A DEFENSE OF THE RIGHTS OF CONSCIENCE

IN BUTLER'S ETHICS

Michael W. Martin

December 3, 1976
Abstract:

In "Nature and Conscience in Butler's Ethics," Nicholas Sturgeon argues that Butler's account of the role of conscience in morality is fundamentally incoherent. Butler's emphasis upon conscience as the most superior principle rendering acts natural or unnatural is inconsistent with his tacit commitment to the "Naturalistic Thesis" that conscience always uses naturalness and unnaturalness as grounds upon which it bases its approvals and disapprovals. I argue that Butler is not committed to the Naturalistic Thesis, and hence his views are saved from incoherence. This Thesis is not entailed, as Sturgeon claims, by two of Butler's central doctrines, and there are reasonable interpretations of the passages Sturgeon cites that do not commit Butler to the Thesis. Butler's view is that the logically primary perception-approvals of acts as virtuous and perception-disapprovals of acts as vicious by themselves can render acts natural and unnatural, respectively, without the need for conscience to rely upon some other superior principle to first determine the naturalness or unnaturalness of acts.
A Defense of the Rights of Conscience

In Butler's Ethics

In "Nature and Conscience in Butler's Ethics," Nicholas Sturgeon argues that Butler's account of the role of conscience in morality is fundamentally incoherent. On the one hand, at the heart of Butler's ethics is the doctrine of the supremacy of conscience: an act approved by conscience is thereby rendered natural, and an act disapproved by conscience is thereby rendered unnatural. On the other hand, it can be shown, Sturgeon argues, that Butler is tacitly committed to the "Naturalistic Thesis": conscience always operates by basing its approvals and disapprovals on their naturalness or unnaturalness. Yet the Naturalistic Thesis renders the Supremacy Doctrine entirely superfluous. Hence Butler's emphasis on the Supremacy Doctrine is inconsistent with his commitment to the Naturalistic Thesis, and it can reasonably be argued that the former doctrine should be abandoned altogether.

I shall argue, however, that Sturgeon fails to establish Butler's commitment to the Naturalistic Thesis. On Butler's view, conscience operates by directly intuiting the virtue or vice of actions. These intuitions or perceptions are inseparable from, and provide the grounds for, conscience's approvals and disapprovals (D.1). It follows that conscience is capable of operating autonomously, without needing to rely upon some other superior principle to first render acts natural or unnatural. Contrary to the Naturalistic Thesis, naturalness and unnaturalness are not, in general, the grounds used by conscience in approving and disapproving of actions.

In the first section, I present a background discussion of the Supremacy Doctrine—i.e., the doctrine of conscience's supremacy—and of the relation between this doctrine and the Naturalistic Thesis. Each of the subsequent three sections is devoted to an examination of

---

one of Sturgeon's three arguments for ascribing the Naturalistic Thesis to Butler.

(1) Butler formulated the Supremacy Doctrine as the backbone of his explanation of "what is meant by the nature of man, when it is said that virtue consists in following, and vice in deviating from it" (P.13). Man's nature, in the strict sense, cannot be defined merely by specifying the presence and relative strengths of his desires, affections, interests, or in general his "principles of action." For unlike the nature of mere brutes, the nature or "constitution" of human beings must be specified in terms of the "superiority" of certain principles and their corresponding faculties. A person follows his nature when he acts in conformity with his superior principles, and deviates from his nature when he fails to act in conformity with the most superior principle operating in a given situation. Butler's only explicit ranking among principles of action places prudence, or "reasonable and cool self-love," superior to the passions, and conscience superior to all other principles. Accordingly, it is unnatural to act on an impulse or passion which violates what prudence indicates as the means to obtaining one's happiness (2.10-11). Again, failing to follow the dictates of conscience for any reason at all constitutes unnatural action, since conscience is supreme among principles of action (2.14).

Butler has two ways of defining the notion of superiority. Sturgeon emphasizes only one of these definitions, which he presents as follows (319): "by the superiority of one principle of action to another 'in nature and in kind' (2.11), as he [i.e., Butler] explains carefully, he means that unique relation between them distinct from any measure of their relative strength, which renders action in accord with the inferior principle and contrary to the superior one disproportionate to the nature of the agent; or, as he prefers to say, 'in the strictest and most proper sense unnatural' (2.10)." Now there is a problem with following Sturgeon in viewing this as Butler's main definition of "superiority." Unnatural acts are simply those which violate human nature, where "human nature" is defined in terms of principles of action standing in relationships of superiority-inferiority. If, in turn, the relation of superiority-inferiority is defined with the term "unnatural," Butler is saddled with a set of uninformative circular definitions.

Butler does, however, provide a non-circular definition of superiority in the course of describing the absolute superiority of conscience. This definition, alluded to and wrongly dismissed by Sturgeon as mere metaphor (319), says that the superiority of principles has to do with their "authority" (P.14). In one place Butler specifies that a principle is authoritative to the extent that it "ought to have" influence upon action (1.8). Following this out, the most superior principle with respect to a given situation is that principle which ought to be acted upon in that situation. Just as in
a political context we distinguish between merely having the power to
do something and having a legal right to do it, so in morality we
distinguish between the motive or principle which is strongest and the
one which has the right to take precedence (2.14). Using this defini-
tion, we can say that where only the principles of self-love and
particular passions are involved, self-love (or the requirements of
prudence) ought to be followed, and the Supremacy Doctrine requires
that conscience ought to be obeyed whenever it makes a judgement.

I will understand the concept of superiority in this second way.
If it is objected that on this interpretation Butler is packing moral
connotations into such expressions as "natural" and "human nature,"
it should be pointed out that this is exactly what he intended to do.
For he claims to be explicating the concept of human nature according
to which St. Paul could correctly say that men are by nature a moral
law unto themselves (2.8), and always, he is interested in what he
calls "the moral constitution of man" (3.2 n.).

Sturgeon defines the "Full Naturalistic Thesis" as the assertion
that "conscience never favors or opposes any action, except on grounds
which include its naturalness or unnaturalness" (328). The crucial
part of this thesis is also expressed in what Sturgeon calls simply
the "Naturalistic Thesis": "conscience always bases its approval or
disapproval of actions, in some essential way, on their naturalness or
unnaturalness" (323). I interpret this to be the following two-fold
thesis: (a) conscience uses criteria or reasons ("grounds") as the
basis upon which it approves or disapproves of actions, and (b) these
criteria always involve the naturalness or unnaturalness of the action.
It is with (b) alone that I shall later take issue.

Although the difference between the Naturalistic Thesis and the Full
Naturalistic Thesis will not play an important role in what follows, it
bears remarking that the distinction rests on a controversial inter-
pretation of Butler's concept of natural acts. Sturgeon asserts that
for Butler "it must be possible for conscience to favor an action with-
out approving of it" (327-8). Conscience favors acts solely on the
basis of their being natural, but it approves of acts on the basis of
their being both natural and the result of a good motive and intention.
Sturgeon cites no textual evidence, and I have been unable to find any,
for drawing a distinction between (a) acts which are natural and (b)
acts which are both natural and spring from a good motive and intention.
He correctly notes that virtuous acts must, for Butler, spring from
good motives and intentions (325-6 n.). But this is consistent with
holding, as I believe Butler does, that the naturalness of an act
includes its springing from what conscience judges to be a good motive
and intention. At the very least, an act is fully natural only if it
accords with all the dictates of conscience, where some of these
dictates impose moral restrictions on the motives and intentions of
actions. Sturgeon's interpretation, as he points out, forces us to
qualify one of Butler's central doctrines. On his interpretation, it
would be false that virtue consists simply in following nature: virtue would consist in following nature plus acting on certain motives and intentions. On my reading, by contrast, Butler can be taken au pied de la lettre.

The incompatibility of the Naturalistic Thesis, in either version, with the doctrine of conscience's supremacy can be seen immediately. If the Supremacy Doctrine is to have any importance in Butler's ethics, the approvals and disapprovals made by conscience must be capable by themselves of rendering acts natural and unnatural, in independence of the operations of other superior faculties and principles. But it is just this possibility that is ruled out by the Naturalistic Thesis. According to that thesis, before conscience can ever approve or disapprove of an act, conscience must first discern the act's naturalness or unnaturality. In order for conscience to be able to make this determination, it must rely upon some other superior principle to render the act natural or unnatural. For it would be nonsense to hold that conscience could approve or disapprove of acts on the grounds that the acts conformed or failed to conform to those self-same approvals and disapprovals. My grounds for disapproving of misconduct cannot be my disapproval of misconduct—to speak in this way would at best be a rhetorical way of saying that I have no grounds upon which I base my disapprovals.

It would require far too much space to retrace all of the steps of Sturgeon's detailed argument showing how the Naturalistic Thesis renders the Supremacy Doctrine superfluous. The next paragraph contains only a brief summary of some of his main points, staying close to his wording, and restricting discussion to approvals and disapprovals of a person's own current acts by his own conscience.

According to the Naturalistic Thesis, my conscience favors or opposes my present actions on grounds which always include their naturalness or unnaturality. According to the definition of naturalness and unnaturality, my action is "rendered natural or unnatural only by the current favor or opposition of a superior principle" operating in me (330). Therefore, before my conscience "can come to favor or oppose" any of my current actions, "the action must already be favored or opposed by some other superior principle" in me. Hence, conscience "charts no independent course of its own," rendering inappropriate Butler's characterization of it as "the guide of life" (345). In favoring and disfavoring actions, my conscience must always first defer to the favor or disfavor of the highest superior principle in my nature, other than conscience, which either favors or opposes my act (346). But then "it makes no difference whatever to the naturalness of any action [of mine] whether or not conscience is superior to any other principle of action" (347). For, the ranking of conscience would make a difference only if either (a) my conscience ever favored or opposed one of my current acts which were not already favored or opposed by any other superior principle in me, or (b) my conscience ever conflicted with the highest principle,
other than conscience, to favor or oppose my act. But the Naturalistic Thesis rules out both of these possibilities. The favor or disfavor of conscience always presupposes that at least one other principle (or faculty) favors or disfavors the act under consideration, and it always relies on the favor or disfavor of the highest such principle.

(11) I turn now to Sturgeon's three arguments for ascribing the Naturalistic Thesis to Butler. Two of the arguments are that the thesis is presupposed in two passages, one in the Dissertation and one in the Second Sermon. The third argument is that the Naturalistic Thesis is entailed by two of Butler's doctrines which are at least as important to his thinking as the Supremacy Doctrine. I will begin with the latter argument, since it is by far the most important.

The Naturalistic Thesis, Sturgeon argues, "is a consequence, on any reasonable understanding of them, of the following two propositions, both of which he [Butler] maintains: (1) that virtue consists in following nature, vice in deviating from it, and (2) that whenever conscience approves or disapproves of an action, it does so on grounds of the virtue or vice of the action" (324-5). That is, according to Sturgeon, Butler's espousal of (1) and (2) commits him to the Naturalistic Thesis: (3) Whenever conscience approves or disapproves of an action, it does so on grounds which include its naturalness (accordance with nature) or unnaturalness (deviation from nature). I shall now argue that this is not so: there is a perfectly reasonable understanding of (1) and (2) according to which they do not entail (3).

Consider the argument in the following form:
(1) Virtue always consists at least partly in following nature, and vice always consists in deviating from nature.
(2) Whenever conscience approves or disapproves of an action, it does so on grounds that the action is virtuous or vicious (i.e., evil).
(3) Therefore, whenever conscience approves or disapproves of an action, it does so at least partly on the grounds that the action is natural or unnatural.

This is roughly the final form in which Sturgeon presents the argument, with one main difference: in the first premise Sturgeon writes "naturalness" and "unnaturality" for where I have written, using Butler's own wording, "following nature" and "deviating from nature." This difference is noteworthy. When Sturgeon says that "vice consists in unnaturalness" (327), and that "vice and unnaturalness are one" (327 n.), his language is misleading. It suggests that the words "vice" and "unnaturalness" are rough synonyms referring to the same property, and Butler nowhere makes such a claim. Butler defines unnatural action (i.e., action which deviates from one's nature) as action which violates a superior principle, but he does not define vicious action in this way. In fact, beyond an enumeration of the types of vicious and virtuous acts, he pretty much treats "vice" and "virtue" as undefined terms.
I suggest that the only sense in which Butler holds that "vice consists in deviating from nature" is that vicious actions are also unnatural actions. More fully, "vice and "unnaturalness" are co-extensive terms applying to exactly the same things, but they are not two words referring to one and the same property. Now, as Sturgeon points out in an important footnote, the argument from (1) and (2) to the Naturalistic Thesis "may seem doubtful" because it involves substitution of identity in a mentalistic context (327 n.). It might turn out, we can add, that the occurrences within the mentalistic context in Premise Two of the predicates "virtuous" and "vicious" are not replaceable salva veritate by "natural" (or "in accord with nature") and "unnatural" (or "in violation of nature"). I want to show that this is in fact the case before discussing the remainder of the footnote just mentioned.

Sturgeon established that if the Naturalistic Thesis is true, then there can be no actions which conscience alone renders unnatural. Applying transposition to this conclusion, it follows that if there were any actions which conscience alone rendered unnatural, then the Naturalistic Thesis would be false. Furthermore, if (1) and (2) can be shown to be consistent with the possible existence of such actions, then in some possible world (1) and (2) will be true while (3) is false, and the argument will be shown invalid. But it is obvious that (1) and (2) leave open the possibility that some acts are unnatural only because they are approved or disapproved by conscience. On the supposition that there were such acts, (1) could be true, since the acts would be both vicious and unnatural, and (2) could also be true, since conscience could disapprove of the acts for their evilness. However, (3) would be false, since it would be impossible that conscience could use unnaturalness as the ground for its disapproval.

For, to make things explicit, what superior principle could render the imagined act unnatural in a way that conscience would be able to employ the unnaturalness of the act as the ground for its disapproval? Certainly no superior principle other than conscience, since ex hypothesi the act imagined is rendered unnatural only by conscience. With equal certainty, conscience could not itself be the principle: it is absurd to suppose that conscience could disapprove of an act on the grounds that the act is thus disapproved by conscience.

We may add that Butler explicitly allows the conceivability of the imagined case, where only conscience renders a given act unnatural or natural. In the Preface, he insists it is logically possible that conscience and self-love could come into conflict, in which case it would be conscience alone which determined the naturalness or unnaturalness of an act (P.26). To be sure, later on he claims that moral duty, as discerned by an honest conscience, and self-interest, as discerned by a reliable faculty of self-love, always coincide in the long run (3.8-9). But he explicitly rests this claim upon a supplemental religious hypothesis that goes well beyond anything expressed in (1) and (2): God will reward the moral person and punish the person who acts immorally.
for the sake of either passion or secular self-interest. Moreover, from this same passage it is clear that the voices of conscience and prudence speak in perfect unison only for those people who understand their "true happiness" and take God's Final Judgment into account in their reflections (cf. 340).

I have suggested that Sturgeon's reasoning can be turned on its head, and there may linger a suspicion that I have simply begged the question. After all, my argument relies upon the possibility that (1) and (2) could be true at the same time there were possible actions which were rendered natural or unnatural solely by conscience's approvals or disapprovals, and Sturgeon argues at length that there could be no such actions (347). Recall, however, that his argument depends entirely on the premise that Butler is committed to the Naturalistic Thesis (330, 344), and it is just this which is in question in the present context. Obviously the counter-example would be ruled out if Butler were committed to the Naturalistic Thesis—but that thesis is the conclusion, not a premise of the argument being examined.

One could easily rule out the counter-example in another way by adding a new premise that conscience and self-love always make exactly the same pronouncements. I do not think this is Butler's considered position, but some passages do suggest the view (cf. 338). Even with this extra premise added, however, the argument would remain invalid. The new set of premises still leaves open the following two-fold possibility: (A) conscience could operate by directly discerning the objective properties of virtue and vice of actions, and simultaneously approving of actions as virtuous or disapproving of them as vicious; (B) these approvals and disapprovals, only as described fully by including a specification of their grounds, render acts natural or unnatural. Part B insures that Premise One is true: each act that conscience approves as virtuous will thereby be rendered natural, and each act that conscience disapproves as vicious will thereby be rendered unnatural. Part A makes Premise Two true: conscience would always use virtue and vice as the grounds for its approvals and disapprovals. But Part A also renders the conclusion false. According to A, conscience operates independently of other superior principles and faculties, having no need to consult them first in order to obtain the specific grounds for making its approvals and disapprovals. The Naturalistic Thesis denies this. For it requires, in Sturgeon's words, that before a person's "conscience can come to favor or oppose any of his current actions, . . . the action must already be favored or opposed by some other superior principle in his nature" (330).

This imagined mode of operation of conscience which shows even the modified argument to be invalid seems to be precisely the one Butler describes as being fact. First, Butler emphasizes that conscience directly perceives virtue and vice, analogously to how "speculative reason directly and naturally judges of speculative truth and falsehood" (D.1 n.). Inseparably linked with these perceptions is an approval of some acts
"under the peculiar view of their being virtuous and of good desert" and a disapproval of others "as vicious and of ill desert" (D.1; cf. 2.8). Clearly this is a description of conscience as functioning autonomously in making moral perceptions and concurrently issuing moral approvals and disapprovals. Second, it is very plausible to interpret the Supremacy Doctrine as requiring that conscience's approvals and disapprovals, only as fully described by including a specification of their grounds, render actions natural or unnatural.

Sturgeon does not provide a full explanation of why he thinks the Naturalistic Thesis is entailed by (1) and (2). In calling the Naturalistic Thesis a "straightforward consequence" of these premises (324), he seems to be assuming that the inference is obviously valid. He does argue, however, in the footnote referred to earlier, that Butler relies on inferences just like this one, and that (presumably indirectly) this commits him to the inference to the Naturalistic Thesis (327 n.). The sole example he offers is the following:

(1') One's own happiness "consists in the gratification of certain affections, appetites, and passions" (11.16).
(2') Self-love is a desire for one's own happiness.
(3') Therefore, self-love is a desire for the gratification of these affections, appetites and passions.

It is true that this argument is similar to the first one in that it involves substitution of identity in a mentalistic context. But there the similarity stops, and I cannot see how commitment to this second, presumably valid argument, commits Butler to the earlier, invalid argument. The second argument concerns desires and their objects, and amounts to substituting one description of the object of a desire ("the gratification of certain affections, appetites, and passions") for another, coextensive description of the object ("one's own happiness"). In this type of argument there are none of the complications involved that arose in connection with reason-based approvals and disapprovals. In particular, there is nothing analogous to how it is possible for conscience's disapproval of an act on the grounds that it is vicious to render the act unnatural—the possibility that insured that the words "vicious" and "unnatural" cannot merely refer to one and the same property of acts.

(iii) Sturgeon's second reason for ascribing the Naturalistic Thesis to Butler is that he "clearly relies" on it in a passage in the Dissertation on Virtue: 
"[Butler asserts] that 'our perception of vice and ill-desert,' which he refuses to separate from our disapproval of them (D.1), 'arises from, and is a result of, a comparison of actions with the nature and capacities of the agent,' in which the action is determined to be 'incongruous,' "unsuitable," 'disproportionate,' or 'unfit' (D.5)" (324). Standing by themselves these excerpts might look like a general statement of the Naturalistic Thesis. They are
taken, however, from a context which dictates a different interpretation.

In Section D.5, Butler is arguing that our moral perceptions and approvals of "creatures" depend upon whether those creatures have the moral conscience of a mature and sane person. We judge actions to be vicious in mature and sane adults which we do not judge vicious "in creatures of other natures and capacities, as brutes." Moreover, we all share "a different sense of harm done by an idiot, madman or child, and by one of mature and common understanding." It follows that "this difference must arise from somewhat discerned in the nature or capacities of one, which renders the action vicious; and the want of which, in the other, renders the same action innocent or less vicious: and this plainly supposes a comparison, whether reflected upon or not, between the action and capacities of the agent, previous to our determining an act to be vicious" (D.5). That is, our evaluations of a creature's actions presuppose judgments about his moral and psychological capacities. These general capacities must first be ascertained in order to insure the "proper application of the epithets, incongruous, unsuitable, disproportionate, unfit, to actions which our moral faculty determines to be vicious" (D.5).

Nowhere in all this is it suggested that in evaluating my own actions my conscience operates by comparing my actions to the dictates of another superior faculty in me before approving or disapproving of them. Nor is it even suggested that my conscience cannot disapprove of another person's actions until after I compare his actions to the dictates of his superior principles and faculties. Here we must be careful to distinguish grounds or reasons for morally approving and disapproving of actions, from general presuppositions rendering such approvals and disapprovals intelligible or appropriate. At least in the vast majority of cases, it is conceptually inappropriate to hold animals morally responsible for their behavior. Moreover, children and insane persons are not as fully responsible, and in some cases not responsible at all, for their actions in the way a normal mature adult is. But from these general presuppositions of moral approvals, it does not follow that each approval and disapproval by one person's conscience of another person's actions must be made on the basis of comparing the actions to the highest principle of the other person which either approves or disapproves of his act. Butler's remarks in D.5 concern the general presuppositions of moral appraisals, not the specific criteria used by conscience in appraising acts.

(iv) The final evidence for ascribing the Naturalistic Thesis to Butler is what Sturgeon presents as the second of Butler's two arguments for the Supremacy Doctrine. The first of these arguments is that the Supremacy Doctrine is an analytic or conceptual truth. We think of conscience, Butler writes, "as being superior; as from its very nature manifestly claiming superiority over all others [i.e., principles of action]; insomuch that you cannot form a notion of this faculty, conscience,
without taking in judgment, direction, superintendency. This is a constituent part of the idea, that is, of the faculty itself: and to preside and govern, from the very economy and constitution of man, belongs to it" (2.14; cf. 2.12). The second argument, which appears shortly after the first, begins with the proposal that we "turn this whole matter another way, and suppose there was no such thing at all as this natural supremacy of conscience; that there was no distinction to be made between one inward principle and another, but only that of strength; and see what would be the consequence" (2.16). The immediate consequence is that in the strict sense in which acts are natural or unnatural (i.e., just in case they conform or fail to conform to a superior principle) there would be no natural or unnatural acts. Put another way, human nature would be defined wholly in terms of the presence of certain principles of action together with a specification of their relative strengths. Using this latter, weaker, sense of "human nature," all acts would be equally natural, since presumably they would all result from the agent's strongest principle operating at the time he acts. If we imagine a man acting on his strongest motive in committing parricide, we must confess that his act corresponds to his nature as much as an act of filial duty on another occasion that also flowed from his strongest motive.

So far, so good. But Butler concludes his argument with the following sentence: "If there be no difference between inward principles, but only that of strength; we can make no distinction between these two actions, i.e., parricide and filial duty, considered as the actions of such a creature; but in our coolest hours must approve or disapprove them equally: than which nothing can be reduced to a greater absurdity" (2.17). Sturgeon argues that there is a problem with understanding this sentence, a problem which can only be resolved by ascribing the Naturalistic Thesis to Butler. To begin with, according to Butler's remarks elsewhere, "distinguishing between actions in a cool hour, and approving or disapproving them, is a function of conscience (1.8). Stated generally, therefore, Butler's crowning conclusion is this: that if no principle of human action has a natural superiority to any other, conscience must approve or disapprove equally of any two actions whatever" (322-3). The problem arises as to why Butler should think this conclusion follows: "If no principle is superior to any other, conscience will be on a par with any passion or appetite; but why should its lack of superiority prevent it, in Butler's view, from distinguishing actions, approving one and disapproving another?" (323). Sturgeon believes the only reasonable explanation is that Butler is presupposing the truth of the Naturalistic Thesis. According to that Thesis, conscience always bases its approvals on the naturalness and unnaturalness of actions, as defined in terms of conformity or non-conformity to superior principles. Since in the imagined circumstances, "naturalness" and "unnaturalness" have no application in their strict senses, conscience is prevented from making discriminating approvals and disapprovals by using naturalness and unnaturalness as criteria.

I find Sturgeon's interpretation of the passage both subtle and, so
long as the passage is viewed by itself, highly plausible. Note first, however, that even if the interpretation is correct there would not be sufficient grounds for viewing Butler as deeply committed to the Naturalistic Thesis. Instead, it should be concluded that Butler became momentarily confused when he wrote the last sentence of the passage: he unwittingly presupposed a view which is incompatible with the numerous passages in which he emphasizes the Supremacy Doctrine. Second, we should consider sympathetically any alternative interpretation of the entire passage which does not convict Butler of this inconsistency. I contend that there is a second interpretation which is both plausible by itself and deserving of preference for rendering the passage consistent with the rest of Butler's account.

My interpretation proceeds by rejecting Sturgeon's innocent-looking initial premise, and then refusing to view the argument set forth in Section 2.17 as independent of the first argument for the Supremacy Doctrine. Sturgeon begins by assuming that since in other passages Butler describes conscience as approving or disapproving actions "in a cool hour," we can take for granted that it is conscience which performs this same function under the circumstances described in the present context. This is an inference on his part, since nowhere in Section 2.17 does Butler explicitly mention conscience. I submit that Butler does not have conscience in mind when he says in the concluding sentence that "In our coolest hours we must approve or disapprove" equally of the acts of filial duty and parricide. The supposition of his argument can be interpreted as ruling out precisely this possibility. Bear in mind that immediately before Section 2.17 Butler had made it clear that it is "a constituent part of the idea" of conscience that it is the supreme principle in man. We must take this seriously! Against this background, when Butler lays down the supposition that there are no superior principles in man, he has, inter alia, supposed that man has no conscience. Stripped of its supremacy conscience no longer exists, since its supremacy is one of its defining properties.

According to this interpretation, Butler is not seeking in Section 2.17 to provide a rigorous proof of the Supremacy Doctrine. Instead, he is asking us to conduct the following informal thought experiment. Imagine that we were transformed into human beings just like ourselves except for one small thing—we would lack superior principles. Since it has already been asserted that conscience, by definition, is a supreme principle, we are being asked to imagine that we lacked consciences or any other principles having conscience-like authority. The absurd result is that we could not make moral discriminations concerning the actions of such creatures, whether they be ourselves or others. This is not because our consciences would lack the relevant sorts of grounds to use

2Sturgeon's interpretation is adumbrated by W.R. Matthews in his prefatory note to Section 2.17: "The view that there is no distinction between principles of human nature except their relative strength takes away the rational ground of moral judgment, and therefore leads to absurdity" (Matthews, p. 47).
in making approvals and disapprovals, but because we would have no consciences at all! Again, the approvals and disapprovals we made in our cool hours would not be the expression of our consciences for the simple reason that we would lack consciences. In effect, we would no longer be human beings, but only conscienceless brutes, a fact which explains why Butler switches to the word "creature" at the end of the passage.

Using my interpretation, the argument in Section 2.17 can be viewed as an extension of the argument Butler gave in the previous sermon for holding that "it cannot possibly be denied, that there is this principle of reflection or conscience in human nature" (1.8). There Butler had his reader first imagine that a man performs two actions: an act of helping an innocent person in great distress, and an act of "greatest mischief" to a former friend who had not offended him in any way. Suppose further that the man subsequently reflected "coolly" and objectively upon his two actions: "to assert that any common man would be affected in the same way towards these different actions, that he would make no distinction between them, but approve or disapprove them equally, is too glaring a falsity to need being confuted" (1.8). In Section 2.17 Butler carries the argument one step further. If human beings lacked all superior principles, in particular the supreme principle of conscience, they would have to approve or disapprove of any two acts equally—no matter how "cool" and objective the hour of their reflections. This absurdity establishes that humans do possess superior principles as an essential part of their constitution.

The nature of my thesis has required a one-sided, negative, criticism of Sturgeon's essay. I trust that this will not conceal my appreciation of the importance of Sturgeon's contribution. Leaving aside numerous insights which it has not been my purpose here to pursue, Sturgeon's formulation of the Naturalistic Thesis and his elucidation of its incompatibility with the Supremacy Doctrine has sharpened, enriched, and quickened our comprehension of Butler's ethics.

3I am grateful to Nicholas Sturgeon and A.I. Melden for their criticisms of an earlier draft of this paper.

Michael W. Martin
Department of Philosophy
Chapman College
Orange, California 92666