faculties, and besides that, in light of the rider Morgan later adds to the Canon, we also need to know what counts as 'independent evidence' for this or that mental capacity. I am not sure how to proceed with this advice. Perhaps there is some principled way to unpack each of these points, but my sense is that the issues involved are quite tricky. An even more pessimistic response would be to say that the Canon is merely stating a truism. *Of course* one should not posit irreducibly emergent faculties, if one can already explain matters in terms of reducible 'resultants'. Again: emergent faculties seem to be just those that resist reductive explanation.

REFERENCES

Allen, Colin, and Marc Bekoff (1997), *Species of Mind: The Philosophy and Biology of Cognitive Ethology*. Cambridge, MA: MIT Press.

Bekoff, Marc, and Colin Allen (1997), "Slayers, Skeptics, and Proponents", in R.W. Mitchell, N. Thompson, and L. Miles (eds.), *Anthropomorphism, Anecdotes, and Animals: The Emperor's New Clothes*? Albany, NY: SUNY Press, 313–334.

Bickle, John (1998), *Psychoneural Reduction: The New Wave*. Cambridge, MA: MIT Press.

Block, Ned (1981), "Psychologism and Behaviorism", *Philosophical Review* 90: 5–43.

Boring, Edwin G. (1950), *A History of Experimental Psychology*. 2nd ed. New York: Appleton Century Crofts.

Burghardt, Gordon M. (1991), "Cognitive Ethology and Critical Anthropomorphism: A Snake with Two Heads and Hognosed Snakes That Play Dead", in C. A. Ristau (ed.), *Cognitive Ethology: The Minds of Other Animals*. Hillsdale, NJ: Erlbaum, 53–90.

Cartmill, Matt (2000), "Animal Consciousness: Some Philosophical, Methodological, and Evolutionary Problems", *American Zoologist* 40: 835–846.

Costal, Alan (1993), "How Lloyd Morgan's Canon Backfired", *Journal of the History of the Behavioral Sciences* 29: 113–122.

Davidson, Donald (1970), "Mental Events", in L. Foster and J. W. Swanson (eds.), *Experience and Theory*. Amherst: University of Massachusetts Press, 79–101.

Dennett, Daniel C. (1976), "Conditions of Personhood", in Amélie Rorty (ed.), *The Identities of Persons*. Berkeley: University of California Press, 175–196.

——— (1987), *The Intentional Stance*. Cambridge, MA: MIT Press.

de Wall, Frans (2001), *The Ape and the Sushi Master*. New York: Basic Books.

Dewsbury, Donald A. (2000), "On Misrepresenting Lloyd Morgan's Canon: Comment on Thomas", http://htpprints.yorku.ca/archive/00000044/00Dewsbury-on-Thomas.htm; accessed on 4/29/04.

Fodor, Jerry (1999), "Not So Clever Hans", *London Review of Books* 21: 12–13.

Galef, Bennett G., Jr. (1996), "Historical Origins: The Making of a Science", in Lynne D. Houck and Lee C. Drickamer (eds.), *Foundations of Animal Behavior: Classic Papers with Commentaries*. Chicago: University of Chicago Press, 5–12.

Kim, Jaegwon (1992), "'Downward Causation' in Emergence and Non-reductive Physicalism", in Ansgar Beckermann, Hans Flohr, and Jaegwon Kim (eds.), *Emergence or Reduction?—Prospects for Non-reductive Physicalism*. Berlin: de Gruyter, 119–139.

https://doi.org/10.1086/505187 Published online by Cambridge University Press

——— (1993), "Supervenience as a Philosophical Concept", in Jaegwon Kim (ed.), *Supervenience and Mind*. Cambridge: Cambridge University Press, 131–160.
——— (1995), "Supervenience", in Samuel Guttenplan (ed.), *A Companion to the Philosophy of Mind*. Cambridge: Basil Blackwell, 575–583.
Kimler, William C. (2000), "Reading Morgan's Canon: Reduction and the Unification in the Forging of a Science of the Mind", *American Zoologist* 40: 853–861.
Mitchell, Robert W. (1986), "A Framework for Discussing Deception", in Robert W. Mitchell and Nicholas S. Thompson (eds.), *Deception: Perspectives on Human and Nonhuman Deceit*. Albany, NY: SUNY Press, 3–40.
Morgan, C. Lloyd (1890), *Animal Life and Intelligence*. London: Edward Arnold.
——— (1892), "The Limits of Animal Intelligence", in *International Congress of Experimental Psychology, Second Session*. London: 44–48.
——— (1893), "The Limits of Animal Intelligence", *Fortnightly Review* 54: 223–239.
——— (1894), *Introduction to Comparative Psychology*. London: Walter Scott.
——— (1903), *Introduction to Comparative Psychology*. 2nd ed. London: Walter Scott.
——— (1906), *The Interpretation of Nature*. New York: G. P. Putnam.
——— (1923), *Emergent Evolution*. London: Williams and Norgate.
——— (1925), *Life, Mind, and Spirit*. New York: Holt.
——— (1932), "C. Lloyd Morgan", in Carl Murchison (ed.), *A History of Psychology in Autobiography*, vol. 2. Worchester, MA: Clark University Press, 237–264.
Newbury, Edward (1954), "Current Interpretation and Significance of Lloyd Morgan's Canon", *Psychological Bulletin* 51: 70–74.
Radick, Gregory (2000), "Morgan's Canon, Garner's Phonograph, and the Evolutionary Origins of Language and Reason", *British Journal of the History of Science* 33: 3–23.
Radner, Daisie, and Michael Radner (1989), *Animal Consciousness*. Buffalo, NY: Prometheus Books.
Regan, Tom (1983), *The Case for Animal Rights*. Berkeley: University of California Press.
Rollin, Bernard E. (1989), *The Unheeded Cry: Animal Consciousness, Animal Pain and Science*. Oxford: Oxford University Press.
Skinner, B. F. (1938), *The Behavior of Organisms*. New York: Appleton Crofts.
Sober, Elliott (1988), *Reconstructing the Past: Parsimony, Evidence, and Inference*. Cambridge, MA: MIT Press.
——— (1997), "Morgan's Canon", in Denise Cummins and Colin Allen (eds.), *The Evolution of Mind*. Oxford: Oxford University Press, 224–242.
——— (2001), "The Principle of Conservatism in Cognitive Ethology", in Denis M. Walsh (ed.), *Naturalism, Evolution and Mind*. Cambridge: Cambridge University Press, 225–238.
Thomas, Roger K. (1998), "Lloyd Morgan's Canon", in Gary Greenberg and Maury M. Haraway (eds.), *Comparative Psychology: A Handbook*. New York: Garland Press, 156–163.
——— (2001), "Lloyd Morgan's Canon: A History of Misrepresentation", http://htpprints.yorku.ca/archive/00000017/00/MCWeb.htm; accessed on 4/29/04.
Vauclair, Jacques (1996), *Animal Cognition: An Introduction to Modern Comparative Psychology*. Cambridge, MA: Harvard University Press.
Wilder, Hugh (1996), "Interpretive Cognitive Psychology", in Marc Bekoff and Dale Jamieson (eds.), *Readings in Animal Cognition*. Cambridge, MA: MIT Press, 29–45.

https://doi.org/10.1086/505187 Published online by Cambridge University Press