Normative Characterization in Biological and Cognitive Explanations*

Mark Bauer

Received: 27/06/2014
Final Version: 10/12/2014

BIBLID 0495-4548(2015)30:2p.271-286
DOI: 10.1387/theoria.11957

ABSTRACT: Normative characterization is a commonplace feature of biological and cognitive explanation. Such language seems to commit the biological and cognitive sciences to the existence of natural norms, but it is also difficult to understand how such normativity fits into a natural world of physical causes and forces. I propose to map normativity onto systems stabilized by counteractive constraints. Such a mapping, I believe, can explain normativity's causal-explanatory role in biological and cognitive inquiry. The common approach in the literature is to derive an account of natural normativity by way of a particular theory of function. I avoid that approach here and attempt to address directly the sort of physical systems that might satisfy naturalizing criteria for normativity. This has the advantages, I think, of allowing an account of normativity without first having to decide the correct theory of function as well as allowing for the theoretical possibility that normative and functional explanation might come apart within empirical explanation.

Keywords: natural normativity; biological norm; cognitive norm; function; biological explanation.

RESUMEN: La caracterización normativa es un rasgo común de las explicaciones biológicas y cognitivas. Ese lenguaje parece comprometer a las ciencias biológicas y cognitivas con la existencia de normas naturales, pero es difícil entender cómo dicha normatividad encaja en un mundo natural de fuerzas y causas físicas. Propongo representar la normatividad en sistemas estabilizados mediante condicionamientos opuestos. Esta representación, en mi opinión, puede explicar el rol explicativo-causal de la normatividad en la investigación biológica y cognitiva. Lo usual en la literatura sobre el tema consiste en derivar una explicación de la normatividad natural a partir de una teoría particular sobre la noción de función. Esquivaré este planteamiento e intentaré abordar directamente el tipo de sistemas físicos que podrían satisfacer criterios naturalizadores de la normatividad. Esto tiene las ventajas, pienso, de que permite una explicación de la normatividad sin tener que decidir primero la teoría correcta de la función, y en segundo lugar, queda abierta la posibilidad teórica de que la explicación normativa y funcional puedan tratarse separadamente dentro de la explicación empírica.

Palabras clave: normatividad natural, norma biológica, norma cognitiva, función, explicación biológica.

1. Introduction

Natural phenomena are often normatively characterized in the biological and cognitive sciences: e.g., livers can be well-formed/malformed; hearts can function properly/improperly; fawns can hide well/poorly; penguins can identify/misidentify their young; ants can calculate/miscalculate their route home; or, deer can perceive/misperceive a predator. The previous activities, behaviors, and structures are characterized or described by how well they

* Thanks to Candice Shelby, Gabriel Zamosc-Regueros, and the anonymous referees for many helpful comments and suggestions.

match up to or correspond with some standard or norm, some specification of how the action, performance, or structural design ought to be. The extent to which some item meets a standard or norm, that item's normative status, seems often to serve a causal-explanatory function in these sciences: the fawn's poor hiding behavior explains its capture by an eagle; poor sperm motility explains a male's low fecundity; a poorly developed heart valve explains the low viability of its possessor; and, misperceiving glass as open space explains avians' tendency to fly into windows. In these types of explanation, the failure of the item to be as it ought seems to have a causal function in explaining some further natural phenomenon.

The biological and cognitive sciences seemingly assume that normative status is a causal-explanatory feature of the natural world. *Prima facie* such an assumption seems rather odd: normative status does not seem readily to fit into the picture of the natural world otherwise presented by the sciences, a natural world of physical forces, chemical reactions, metabolism, and so on. The oddity of that assumption would be quickly made a non-issue if we were to adopt a global non-literal construal of the normative language present in these explanations. Davies (2001), for example, suggests that we do just that: the pervasive presence of normative language within the biological and cognitive sciences is, he thinks, a mere artifact of human psychology, just a cognitive trick for the human mind to grapple with complex phenomena. Yet that easy remedy itself seems to be at odds with the broader picture of the natural world presented by the sciences.

Along with electrons, gravitational fields, and enzymes, that world purportedly contains a diversity of organismal representational systems. As evidenced by much of the cognitive ethological literature, such representational systems seem to play a significant role in explaining a broad range of biological behavior and response. Unlike pure information-bearing systems, the tokened states of representing systems, whether in the skull or out in public, are supposedly contentful independent of whether the relevant represented state exists. That content and, say, truth-value can come apart makes possible representational and cognitive errors: e.g., misperceivings, false believings, false signals, or mistakes in categorizing, judging, or evaluating. Since this independence between the content of the representing and the status of the represented seems to be a constitutive and distinguishing feature of representing systems, positing a representing system seemingly presupposes that normative status is a genuine feature of the natural world. Or, simply representational activity just seems to be the sort of activity that can be performed rightly/wrongly, correctly/incorrectly, or well/poorly. A global non-literal construal of the normative language present in biological and cognitive explanation would seem then to eliminate the possibility of genuine representational and cognitive errors, and that would seem to imply, in turn, that there just are not really representational systems at all. While representational explanations of organismal behavior might be supplanted by alternative explanatory forms (e.g., by dynamic systems or connectionist explanations), the oddity of their presuppositions (i.e., that cognitive and representational systems literally commit mistakes) is no reason on its own to reject the representational explanations offered as literally true.

The state of play seems to be this: ascribing normative status to natural phenomena is both prima facie at odds with the picture of the natural world presented by the sciences and integral to it. What seems to be needed is some way to map the normative language present in biological and cognitive explanation onto otherwise accepted natural processes in such a way that the normative language can play its assigned causal-explanatory role. I want to suggest here one way to do such a mapping. Normative language looks, I think, to map well onto certain sorts of stabilized systems, and stabilized systems of these sorts seem to be the key targets of biological and cognitive inquiry.

I will take a somewhat different approach than is common in the literature. Concerns about the naturalistic status of normative ascription tend to be contained within concerns about the naturalistic status of functional ascription. This focus is reasonable. Functional ascription is a prominent feature of biological and cognitive explanation, and functions seem to be the sort of thing that might be performed more or less well. *Prima facie* a theory of natural function should accommodate normative characterization. I propose to go forward here, however, with minimal attention to function. The intent is to methodologically isolate the normative question from the functional one. The reason to take this methodological approach is two-fold. First, it might provide a way to address directly the question of natural normativity without having first to adjudicate the contentious space of functional theory. Second, it provides a way to avoid a common but, I think, questionable assumption, namely that normative characterization is limited to functional explanation. Given its commonality, let me briefly motivate why that assumption is better avoided at the outset.

In at least the mundane case of human cultural norms, the normative and the functional hypothesis are independent hypotheses. We can recognize the presence of a wide range of human standards, e.g., standards of dress, greeting rituals, personal distance, gender specific behavior, pronunciation, etc., without first identifying any particular function that such a standard might serve. Further, we can appeal to the presence of a cultural norm in order to explain the stability of communal conforming behavior as well as responses to deviance without first hypothesizing any function. And, while it is perfectly plausible that a given norm performs a social function, it is also plausible that a norm might perpetuate independent of any particular function. For example, a standard of dress might remain stable despite shifts in function, e.g., shifting from the function of gender identification to profession identification. Or, a standard of pronunciation can emerge in an isolated community without any function; it might simply be the consequence of language learners conforming behavior to expert speakers. That standard of pronunciation might come to possess a function, e.g., in-group identification, as our isolated community has increasing contact with others. In at least the ordinary and mundane case of human social norms, the norm need not be derivative from function.

The two can apparently come apart, because the human cultural norm is often seemingly just a certain sort of replicated behavior. For standards of dress, personal distance, greeting rituals and the like to emerge and perpetuate, what is required are replicating mechanisms, namely social learning mechanisms which can conform naïve and deviant members to displayed behaviors. Just as the stability of a phenotypic trait in a population may be solely the result of the replicative engine, so too human cultural norms may emerge and be sustained by psychological replicative or conforming mechanisms independent of function.

There is no reason to suppose that the full range of human cultural norms are maintained or perpetuated through high level cognitive functioning or anything like explicit conscious judgment. Imitative and emulative mechanisms more than suffice to standardize accents, surface grammar, greeting rituals and the like. But, imitative and emulative mechanisms for social learning are by no means the sole province of the human animal. In fact, the biological literature already provides a strong analogue of the human cultural norm, namely the "tradition". A tradition is just any socially proliferated trait or behavior – that is, it is a trait proliferated through interaction with conspecifics and not through reproduction.[1] A wide range of more or less cognitive mechanisms have been identified as proliferating traits socially within biological populations. For example, local groupings of rats exhibit food traditions which proliferate through behaviors such as responding to odorants on the breadth of encountered rats. (Galef & Laland 2005) Or, there are the localized song dialects of hummingbirds (Apodiformes), parrots (Psittaciformes), and the true songbirds (oscine Passeriformes) which proliferate through imitative mechanisms as well as the fine-tuning of inherited auditory song templates to presented songs.[2] (Janik & Slater 2003; Shettleworth 1998, chap. 10) Or, imitation and emulation might account for a wide array of localized behaviors within chimpanzees. (Whiten et. al. 2001). The key to the emergence of a tradition is the presence of a social learning mechanism, e.g., imitation or emulation, that replicates and proliferates behavior exhibited by conspecifics. Whether a given tradition serves a function as well is a further independent empirical hypothesis. For example, the distinctive rain dances performed by chimpanzee troops might lack any function. (Nakamura 2002))

My point is not that biological traditions are examples of norms. They may or may not be. Instead, the point is just to notice that it should not be taken as a given that normative characterization is the servant of functional explanation. Human cultural norms come apart from function, because the presence of the human cultural norm is often seemingly just dependent on the presence of certain behaviorally conforming mechanisms. Plausibly the mechanisms implicated in many human cultural norms do not involve higher level cognitive capacities but involve social learning mechanisms shared by many other organisms. If so, plausibly normativity might come apart from function in biological explanations extending past human behavior for just the same reasons as they come apart in the human case. So, if we are going to take the possibility of natural normativity seriously, I think that it is better at the outset to avoid the assumption that normativity's role is limited to functional explanation.

Here's the plan. I will first set out some naturalistic criteria specific to normativity. I will then offer an abstract model of a causal set-up built for the express purpose of satisfying those naturalistic criteria. That model is, however, just an abstract model. The next step will be to show that the model is realized by systems at the core of biological inquiry and that its realization by such systems can account for normativity's apparent causal-explanatory role. Last, given this model for natural normativity, I will address its potential implications with respect to naturalistic theories of function.

2. *Naturalistic criteria*

Before presenting a model of natural normativity, it would help to have in place some criteria for appraising such a naturalistic model. I will suggest two general criteria for any naturalizing project and two further criteria specific to normativity.

Call any explanatory-cum-classificatory practice to be naturalized the 'target practice'. Plausibly any naturalizing project should provide 1) a means to map some range of the tar-

[1] See Fragaszy& Perry (2003) for an excellent introduction to the biological investigation of traditions in non-human populations.

[2] According to Janik and Slater (2003), these three groups represent half of all present bird species.

get practice's phenomena onto otherwise accepted natural phenomena and 2) that mapping should accommodate the explanations offered by the target practice. A model need not provide for a perfect or complete mapping between the target practice and otherwise accepted natural phenomena. An incomplete mapping, a mapping that only accounts for some of the target practice, might result for a couple of reasons: 1) the target practice has overextended its classificatory practice to cover phenomena which simply do not fit into the natural world; or 2) the target practice classifies phenomena in such a way that no single mapping to other phenomena will do. A model that only maps some of the target range is not wrong or mistaken; it just simply isn't the end of the story. A model is not falsified by pointing out that it does not cover this or that case. Models are, instead, accepted or rejected given their comparative virtues with other offered models.

To extract plausible naturalistic criteria specific to normativity, I will concentrate on normativity's paradigmatic role in evaluative judgment. In that role, there is nothing particularly mysterious about normativity. A norm is simply a represented measure or standard used to evaluate states of the world. The normative status ascribed to some item (that is, x is good/bad, is right/wrong, is correct/incorrect, etc.) is simply the judgmental result.

For example, assume 'Fire exits are unblocked' is tokened in a scenario in which fire exits are blocked. The deviation between the indicated and actual states of affairs makes possible two distinct logical consequences: the representation is incorrect, or the actual state of affairs is. Whether the representation is factive or normative determines which logical consequence obtains. If factive, the representation is false, incorrect, wrong, inaccurate, and so on. The blocked fire exits are neither right nor wrong; they just are. If normative, the actual state of affairs is incorrect or wrong, not the representation. Whether factive or normative, the representation is, nonetheless, an indicative representation: it indicates a state of affairs that might deviate from the actual state of affairs. For the factive, the degree of deviation of the actual state from the indicated state determines the representational status (e.g., true/false, right/wrong, etc.) of the representation. For the normative, the degree of deviation of the actual state from the indicated state determines the normative status of the actual state of affairs. That's the indicative, so consider the imperative 'Lock the door.' That imperative might be given a non-normative or a normative construal. For example, let that imperative be the expression of my desire and assume that the door just will not lock despite my best efforts. It is my desire, not the door, which has been frustrated. As with the indicative, it is the representational status of the desire which is affected. (The representational status of an imperative is clearly not a truth-value but, instead, is commonly characterized as a satisfaction-value (i.e., satisfied or unsatisfied).) Consider that same imperative statement now to be a standing rule at Grandma's house. If I fail to lock the door, I am in the wrong, not the rule. Just as with the indicative, the logical consequences of the normative imperative fall on the actual state of affairs, not on the representation.

Both representational status and normative status indicate, then, the distance between the represented state of affairs and the actual state of affairs. The positive language of representational status (e.g., true, correct, right, accurate, etc.) and normative status (e.g., right, good, correct, etc.) indicate null distance/deviation, and the negative language of representational status (e.g., false, incorrect, wrong, inaccurate, etc.) and normative status (e.g., incorrect, wrong, bad, etc.) indicate some distance/deviation. Whether described as distance, deviation, or degree of correspondence or alignment, all permit that representational or normative status be continuous, not discrete. Just as we might allow continuous truth-values in fuzzy logic, so too we can allow things to be more or less good/bad.

Given normativity's paradigmatic role within evaluative and comparative judgment, the most straightforward path to a legitimately naturalized normativity would be to provide some naturalistic analogue to evaluative judgment. A naturalistic setup will clearly not issue in judgments of the form "x is good" or "x is bad." But, a naturalistic setup might produce something functionally analogous to the consequences of those judgments. In the cognitive case, evaluative judgments do not exist in a vacuum. Instead, the results of such a judgment provide *ceteris paribus* a reason to act. The judgment "x is false" is *ceteris paribus* a reason to correct the statement, and the judgment that some x is out of alignment with a standard is *ceteris paribus* a reason to correct the world, i.e., to bring it back into alignment with the standard. And, in the inverse, just as there is no reason *ceteris paribus* to correct a statement judged to be true, there is no reason *ceteris paribus* to correct the world when it is in line with the norm.[3]

To have a naturalistic analogue to the cognitive characterization of normativity, what we need is 1) a naturalistic setup with a null point from which items might deviate and 2) that setup be so structured to return deviant states to the null state. The second conjunct is the key to making the case that we have a naturalistic analogue to cognitively characterized normativity. That x is wrong should not only have causal consequences within the setup, but those consequences should be "corrective" in order to be analogous to the practical consequences of normative judgment.

A further criterion should be plausibly added as well. Tokens of, say, the same behavioral type can vary in their normative status depending on when, where, in what culture, environmental setting, etc. that they take place. For example, a man holding a door open for a woman can be correct and upstanding behavior, incorrect and offensive behavior, or neither depending on this or that culture at some historical point. This suggests that, whatever naturalistic setup realizes a normative setup, that setup should be something that can be only locally present, thereby allowing tokens of the same behavioral type in different contexts to vary in their normative status. That feature cannot be, then, some 'internal' fact about this or that token but must be some 'external' local fact which has causal effects on local tokens.

3. *Setting the Stage*

With criteria specific to normativity in hand, the most direct route to a model of normativity is to build a model system that would satisfy those criteria. In this section, I will construct an abstract schema of a particular subtype of stable system. In the subsequent section, I will suggest that normative language can be mapped onto that system in such a way as to meet the specific naturalistic criteria for normativity. But, since all of this is only abstract model construction, I will then try and show that this abstract system type is instantiated in two key types of causal setups for biological and cognitive inquiry.

Before characterizing the relevant subtype, it would be helpful, I think, to have on the table a rough sense of a stable system. For contrast, let's start with a (purely) unstable

[3] For a similar discussion of the naturalistic criteria for normativity, see Bauer (2009).

system. Assume that some system ('S') is in some initial state 'α', and assume that there is some set of possible and equally probable state transformations of S from α. For each possible state transformation, allow some further set of possible state transformations. So, S will over time track some trajectory through possibility space. To allow that S might remain in α, include α in the set of possible transformations from α. Sα at t1 to Sα at t2 would be a trivial transformation in state. The only reason to include it as a 'transformation' is that it is a categorically similar and equally probable event as S entering any other state at t2. α's inclusion among possible transformations is not, however, trivial. It allows the possibility that S never leaves α, and that in turn makes available the difference between looking to be and being stable. Or, the inclusion of α in the possible transformation set allows S to present the mere appearance of stability.The setup so far should be clearly an unstable setup as each possible transformation is equally probable. S's instability is increasingly likely to be observed over time, but S's instability is not constituted by time passing. It merely depends on the presence of a setup in which S is at least as likely to move away from α as stay in it.

The (purely) stable system just requires a small modification to the (purely) unstable system. On the possibility space available at Sα, overlay external constraints reducing to null the probability for any transformation excepting the trivial one to α. Assume those external constraints are present at every time point. S will track a single trajectory through possibility space as a consequence, namely staying in α. S is stable, because, given the external constraints, there is no probable transformation except the trivial one. The constraints are 'external', because the constraints are logically independent of the characterization of S or the set of possible transformations. The constraints are, then, something that can be added to or subtracted from (at least logically) the possibility space in which S exists. The logical independence of the constraints allows them to serve to explain the behavior of S: e.g., if it had not been for x-constraints, S would have likely moved along a different trajectory. The presence of the constraints allows there to be a legitimate distinction between apparently and really stable systems. Just as with the unstable system, the stability of S is determined by the constraints at a time, so, though stability is a diachronic notion, no time need pass for a system to be stable. (Quick modifications can generate other models of a stable system. For linear processes, simply chart a non-trivial trajectory through possibility space and place constraints at each point to funnel the process. For circular or 'autocatalytic processes', chart a non-trivial trajectory inclusive of a repeating string of transformations.)

The above models are pure models, because they treat stability as a discrete category. There are a number of ways to take those pure models and loosen things up to allow stability to be, as it is more commonly used, continuous or fuzzy. For example, rather than assuming that constraints limit transformations to just one from the set, we can open up the range of permitted transformations. E.g., the external constraints limit transformations to α and β but transformations within α and β are equally probable. The system is stable given the limited trajectory due to the external constraints, but, within the range of permitted transformations, the system is unstable – that is, there is no reason why S picks the particular trajectory through α and β transformations that it does. Instead of altering the range of permitted transformations, we might make external constraints temporally limited (e.g., at t6 the constraints will fail) or temporally decaying (e.g., at each t, there is an increasing probability that S will escape the constraints). Alternatively, we might allow that the constraints are permeable in that at each time there is some consistent probability of escape. S would be increasingly likely to be unstable over time, though the constraints are the same throughout.

While many further variations are possible, I want to concentrate on just one further and quite different variant. The constraints so far prevent S from entering particular states. Let's switch from these preventative constraints to counteractive constraints. Remove the external constraints and allow S to again move with equal probability in any direction within possibility space. For each possible state transformation from α, there is again some possible state transformations; call these 'the secondary sets'. E.g., from α, there is a set of transformations (α, β, γ), and from β, there is the secondary set of transformations (α, β, δ). Assume that each secondary set includes α, thereby allowing that S can return to α should it leave. Add now a factor to possibility space with the following effect: every possible transformation in a secondary set other than α is reduced to null probability. The result is that S in α is unstable: it is at least as likely to leave α as to remain. However, if S should leave α, this new factor, a counteractive constraint, sends S back to α. Though S is unstable in α, S will spend more time in α than in any other state. Or, a trajectory centered on α is stable for S. Further variations could be introduced by varying the weighting of the factor on secondary sets or varying which possibilities it affects, and further variations still can be introduced by coupling counteractive and preventative constraints.

4. *Normative mapping*

Taking those quite abstract models, normativity seems appropriately mapped to some but not others. Assume S is subject to only preventative constraints, and assume some trajectory through possibility space as a result of those constraints. Call those states on that trajectory 'favored' and those off the trajectory 'unfavored'. The favored/unfavored distinction reflects or is a way to describe the shape of the possibility space given the relevant external constraints. However, the distinction has little explanatory role outside of a short-hand way to characterize on what side of the constraint a possible transformation lies. There is no causal implication to being in an unfavored state. The state is simply out of the bounds of the constraints, and, should S make it into such a state, the constraints are now causally irrelevant to its subsequent trajectory through possibility space. Mapping normative language onto preventative constraints just fails to provide the causal explanatory role seemingly presupposed for normative status. However, if one assigns 'favored' and 'unfavored' in the same way for counteractive constraints, there seems to be a different result. There are now causal consequences to being unfavored: any S in an unfavored state is more likely than not to return to the favored state due to the presence of counteractive constraints, and those causal consequences are significant to explain why the long-term trajectory of S centers on, say, α. That provides at least a possibly naturalistic and causally significant basis for a favored/unfavored distinction. The favored are plausibly favored as the causal setup is such as to drive S into those states, and the unfavored are plausibly such as the causal setup is such as to drive S away from those states.

With at least a possibly naturalistic and causal distinction marked by the favored/unfavored distinction, that latter distinction might serve to map norms and normative status. Take the norm to be the favored state(s) for S, the characterization of which will be an indicative such as 'S is α'. (The imperative would clearly be wrong, since this would provide a prescriptive rule that the overall setup is striving to follow.) The truth of 'S is α' is constituted by the causal setup governing S. The putative norm 'S is α' does not imply that S is α at any given time. This is what we should expect of a norm, namely that such and such norm is the case does not imply that the world is line with that norm. Additionally, rather than saying 'S is α', we would be justified in saying 'S ought to be in α'. The latter is not a long-run predication, but it says what is 'owing' given the setup in which S exists. There is nothing, then, mysterious about there being a norm here: it is simply to recognize that a particular kind of causal setup is in place. The next step is to map normative status onto the favored/unfavored distinction. This seems to track the analogous everyday case. To be in the right or to be good is to be acting in some way in agreement with some norm, and it is to be acting in a way that should be encouraged or not discouraged. To be in the wrong or to be bad is to be acting in some way out of agreement with a norm, and it is be acting in a way that should be discouraged or not encouraged. It is to be acting in a way that makes one liable to correction.[4] Normative status mapped onto the state of an S with a counteractive constraint setup seems to be strongly analogous to those ordinary everyday thoughts. The causal setup is so structured to discourage deviant states by correcting the S to come in line with the norm and so structured not to discourage states in line with the norm. It does not seem, then, a significant conceptual stretch to map normative language onto such setups, because such setups provide consequences analogous to the consequences of practical judgment.

Let's connect this proposed normative mapping with the two specific criteria for naturalized normativity. A stable system with counteractive constraints provides a natural null point given the facts of the causal setup centering S along a trajectory in possibility space. That, in turn, provides a naturalistic basis to assign normative status to states within the possibility space, and those assignments carry very specific and normatively relevant causal implications: systems out of alignment with the norm are likely to be brought back into line, and those in alignment with the norm are likely to be kept in line. Additionally, the normative status of a tokened S is a property determined by the particular existing causal setup in which it resides. It is not a universal property of the S-type in some

[4] Phrases such as "should be encouraged" or "liable to correction" might seem to introduce a problem. They might be plausibly read as indicating that encouragement or correction is either appropriate or likely. Only the likelihood reading will be right for a naturalistic setup. However, if the right reading in a normative context is that correction is appropriate, then the plausibility of the proposed normative mapping would be specious; it would trade on an equivocation. It is the case in examining the range of human norms that there are often further standards governing the appropriateness of correction, standards governing how, when, and by what parties corrective action might be taken. That said, it is also the case in many mundane examples of human norms that the likelihood reading more than suffices. With human social norms, such as standards governing eye gaze, personal distance, or pronunciation, there might be further standards governing the propriety of corrective forms or which persons might engage in correction, but there need not be. To think that such norms exist in such a population, it suffices that people do respond to deviant and naïve behavior and that such responses generate conformity. Whether correction is or is not appropriate is often beside the point. Or, to say it in another way, whether corrective activity is appropriate is not a further question that the anthropologist or sociologist need ask and answer before describing a social norm.

state type – that is, there is no implication that all S's in some state have the same normative status or any normative status. As the localized normative status depends on the local presence of counteractive constraints, the causal implications of normative status are localized as well.

5. *Satisfying the General Naturalistic Criteria*

To satisfy the general naturalistic criteria, it should be plausible that some appreciable frequency of biological and cognitive systems are stabilized by counteractive constraints.

Many individual biological and cognitive systems do seem plausibly stabilized by way of counteractive constraints. Clear candidates are to be found among the varied homeostatic, regulatory, maintenance, and repair systems common to organisms. Such systems often operate not by preventing a change in state but by returning a system back to some target state. For example, sweat production in humans or panting by a canine do not prevent high core body temperature but, instead, are induced to lower core body temperature to a target range. The previous proposed normative mapping should plausibly apply to such cases. The state target maintained by these auxiliary counteracting systems would be the norm. A low or high body temperature would be incorrect, because the body at those temperatures is apt to be corrected so as to be brought into line with some temperature target for the body. (Notice the low or high body temperature is not bad or incorrect because it is deleterious to life but is only so due to the presence of counteracting stabilizing systems. That a low or high body temperature is deleterious can explain, instead, the presence of such counteracting systems.)

The commonality of counteracting systems in organisms is not a mere coincidental fact about the present state of terrestrial life. Any living system, even the simplest unicellular organism, requires a number of complex, extended (and usually cyclic) biochemical and physical events to take place in order to remain alive from one moment to the next. Environmental variability and internal degradation can not only knock those processes off path but will tend to occur at a rate faster than the lifespan of the organism. (Simple cases: considering the total temperature range that humans experience in their lifespan, without some counteracting measures (e.g., coats, hats, or shelter) most humans would quickly freeze or overheat; or, considering the high probability of a cut or injury due to moving about during a human's lifespan, humans would quickly die from blood loss without platelets.) Organisms incapable of counteracting environmental variability or internal degradation would simply fail. It should be expected, then, that any examined living system will have an array of counteracting systems. Identifying and explaining the role of such counteracting systems should be a key part of explaining how organisms make a living in the world. Here's the reason to note these large-scale biological hypotheses: given the centrality of maintaining target states for organismal operation, we should expect normative characterization to be widespread in the biological and cognitive sciences given the suggested mapping above.

Biological and cognitive explanations are frequently directed at populational phenomena, e.g., the frequency distribution of genders, alleles, or specialized behaviors within a population. The general aim with such population level phenomena is to account for the stability of a frequency distribution or, for evolutionary explanations, a stable populational trajectory through, say, morphological space.[5] The population is most commonly a reproductively-bound population but it need not be. In ecology, the considered population might be the whole set of individuals within an ecosystem or populations characterized by an ecological role (e.g., predators, scroungers, canopy-dwellers, canopy-creators, etc.); either way, the total members of these ecological populations do not stand in reproductive relations to one another.

Populational stability can reflect factors that might be characterized as preventative. For example, the fidelity of a reproductive process prevents the emergence of variant reproductive products. But, factors significant to populational stability are often better characterized as counteractive. The most familiar case is provided by natural selection. In a case of selectional stability, there is both a stable populational frequency distribution of, say, genes, morphology, or behavior as well as an overall reproductive tendency to drive the population away from that frequency distribution. The latter would reflect that reproductive output by the population at any given generation is likely to be at odds with the long term stable frequency distribution. So, in the absence of any further constraints, the frequency distribution within the population should be different than the stable frequency distribution. The selective process serves to correct or to counteract the tendency otherwise produced by reproduction. With environmental factors affecting the relative fitness of individual variation (whether individuals are genes, organisms, or groups), individual variation that would shift the frequency distribution over the long run is pushed out of the reproductive game. Selection is one significant factor that might be characterized at the populational level. But, there are non-selectional cases as well. For example, a setup similar to selection is one which reduces viability of non-heritable variation to null. The distinctive shape of the human heart, it has been suggested, is not a consequence of heritable factors but, instead, is the developmental result of a proto-pump responding to the shearing force of blood passing through it. (Hove et al. 2003) Simple changes in factors during development can produce very different shaped hearts from the same underlying genetics. However, variations on the shape of the heart significantly affect pumping capacity and tend to be lethal. A population will have a stable heart shape over time but not as a consequence of selection. (It cannot be selection, because individual variation is non-heritable.) Alternatively, in ecology, a particular predator-prey distribution within a locale might be stable over time, because fluctuations from that distribution increase lethality on one side or the other to bring things back into line. (Again, this cannot be selection, because the ecological population does not involve a single reproductive population with a heritable trait range.)

Stable populational phenomena can be (and often are) instances of systems stabilized by counteractive constraints. The proposed normative mapping should, therefore, apply to such populational phenomena. As it is a frequency distribution (e.g., a 1:1 male-to-female ratio) that is stabilized, that frequency distribution would be the relevant norm. Whole populations are, then, correct or incorrect relative to that norm, and only derivatively would normative status accrue to individuals. So, being male is neither right nor wrong,

[5] This is not to say that the interest here is in merely what is common or frequent in a population. The interest is in the stability of a frequency or frequency distribution, and that is inclusive of infrequent phenomena.

but being a member of an overly large male population would be wrong or, better, contributing to what is wrong about the present population. Though that might seem strange at first, ordinary everyday normative reasoning already provides logical space for this sort of thought. Burning your leaves in the fall is not on its own good or bad. If the overall population is small, there are no negative consequences on yourself or your fellows. If the population of leaf burners is sufficiently large, there can be real negative consequences to the health and welfare of yourself and your fellows. In this case, your action is only derivatively in the wrong, because you are contributing to a bad situation.

Given the prominence and significance of homeostatic individual systems as well as populational phenomena such as stabilizing selection, it should be reasonable to conclude that the proposed normative mapping captures at least a sizable section of the target practice. The proposal plausibly satisfies the first general naturalistic criterion. Let's turn to the second. That criterion requires that the naturalistic proposal accommodate the causal-explanatory practices of the target practice.

Normative status, on the above proposal, indicates some item's place within a broader causal setup structured to act upon it. Whether an item is correct or incorrect has differing causal implications within that setup. Negative normative status in particular has the implication that 'external' causal structures should move the respective item back into the target stable state. So, for example, a flatworm out in the open is in an incorrect state. The state is not incorrect, because flatworms out in the open are much more likely to be eaten. Instead, the light-sensing cells of the flatworm will affect a change in state until it is no longer out in the open. That is, there is a causal structure in place that acts to bring the flatworm back into a target state. That flatworms out in the open are likely to be eaten would be relevant to the explanation for why flatworms have a behavioral response system to move them out of the light.[6] So, the incorrectness of the flatworm has clear causal implications: that a flatworm is in an incorrect state explains why it is subsequently moved to cover (or a correct

[6] The line suggested here for the flatworm might strike one as counterintuitive. Intuitively, being out in the open is bad for the flatworm, because it is apt to be eaten. Further, it is counterintuitive to think, as implied by the above, that being out in the open would not be an incorrect state for the flatworm if the counteracting mechanisms were suspended or disabled. These intuitive judgments do reflect a common way of speaking about biological organisms, but this way of speaking is also best understood, I think, non-literally. Organismal processes or behaviors increase viability or fecundity or they do not. That is, processes or behaviors either increase or decrease the probability of ontogenic or phylogenic continuation. The "good for" or "bad for" generated by the above intuitions seems to be simply a way of saying that a process, say, increases or decreases, respectively, the probability of system continuation. After all, by those intuitions, a hidden flatworm is acting correctly insofar as it is keeping away from predators and, thereby, increasing its viability; a flatworm out in the open is acting incorrectly insofar as it is exposed to predation, thereby reducing its viability. But, then, this intuitive application of normative language assumes that physical system continuation (whether in the form of the organism or its lineage) is a universal norm of the biosphere. It is hard to see, however, why such a norm should be taken literally. To take it literally would be to assume that somewhere in nature there exists a system roughly analogous to evaluative judgment operating over the whole of biology within the universe. Or, in a weaker version, there is some such analogue operating over and on each distinct biological lineage. As far as I can tell, nothing in the present state of biology indicates that anything might serve that role. But, more importantly, it is unclear what explanatory advantage is gained by a literal construal. Organisms act in ways that either increase or decrease their viability or fecundity, and so, organisms and their lineages ei-

state). The explanatory target in this case is the alteration of some item to a target stable state, but the examples at the outset (e.g., a bird striking a window as a consequence of misperceiving, or a fawn captured by an eagle due to poor hiding behavior) were of a different character. In those examples, normative status does not explain the return to a target state but explains, instead, some further independent event.

Here's how I think this further type of explanation can be accommodated. Assume that S's being in α is stabilized by counteractive constraints as well as that some event E is explained by S being in α. Given that Sα explains E, the following should be reasonable implications: *ceteris paribus*, E would take place if a tokened S were in α, and E would not take place if a tokened S were not in α. Not-α would not be the absence of state but would be S in some further state, such as β,γ, δ, etc., because S cannot exist except in some state. Further, not-E would not be the absence of any event but would be some further event, such as F, G, H, etc., as even the absence of change is a describable event. By the proposed normative mapping, a tokened S in α possesses a positive normative status and vice versa. Or, things are as they ought to be when S is α, and things are not as they ought to be when S is not-α. So, E should take place, *ceteris paribus,* when things are as they ought to be, and not-E should take place, *ceteris paribus*, when things are not as they ought to be. Citing the positive/negative normative status of some tokened S can explain, then, some subsequent event, because that status merely refers to whether S is or is not in the stabilized state.

Take, for example, a fawn caught by an eagle due to poor hiding behavior or a robin striking a window as a consequence of misperceiving it as open space. Hiding behavior for pronghorn fawns consists in lying very still in the open prairie. This inherited behavior is exhibited at a high frequency among pronghorns and is selectively stabilized by heavy predation. As the stabilized behavior, a still fawn would be hiding well, and a jittery fawn would be hiding poorly. A poorly hiding fawn should not be, *ceteris paribus*, passed over by a predator. The capture of a tokened fawn by an eagle can be explained then by its poor hiding behavior. (Byers 2002) Or, to take another example, robins make use of sight to avoid obstacles in flight. Representational errors concerning such obstacles are potentially catastrophic, so it is not too much of a speculation to think that the representational system is stabilized to represent obstacles where they are there and not otherwise. So, a robin with a veridical representation of an open space is representing well or correctly insofar as there is a system reinforcing the tokening of representations like that in those conditions. A correct representation can serve, then, to explain the robin's avoidance of obstacles, and similarly a false or incorrect representation can explain its failure to avoid an obstacle.

6. Capturing too much

The model for mapping normative language onto systems stabilized by counteractive constraints might seem to be just too broad a proposal. Abiological physical systems might also

ther persist or they don't. Without invoking some classical and mystical form of teleology at this point, it is hard to know what explanatory function such a continuation norm might serve.

instantiate such stable setups. Yet, it would actually be quite odd if they could not: not only is the biological/abiological divide continuous but, as pressed by Ganti (2003) and Kauffman (2000), the essential feature of living systems as replicating autocatalytic collectives is not unique to life. Independent of the naturalizing account given, what should be expected, then, is to find norms outside of the paradigmatic boundaries of the biological. That said, normative language is not generally used for explanatory purposes in the physical and chemical sciences. If so, why is it that physicists and chemists do not regularly mobilize normative characterization if the model is on the right track?

That question is a psycho-sociological question. The question is just why have biologists gone one way and physicists another. Here is, I think, a plausible hypothesis. System stability is at the core of biological and cognitive inquiry. Just how an organism can exist from one moment to the next and, in so doing, counter environmental variability and internal degradation are ground level concerns. As a consequence, the normative mapping proposed here would be broadly applicable at the key focal point of inquiry. In contrast, stable systems do not comprise the core subject matter of physics and chemistry. So, for purposes of uniformity of explanation, one would not expect normative language, even when appropriate, to be mobilized in those fields.

7. Concluding Remarks

The widespread use of normative language in biological and cognitive explanations presents a prima facie problem: natural norms and normative status do not seem to fit into the broader picture of the natural world. To dissolve that problem, the suggestion offered here was that, first, we map normative language onto systems stabilized by counteractive constraints, identifying the norm with the stable state and normative status with distance from the stable state. Doing so allows for the satisfaction of the two criteria specific to normativity by providing a localized feature determinative of a natural null point and providing the appropriate causal implications to normative status. Second, relevant real systems seem to provide plausible instances of that abstractly specified stable system type. In fact, such real systems seem to form the core target of biological and cognitive inquiry. By adopting the suggested model, the apparent failure of fit would be merely apparent.

As noted at the outset, much of the interest in natural normativity within the literature has been restricted to norms of functional performance. Normativity for the functional theorist arises as an issue, because plausibly some items should be functionally classified even when they do not or cannot perform the relevant functional effect. So, the functional theorist needs to provide some basis for functional classification independent of actual tokened performance. Once provided, items can be normatively characterized by reference to the function that they ought to perform as members of a functional category. But, by methodologically isolating the normative question from the functional one, the paradigmatic example of normativity shifted from the success or failings of some device to the broader case of normative judgment and its consequences. (It is broader, because the evaluation of a device is just one example of normative judgment.) Using normative judgment and its consequences as the paradigmatic case required that the relevant naturalistic analogue have something like a "corrective mechanism". That requirement goes beyond what is necessary for a successful functional account. That the requirements for a successful functional account and for a successful normative account can conceptually come apart suggests four different points.

1. The success of a functional account does not hinge on providing literal norms of performance. Some supplied naturalistic basis can plausibly serve the needs of a functional theory without meeting the criteria for normativity. Given the adoption of some functional theory, we can have grounds to take its described functions as real natural phenomena without yet having a reason to take its described norms of functional performance as anything more than a heuristic device.
2. Whether some particular norm of functional performance is a norm in fact is contingent. Whether the appropriate naturalistic setup is in place for a literal natural norm is a fact conceptually independent from the naturalistic basis for functional ascription. Whether the described norm of functional performance is a norm in fact is, therefore, an independent empirical hypothesis.
3. It may turn out that all natural norms are in fact only norms of functional performance. That would not imply that all norms of functional performance are themselves literally norms given the conceptual divide in functional and normative requirements.
4. It is possible that not all natural norms are norms of functional performance. Given the shift to the broader category of normative judgment and its consequences, the specification of a naturalistic norm did not need to make any reference to function. That permits accommodating the very real possibility that some human as well as animal traditions might exist as norms without serving any function.

The model offered here should not be seen as conflicting with or even favoring any particular functional account. Instead, it just provides an independent means to assess the extent to which we should take some norm of functional performance to be literally a norm.

REFERENCES

Bauer, M. 2009. Normativity without artifice. *Philosophical Studies* 144: 239-59.

Byers, J. 2002. The Ungulate Mind. In *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, eds. M. Bekoff, C. Allen and G. Burghardt, 35-40. Cambridge: The MIT Press.

Davies, P.S. 2001. *Norms of Nature: Naturalism and the Nature of Functions*. Cambridge: The MIT Press.

Fragaszy, D. & S. Perry (eds.). 2003. *The Biology of Traditions: Models and Evidence*. Cambridge: Cambridge University Press.

Galef, B. & K. Laland. 2005. Social Learning in Animals: Empirical Studies and Theoretical Models. *BioScience* 55: 489-99.

Ganti, T. 2003. *The Principles of Life*. New York: Oxford University Press.

Hove, J., R. Koster, A. Forouhar, G. Acevedo-Bolton, S. Fraser & M. Gharib. 2003. Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. *Nature* 6919: 172-7.

Janik, V. & P. Slater. 2003. Traditions in Mammalian and Avian Vocal Communication. In *The Biology of Traditions: Models and Evidence*, eds. D. Fragaszy and S. Perry, 213-35. Cambridge: Cambridge University Press.

Kauffman, S. 2000. *Investigations*. New York: Oxford University Press.

Nakamura, M. 2002. Grooming-hand-clasp in Mahala M Group Chimpanzees: Implications for Culture in Social Behaviors. In *Behavioral Diversity in Chimpanzees and Bonobos*, eds. C. Boesch, G. Hohmann, and L. Marchant, Cambridge: Cambridge University Press.

Shettleworth, S. 1998. *Cognition, Evolution, and Behavior*. Oxford: Oxford University Press.

Whiten, A., J. Goodall, W. McGrew, T. Nishida, V. Reynolds, Y. Sugiyama, C. Tutin, R. Wrangham & C. Boesch. 2001. Charting Cultural Variations in Chimpanzees. *Behaviour* 138: 1481-1516.

Mark Bauer is a Research Assistant Professor at the University of Colorado Denver who works primarily on issues of biological and cognitive explanation.

Address: Philosophy Department, The University of Colorado Denver, 955 Lawrence St, Plaza M108, Denver, Colorado, USA 80217. Email: mark.bauer@ucdenver.edu