

## SEARLE ON INSTITUTIONAL OBLIGATION

### I

John Searle sets out five statements, the relation of each to its successor being one of entailment (given the addition of certain tautologies).

- (1) Jones uttered the words 'I hereby promise to pay you, Smith, five dollars'.
- (2) Jones promised to pay Smith five dollars.
- (3) Jones placed himself under an obligation to pay Smith five dollars.
- (4) Jones is under an obligation to pay Smith five dollars.
- (5) Jones ought to pay Smith five dollars.<sup>1</sup>

Searle's avowed aim is to bridge the so-called fact-value dichotomy. Thus,

... no set of *descriptive* statements [it is often claimed] can entail an *evaluative* statement without the addition of at least one evaluative premise. To believe otherwise is to commit what has been called the naturalistic fallacy.

I shall attempt to demonstrate a counter-example to this thesis ... if we can present a plausible counter-example and can in addition give some account or explanation of how and why it is a counter-example, and if we can further offer a theory to back up our counter-examples—a theory which will generate an indefinite number of counter-examples—we may even incline ourselves to the view that the scope of that thesis was more restricted than we had originally supposed.<sup>2</sup>

We are, then, to be offered an actual deduction of an evaluative state-

1. John Searle, "How to Derive an 'ought' From 'is'," p. 102, originally in *Philosophical Review*, 1964. My quotation page numbers refer to the pagination of the article in Philippa Foot, ed., *Theories of Ethics* (Oxford: Oxford University Press, 1968). The derivation is restated, with some changes, in John Searle's book *Speech Acts* (Cambridge: Cambridge University Press, 1969), Chap. 8.

2. "How to Derive an 'ought' From 'is'," p. 101.

ment from factual premisses, and a theory as well, which will place us "in a position to see how we can generate an indefinite number of proofs." Although, as we shall see, the deduction Searle offers us concerns itself with the institution of promising, Searle says that the same trick, of deriving normative statements from factual ones, can be carried through for *any* institution which is constituted, or defined, rather than merely regulated, by its rules. Other examples of such deductions are: that I ought to leave the field when the umpire shouts 'Out' following from the rules of baseball; 'One ought not to steal' as a rule of private property; evaluative statements following from the rules of marriage, loans, and debts.

The fact that Searle's example is concerned with promising may well introduce unnecessary complications which would not appear had the derivation actually been presented using for example the rules of baseball. What is important is that the derivation proceeds from rules (constitutive rules) to evaluative statements. There is no suggestion that the rules used as premisses of the argument be moral rules; indeed, were the rules *moral* rules, then Searle's claim to have bridged the fact-value gap would be immediately vitiated, since some of the premisses would then be evaluative. Thus, the constitutive rules are not moral rules, nor are the 'ought' statements, or statements of obligation, so derived, held to be statements with occurrences of moral 'ought' or moral 'obligation'. As Searle claims, "we are concerned with 'ought', not 'morally ought'."<sup>3</sup> Searle's use of promising might be misleading to the reader, for after all aren't the rules of promising *moral* rules, and hence any derived evaluative statement about the obligation to keep a promise a *moral* statement? Not according to Searle, for he does not consider promising an essentially moral-obligation-producing institution.

I think . . . that the obligation to keep a promise probably has no necessary connection with morality. It is often claimed that the obligation to keep a promise is a paradigm case of a moral obligation. But consider the following very common sort of example. I promise to come to your party. On the night in question, however, I just don't feel like going . . . I just don't go; am I immoral? Remiss, no doubt. If it were somehow very important that I go, then it might be immoral of me to stay home. But then the im-

3. *Speech Acts*, p. 176.

morality would derive from the importance of my going, and not simply from the obligation undertaken in promising.<sup>4</sup>

What Searle is offering then is an example of, and a method for further production of examples of, the derivation of a (nonmoral) evaluative statement from nonmoral constitutive rules. The use of the example of promising should not mislead us into thinking of the rules in the premisses of the derivation as *moral* rules, for Searle explicitly denies that they are in the case of promising, and in any case the use of moral rules as premisses would introduce evaluative premisses into the argument, vitiating at once Searle's attempt to bridge the fact-value gap.

## II

If we think of obligation-producing rules, we can be led to ask: who is the subject of such obligations? We could say more simply: whom does the rule apply to?, but we must be careful here. For instance, consider the following. There is a rule in Ireland, 'Drive on the left hand side of the road'. But no one thinks that the rule implies a statement of obligation for me, since I am a resident of the Commonwealth of Massachusetts. Malaysian rules are not thought to obligate Thais, nor Soviet law obligate Turks, nor canon law Jews. But we must understand 'apply to' or 'is the subject of' in a very special way. After all, in one sense of 'apply to', the cited Irish rule won't apply to nondriving Irishmen. Still there is a difference between the Irish law not applying to a nondriving Irishman and the Irish law not applying to a driving resident of the Commonwealth of Massachusetts, and, in the sense of 'apply to' or 'is the subject of' that we are using, we mean Irish laws to apply to the former but not the latter, even the law about which side of the road to drive on. We mean something like: who is the law addressed to, or to whom was it promulgated?

In considering to whom a law applied, or who are its addressees, there is a temporal dimension as well as a spatial one which must be considered. The rule 'One ought to burn witches promptly after their trial' is not addressed to me, or does not apply to me, although I am a citizen of Massachusetts and the law did apply to residents of seventeenth-century Massachusetts.

4. *Ibid.*, p. 188.

These remarks, utterly trivial as they are (although the notions of promulgation and addressee of a law are enormously complex), do force Searle to answer the question: whom can the constitutive rules obligate, whom do they apply to, to whom are they addressed?

In his original article, Searle disregarded this question, except for one remark: "By undertaking to play baseball, I have committed myself to the observation of certain constitutive rules."<sup>5</sup> In his book *Speech Acts*, however, the problem receives a great deal more attention.<sup>6</sup> For instance, Searle considers the possible rejoinder that, as far as his doctrine goes, he could be obligated by some institution in Australia, wholly unknown to him. Searle replies that "the notion of an obligation is closely tied to the notion of accepting, acknowledging, recognizing, undertaking, etc., obligations in such a way as to render the notion of obligation essentially a contractual notion." Thus, with regard to this hypothetical group in Australia which was presumed to be capable of obligating me to do some action, "Unless I am somehow involved in the original agreement, their claims are unintelligible."<sup>7</sup> Since according to Searle, obligation is a contractual notion, I must have *accepted, agreed to, or acknowledged* that those rules are to govern my actions.

In explicating the notion of institutional obligation in this way, I take it that Searle means to be offering necessary and sufficient conditions for being obligated by an institution. Presumably, to claim that "the notion of [institutional] obligation [is] essentially a contractual notion" is to claim that, if I have made a contract with an institution to do something, then I have an obligation to do it, and if I have an obligation to do something which an institution requires (apart, I suppose, from any separate moral obligation I might have to do the act in question), then I have made some sort of contract. Searle describes the contractual element in various ways: accepting, acknowledging, recognizing, or consenting to something in some way. Now, since Searle could not want in all seriousness to maintain that I can be obligated only by institutions with which I have expressly made a contract, or by which I have expressly undertaken to abide, he is forced to go the way of consent theorists in political

5. "How to Derive an 'ought' From 'is,'" p. 113.

6. *Speech Acts*, pp. 188-198.

7. *Ibid.*, p. 190.

philosophy and develop a doctrine parallel to that of tacit consent or unwritten contract. How do we give our consent to, or make a contract with, these institutions? As an example, Searle claims that "in the case of linguistic institutions like promising (or statement making) the serious utterances of the words commit one in ways which are determined by the meaning of the words."<sup>8</sup> I do not *explicitly* consent to, or undertake to abide by, the rules of English or promising. Rather, insofar as I make a serious utterance of a certain sort, I (tacitly) undertake or commit myself to follow the rules of English or promising. I do not need to actually say 'I undertake to . . .'. Indeed, in the case of speaking English or promising, it is not clear that I could commit myself to speaking English or promising by making a promise in English. Rather, I commit myself to, or undertake to abide by, these institutions in other ways:

When I do assert literally that he made a promise, I do indeed commit myself to the institution in the sense that . . . I . . . undertake to use the word 'promise' in accordance with the literal meaning, which literal meaning is determined by the internal constitutive rules of the institution.<sup>9</sup>

Finally,

The point is merely that when one enters an institutional activity by invoking the rules of the institution one necessarily commits oneself in such-and-such ways, regardless of whether one approves or disapproves of the institution.<sup>10</sup>

As I indicated above, Searle does use several different words or phrases as more or less synonymous ways of describing the relation which must hold between the constitutive rules of an institution and the person obligated: undertaking to abide by, committing oneself to the observance of, accepting, recognizing, acknowledging the rules of the institution or the obligations which they impose. But since Searle explains 'commit', for example, in terms of 'undertake', and the notion of 'undertaking' seems to occupy a central place in Searle's account, it is with this last notion that I shall deal, rather than the notion of a tacit contract for example.

8. *Ibid.*, p. 189.

9. *Ibid.*, pp. 194-195.

10. *Ibid.*, p. 189.

It is important to get clear on what the notion of 'undertake' is and what Searle uses the notion for in the derivation itself. 'Undertake' itself has a range of meanings in the dictionary, some of which make it clearly a moral concept and others which do not. For example, three of the meanings are: 'enter into or upon (a task, journey, etc.)'; 'to give a promise or pledge that'; 'to make oneself responsible for'. The first meaning is a nonmoral use of 'undertake', the latter two are moral, or connected with moral concepts in certain ways. Searle must intend 'undertake' in something like one of the latter two senses, for he criticizes Hare for not recognizing that 'commit' (which Searle explains in terms of 'undertake') is normative:

It is a matter of immense fascination to me that authors who are 'anti-naturalists' when they think about it, tacitly accept the derivations of evaluative from descriptive when they are just doing philosophy and disregarding their ideology. Consider the following passage from R. M. Hare:

"If a person says that a thing is red, he is *committed* to the view that anything which was like it in the relevant respects would likewise be red."<sup>11</sup>

This then is Searle's reply to our scope question. Who does a rule apply to? Who are subjects of institutional obligations? Using one convenient dictionary entry for 'undertake': anyone who has made himself responsible for abiding by the rules. But if this is the case, surely an extra premise should be added to the derivation. Thus:

- (1) One ought to leave the field after three strikes.
  - (2) Jones has just had his third strike.
  - (3) Jones has made himself responsible for abiding by the rules (or, . . . is committed to abiding by the rules).
- ∴ (4) Jones ought to leave the field.

But, as stated, the derivation won't do for Searle's purposes, since one of the premisses, (3), is evaluative. What we now want to look at is what (if any) descriptive premisses could entitle Searle to (3), and which could then replace (3) so that the derivation could proceed from purely descriptive premisses to an evaluative conclusion.

If we recall the earlier quotes from Searle, when I claimed that

<sup>11</sup>. *Ibid.*, p. 190.

he was forced to develop a notion of *tacit* consent or *implicit* contract, it is plain that for Searle the criterion of implicitly undertaking to abide by the rules, or of making oneself responsible for abiding by the rules, of a constitutive rule-governed institution is simply the doing of some acts which are defined by the institution. Thus, Searle claimed that the criterion of making oneself responsible for abiding by the rules of English was simply the making of a serious English statement; the criterion of making oneself responsible for abiding by the rules of baseball was playing baseball (stealing a base, batting, or whatever). It is no inessential feature of Searle's position that the criterion of making oneself responsible for abiding by the rules is so weak, for the criterion must embrace all those whom we do hold to be obligated by institutions (and thus *explicitly* promising or making oneself responsible would be too strong by ruling out people whom we do think obligated) and further the criterion must not include any evaluative element (otherwise some of the premisses in the derivation would be evaluative). However, if Searle's weak criterion of making oneself responsible for abiding by the rules, which is nonevaluative, works, then we are entitled to (3) above by descriptive premisses, and thus the derivation will have no irreducibly evaluative premise.

But Searle's weak, nonevaluative criterion for making oneself responsible for abiding by rules does not work. The relation between the fact that someone has performed some institutionally defined action, or observed some constitutive rule on some occasion or set of occasions, and the evaluation that the person is responsible for abiding by those rules is an extremely complicated relation. Suppose for example that Jones is coerced into playing baseball by certain threats. Then Jones is not responsible for playing baseball, or continuing to follow the rules of baseball once the threats have been removed.

But even if we restrict our consideration to acts where no threat is part of the agent's reason for doing the act,<sup>12</sup> there remain certain

---

12. Even if this would involve no moral concepts, which it does seem to do. See, for example, the notion of 'morally expected course of events' in Professor Robert Nozick's "Coercion," in *Philosophy, Science and Method: Essays in Honor of Ernest Nagel*, ed. by Sidney Morgenbesser et al., (New York: St. Martin's Press, 1968).

other important complications. Suppose for example that I find myself as an officer in a concentration camp whose duty it is to execute Jews. At some point I come to see that what I am doing is evil, and I decide not to carry out certain duties I have been given, connected with the executions. It is not the case that I withdraw from the institution (were this possible), for resigning or withdrawing from an institution does normally cancel any obligations one might otherwise have had.<sup>13</sup> Rather, I remain in the institution, but lie about my having performed the executions. Now the question arises: have I in any sense whatever made myself responsible for those executions, simply because I have performed the executions in the past and remain a bona fide member of the institution? Do I have any responsibility to not lie and to execute the Jews? The answer is plainly 'no' to these questions, and this can be seen in the way in which we want to answer the following kinds of questions. Even supposing I would not be punished for refusing to execute the Jews, do I owe my superiors any excuse, apology, or explanation for my refusal to abide by the rules of the concentration camp? Would it be morally better for me not to execute the Jews and not break camp rules (perhaps by changing my duties by being transferred to another department) than to not execute the Jews and not fulfill my duties as an executioner? For me to be responsible for abiding by rules, and for you to be entitled to expect me to abide by them, it would seem that the rules would have to constitute a morally permissible practice or institution. However, this qualification to Searle's criterion for making oneself responsible for abiding by the rules of an institution appears to introduce an irreducibly evaluative element into the derivation. One premise, the one which states the conditions under which institutionally defined action (e.g., the serious making of an utterance) constitutes commitment to observance of the rules of that institution in the future, concerns itself with the moral permissibility or impermissibility of engaging in that practice or institution, and thus seem normative in its import.

13. Sometimes, of course, institutions do not recognize or allow ways to opt out of them voluntarily. Suppose in the case I am imagining that there is no recognized way to disassociate myself from the concentration camp, and *not* just that I am afraid of what will happen to me if I do disassociate myself voluntarily.



## III

But even apart from the criticism that Searle's derivation contains at least one evaluative premise, in what way, we might ask, are we still on the trail of nonmoral, institutional obligation? What importance do constitutive rules have for the reconstructed derivation? Whatever conclusions we derive about obligation rest, as it were, on the fact that someone has made himself responsible for the doing of something, or morally entitled others to believe that he will do something. In whichever way we choose to word this, the obligation is a moral obligation, the duty of fidelity to my commitments, or execution of my responsibilities, or whatever. If I make myself responsible for doing *x*, then I ought to do *x* *morally speaking*. It does not matter whether '*x*' is 'observe a constitutive rule', or 'observe a set of regulative rules', or 'perform a particular action'. The bulk of Searle's argument, relying as it did on constitutive rules and institutionalized obligations, has become thoroughly superfluous, for according to Searle the derivation of 'ought' or the obligation statement from 'is' was supposed to depend on constitutive rules and the special kinds of obligations they imposed. But if the obligation is now to be based on the moral responsibility I may have assumed by past actions of mine to do certain other things, then the fact that *what* I have made myself responsible for is obedience to certain constitutive rules is irrelevant to the question of what kind of obligation I have assumed, for it is the responsibility and not the constitutive rules which, as it were, is generating the obligation. The derived 'ought' or obligation statements will be moral claims, not (so-called) nonmoral, institutional ones, for it is the same obligation that commitment or responsibility of any sort gives rise to whether or not what I make myself responsible for is the observance (in the future) of certain constitutive rules. Of course, the *way* in which I make myself responsible may have something to do with rules. For example, past habitual action of mine, even if morally permissible, normally gives no one a right to expect, nor me a responsibility to engage in, the same habitual action in the future. It is *principally* in the case of rules where we can argue from past actions to future obligations to action, but this last point applies to all sorts of rules and not just constitutive rules.<sup>14</sup> But the ways in which I can make myself re-

14. See my paper "Tacit Promising," forthcoming in *Ethics* (October, 1972) for a defense of this point.

responsible for the doing of something, or commit myself to do something, are many (even apart from the case of generating responsibility from my past actions in rule situations), and all such ways create precisely the same moral obligation, whether the situation is a rule situation or not, namely the duty to do whatever I am responsible for doing.

Searle's claim to have bridged the fact-value gap is without foundation, for the derivation, insofar as the conclusion really is that one has a duty to do something, rests on the suppressed moral premise that one ought to do what one has a responsibility for doing and the moral claim that the agent has a moral responsibility to do something. How can Searle not have seen that, in answering the question of the scope of the constitutive rules under discussion, he had allowed into the derivation evaluative premisses? And that, in so doing, moral obligations were being derived in the conclusion and not institutional ones? And that the discussion of constitutive rules and institutionalized obligations was being rendered otiose? I think that Searle's mistake can be traced to a confusion over two distinct senses of 'duty' and 'obligation'. To see this even more sharply, suppose some Searlean offered us the following example of a derivation which bridged the fact-value gap.<sup>15</sup>

- (a) 'x is a concentration camp marshall' =df. 'x has the duty of killing Jews in the concentration camp.
- (b) Jones is a concentration camp marshall.
- (c) Jones has the duty of killing the Jews in the concentration camp.
- (d) Everyone ought to do their duty.
- (e) Jones ought to kill the Jews in the concentration camp.

Since (a) and (d) are analytically true, and (b) and (c) seem, as a matter of fact to be true, (e) must be true. But since Jones has quite clearly, no reason at all to kill Jews, (e) is either false and hence does not follow from (a)-(d), or true, but only to be

15. I take it that it is not unfair to claim that a Searlean could offer this derivation, assuming of course that the rules of the concentration camp *define* certain roles and activities, e.g., being a camp marshall or calling evening role, and do not merely regulate them. Searle claimed that "we can generate an indefinite number of proofs," and presumably such a proof can be carried through for *any* institution which is defined or constituted by its rules.

taken in a descriptive sense, such that even if it is true that one ought to do something, it is not necessarily true that one has any reason whatever to do the act in question.<sup>16</sup>

What Searle has apparently done is to take the occurrence of the words 'duty' or 'obligation' in the premisses (statements (3) and (4) in the promising derivation) as normative occurrences of those words, when in fact these occurrences are only descriptive in the sense explained. That is to say, *IF* we think of institutions only as sets of rules, then the institutionalized obligations or duties which are part of the definition of the institution and are internal to it are descriptive duties and obligations which a man may have absolutely no reason to carry out. Using Searle's terminology, his mistake has been to confuse the institutionalized duties and obligations a man might have in virtue of the rules internal to the institution, with the 'external' statements of duty and obligation, only the latter of which are normative or reason-producing, and which do occur in the derivation once the notion of responsibility (or commitment) has been introduced.

'Duty' seems clearly to have both, a descriptive and an evaluative sense. I think that 'obligation' can bear both senses as well, for in ordinary parlance to say that a man had a legal obligation does not necessarily imply that the man had any reason whatever to do the action in question. Not very much hangs on whether or not I have interpreted ordinary parlance correctly. All I want to claim, is, that Searle must be understood to be using 'obligation' in this way, regardless of whether or not it is correct to do so in ordinary language. However, I think 'ought' does *not* bear both senses; if it is true that a man ought to do something, then it does follow that he has some reason to do it. If 'ought' had a descriptive sense, we could have held that (e) was true and did follow from (a)-(d), provided of course that it bore that descriptive sense in (e). Since it seems that 'ought' never does have such a sense, it seems more plausible to hold (e) false and hence deny that it follows from (a)-(d). (e) doesn't follow because 'duty' is used equivocally in the argument. In (d), 'duty' has its evaluative or normative sense, whereas in (c) it bears

16. For an elaboration and defense of the distinction between normative and descriptive senses of 'obligation' and 'duty', see my paper "Natural and Positive Law Revisited," *The Modern Schoolman* (May, 1972).

a descriptive sense since it occurs in a rule of the institution. *Pari passu*, in Searle's original derivation, (5) doesn't follow from (1)-(4), because 'obligation' is used equivocally in (4), where it bears a descriptive sense, and in (4b),<sup>17</sup> where it bears a normative sense. One has some freedom in choosing precisely where to place one's wedge in these derivations, but I have chosen what I think are the most plausible places.

Finally, throughout this paper I have assumed, with Searle, that promising is an institution *like* baseball or chess in certain important logical respects. I have tried to show that some very general things Searle claims about institutions constituted or defined by their rules won't bear up under examination. For example, I argued that the obligation 'internal' to baseball is not normative in my sense, but that, in order to obtain a normative statement about what someone playing baseball has reason to do, certain other premisses which did not concern baseball at all (for example, a premise about responsibility) had to be added. One might cavil at this argument when applied to promising itself. When I promise, I am obligated to do what I promise to do. Surely *that* is a normative conclusion, and no further premisses need to be added. I have no objection to this, and perhaps an example in terms of promising really does close the fact-value gap. If one is inclined to this view, my argument can be construed as pointing out the enormous difference between promising and the other examples Searle uses. In any case, if Searle were right about the particular case of promising, his point is not generalizable to constitutive rule-governed institutions. And given the important differences here, we may come to wonder whether it is *at all* enlightening to think of promising on the model of an institution defined by a set of rules.

DAVID-HILLEL RUBEN

UNIVERSITY OF GLASGOW

17. "(4b) one ought to do what one is under an obligation to do." (4b) is analytic only if 'obligation' is normative. Searle doesn't number this as "(4b)," but in passing from (4) to (5) he claims that one needs a tautology analogous to (3b), and says that the relevant tautology is the one I have numbered (4b).