

Autonomy and self-respect

THOMAS E. HILL, JR.
The University of North Carolina at Chapel Hill

U.C.L.
INSTITUT SUPERIEUR DE PHILOSOPHIE
Bibliothèque
Collège D. Mercier
rue Cardinal Mercier, 14
1348 Louvain-la-Neuve

 **CAMBRIDGE**
UNIVERSITY PRESS

Promises to oneself

When one makes a promise to others, one puts oneself under obligation to them. Can one then incur an obligation to oneself by making a promise to oneself? If not, why not? Do the disanalogies between promises to others and "promises to oneself" show that "promises to oneself" are morally insignificant or, more generally, that obligations to oneself are impossible?

There are several reasons for raising these questions.

First, the idea of promises to oneself is familiar but puzzling. We often talk of such promises in jest, but sometimes quite seriously. Some find the idea quite natural, while others dismiss it as absurd. These phenomena suggest conceptual tensions of some philosophical interest in themselves, apart from any implications for larger issues.

Second, "promises to oneself" are frequently invoked in the controversy about whether one can have moral obligations to oneself. For example, some philosophers ridicule the idea of obligations to oneself by noting that I am not *morally* to blame if I "promise myself" a treat and then fail to take it. Obligations to oneself, it is assumed, would be self-serving requirements whereas morality is concerned with interpersonal relations. Other philosophers use the paradoxical idea of a "promise to oneself" to illustrate and support their general contention that "obligations to oneself," construed literally, are logically impossible.¹ The example is important to their case because promising is a paradigm of putting oneself under obligation to someone in the most *literal* sense.² To assess these arguments, then, we need to look closely at the analogy between promises to others and "promises to oneself."

Third, reflection on this analogy gives us a new angle from which to rethink the old question, why are promises to others morally binding? If, as I suppose, promises to others generally result in moral obligations and "promises to oneself" do not, what is the difference? Contrary to common assumptions, I suggest that the explanation is neither that obligations to oneself are logically incoherent nor that morality is concerned only with our relations to others. Moreover, there is

¹ See, for example, M. G. Singer, "Duties to Oneself," *Ethics* 69 (1959): 202-5.

² It is controversial, for example, whether we have obligations that are, in a literal sense, obligations to animals, deceased persons, or future generations. It may also be doubted whether obligations of charity are, literally, obligations to the needy persons who are the beneficiaries of their performance.

reason to doubt that rule-utilitarian and contractarian theories adequately explain the difference, for a "practice" of promising to oneself is possible and might prove useful.

Finally, though I agree with the common view that "promises to oneself" do not *generally* create moral obligations, the apparent exceptions to this negative conclusion lead to a more constructive suggestion. That is, if there are obligations to oneself, they are not simply obligations to promote one's own welfare or to follow the rules of a useful practice; they are more likely to be found among the moral requirements of self-respect.

My discussion will be divided as follows. In the *first section* I try to press as far as possible the idea that "promises to oneself" are, or rather could be, analogous to promises to others. There are, I suggest, far more similarities than one might suspect, and they help us to see what might be meant by speaking of an obligation as "to oneself." The *second section* takes up disanalogies, with particular attention to two that have been thought to make obligations to oneself logically impossible. While acknowledging differences, I argue that the disanalogies do not establish the logical impossibility of promises to oneself. In the *third section* I consider the moral significance of "promises to oneself." Although I agree with the dominant view that *these* do not *in general* create moral obligations, I suggest that this is for reasons other than the reasons that are supposed to exclude *all* obligations to oneself. Moreover, the apparent exceptions point to the sort of consideration that could more plausibly ground an obligation to oneself.

I

So that we can compare "promises to oneself," let us first consider some of the salient features of promises to others.

First, under what conditions do we say that a promise has been made? Clearly the person who makes the promise and the person to whom it is made must be at least minimally rational, capable of understanding a language, foreseeing to some degree likely consequences of their acts, *etc.* They must be conscious and aware of what they are doing at the time the promise is made: for example, not talking in their sleep, babbling in a drugged condition, and so on. One person, A, says something to another person, B, about a future act: typically, "I promise to X," but other expressions suffice when the context makes clear that this is what is intended. A and B must take what is said seriously, not in jest, as a line in a play or an example in philosophy class, *etc.* But, of course, A may not really intend to do X: lying promises are still promises. To distinguish promises from threats (e.g., "I promise I will kill you if you do that"), we should perhaps add that A believes that B has some interest in X and that B does have such an interest or at least lets A believe so. Both A and B must understand what promising means. This implies that they understand that, unless B releases A

AUTONOMY AND SELF-RESPECT

(e.g., says, "Forget it") or some other standard excusing condition obtains, then, by the conventional understanding of promising, A will be wrong not to do X and B will be entitled to complain. Though A may not actually intend to do X, A must intend that B believe that he intends to do X (or at least that B takes what he says as an attempt to have B believe this).³

Second, under what conditions is a promise, once made, fully binding? The concept of promising allows that if various circumstances arise, a promise is no longer binding. Most obviously, if the person to whom the promise was made, B, releases the promisor, the promise is cancelled. Again, unforeseen circumstances may arise so that it is obvious without B's saying anything that the original point of the promise will not be served by carrying it out. The promisor may, through no fault of his own, become incapable of fulfilling the promise. The promise may have been to do something morally wrong on other grounds, or conditions may have changed so that the only way to fulfill it is to violate some more stringent moral requirement. In these cases it is generally accepted that the original promise need not be kept, even though the promisor in some instances may still owe something to the promisee.

Third, what conditions must obtain in most cases if the practice of promising is to seem reasonable and to survive as a convention? Though particular promises could be made without the following conditions being satisfied, if these were not present as a rule, I think, we would not continue to make promises. (1) B actually has an interest in X. (2) There is some reason for A to make the promise, for example, A gains something in exchange or believes that B ought to be assured of A's Xing. (3) B believes that A is more likely to X if A promises. (4) A is really more likely to X if A promises. (5) In saying "I promise to X," A seriously resolves, makes it a matter of principle, to X; A regards this as more than a casual plan or intention and is willing to incur feelings of guilt and lower self-esteem in case of failure, except when there are conflicting obligations or excuses. (6) Others tend to think less of A if A fails to do as promised. (7) B remembers the promise and will be disappointed and prone to criticize A if A fails to keep it.

By analogy, then, what would be required for me to make a binding promise to myself? I must, of course, be sane and conscious. I would need to say something, at least silently, to myself. The words could be "I promise myself that I will X." I must not be joking, merely amusing myself, illustrating a philosophical point, *etc.* I must have some reason for wanting to do X, for example, a belief

³ This last qualification seems required to account for the following case. A tells a friend that he does not intend to do X and asks him to convince B of this without saying how he knows. Then A says "I promise to do X" to B, not intending to cause him to believe that he intends to do X but meaning for B to think him to be trying to fool him with a false promise. B, for reasons of his own, pretends to believe that A will do X. A promise, I think, has been made. The example is from Holly Thomas.

PROMISES TO ONESELF

that it promotes my interests or furthers some morally worthy end. I must intend to do X, for, unlike the case of promises to others, there is no possibility of getting the promisee (myself) to believe that I intend to do X unless I really do, at least if we set aside the half-believing and half-disbelieving condition of self-deception. Crucially, if what I say is to be more than an expression of intention, I must understand that my words "I promise myself" imply that I will be wrong and liable to legitimate self-criticism if I fail to do X, barring standard excuses and justifications.

Consider next conditions which must be met for a promise, once made, to be fully binding. One can easily imagine promises to oneself analogous to promises to others insofar as they would not be regarded binding under certain conditions, for example, if circumstances so change that keeping the promise becomes obviously pointless, if it conflicts with more stringent obligations, or if it turns out to be impossible through no fault of the person who makes the promise. We can even imagine that those who make promises to themselves have a well-understood procedure for releasing themselves from the promise when necessary. Suppose, for example, it is understood that one can cancel the promise if, but only if, one has *reconsidered* it carefully in the light of *new*, unanticipated *information*, without the *immediate pressure* of expected temptations, and one has reached a definite decision that it is best, all considered, that the promise no longer remain in force. Taking the point of view of the recipient of a promise to myself, I could cancel it when unforeseen events make it clearly detrimental to my original aims and values; but taking the point of view of the promisor, I could not escape my obligation at will but only when special circumstances arise. These circumstances might be understood to include an explicit statement to myself that the promise is cancelled, though, as we shall see, they must include more than this. In the absence of these special circumstances, failure to do what one promised oneself to do would be viewed as breaking the promise, not releasing oneself from it.

So far, setting aside for now possible disanalogies, it seems as if promises to oneself are conceivable. But would anyone have adequate reasons to make them? Could such promises, understood as I have described, serve the purposes of enough people so that they would continue to understand the words "I promise myself" in the required way? Let us review the conditions that must obtain in general if we are to continue to make promises to others. Are the analogous conditions met for promises to oneself?

(1) Typically, I suggested, B must have an interest in the fulfillment of the promise. This would be so for promises to oneself if persons typically promised themselves to do only what in their best deliberative moments they wanted themselves to do at a later time. An example would be the reformed drug addict who promised herself to refuse all offers of the drug she used to crave. She has an interest in doing what she promised herself.

(2) Promising would continue only if there are reasons for A to promise, e.g., sometimes A believes that promising will promote A's own interests. This condition is easily satisfied in the case of promises to oneself provided the previous condition is met and that making the promise substantially increases the likelihood that the desired future act will be performed (see [4] below). Since, by (1), B stands to benefit from the proposed act and, by hypothesis, A and B are identical, it follows that A has an interest in making the promise if doing so makes it more likely that the act will be done.

(3) B typically must believe that A is more likely to perform if A makes the promise. This condition will be met if the next condition, (4), is satisfied, provided only that the persons who make promises to themselves are reasonably well informed. That is, informed persons will believe themselves more likely to do something if they promise themselves to do it, provided that in fact this is true.

(4) Promising would not survive as a practice unless promising typically makes the promisor more likely to do the thing promised. Now we have supposed that promises to oneself are usually made when the agent, on due deliberation, wants to perform some future act and is capable of performing it. Let us suppose further that such promises are normally made only when the agent, in deliberative moments, anticipates impulses and temporary pressures that will be temptations to deviate from the plan judged best. In this situation one would have good reason to want to ensure, or at least make more probable, that one will overcome the anticipated obstacles and act on the plan that, under deliberation, seemed best. If, by making a promise oneself, one makes oneself liable to feelings of guilt for yielding to contrary impulses, one increases the likelihood that one will do the desired act. Making the promise to oneself also strengthens one's motivation by making one liable to the disapproval of others who may know of the promise or liable, if others do not know, to the unpleasant feeling of hiding to avoid disapproval. By hypothesis, if I make a serious promise to myself I must believe that noncompliance, barring special circumstances, is wrong, and so I should expect to feel guilty. Others who accept promises to oneself in the same way and who learn of my noncompliance must also believe that I have done wrong and so can be expected, in general, to think less of me for it.

It should be noted that even if the viability of promises to oneself depends in this way on the sanctions of others, this does not imply that the resulting obligation is *to* others. Liability to criticism by persons other than the promisee is also a supporting feature of the practice of promising to others, and this does not mean that all such promises are *to* society or *to* the would-be critics.

(5) Promising would continue, I suggested, only if typically the promisor genuinely intends to do as promised. In the case of promises to oneself this condition is readily met, for one cannot promise oneself to do something that one has no intention of doing. For A must intend for B to believe that A intends

to comply, and this is impossible if A knows that B (here = A) knows that A has no intention to comply. In order to make a promise to oneself one *has* to intend to do as one promises.

(6) The maintenance of promising, I have assumed, is conditional on the sanction of public opinion. This sanction could work to motivate us to keep promises to ourselves as well as promises to others provided that there are others who believe that breaking a promise to oneself is wrong and that they know when these promises are made and broken. We could keep our promises to ourselves secret; but insofar as their purpose is to increase the likelihood that we will resist anticipated temptations and carry out best deliberative plans, we would always have a reason not to be secretive about such promises. Others may have less motivation to express disapproval when one breaks a promise to oneself than when one breaks a promise to another, but most people are already so willing to condemn the weakness and imprudence of others that we can hardly expect that they would refuse to add pressure to those known to have made promises to themselves.

(7) Because one will normally have a self-interested reason for wanting to see one's promises to oneself fulfilled, one should be even less inclined to forget them than to forget promises to others. But, one may wonder, would people have adequate motive to impose sanctions on themselves for breaking promises to themselves? In the case of a promise to oneself the person especially entitled to complain about nonfulfillment, the promisee, must be the very person who made the promise. To be an object of complaint and criticism, especially self-criticism, is not pleasant, and so we might expect unfulfilled promises to oneself to provoke less spontaneous and vehement criticism than is typically aroused by broken promises to others. When one offends against oneself, one is tempted to forgive if not forget. Nevertheless, though painful, self-criticism is a familiar fact of life, and in fact people are often harder on themselves than on others. Thus it is reasonable to suppose that people would often be self-critical upon breaking promises to themselves and knowing this would increase the likelihood of their fulfilling such promises.

II

The discussion so far suggests that genuine promises to oneself are possible; but have we overlooked crucial differences between promises to others and "promises to oneself"? Surely there are differences, and these may make all the difference.

The most obvious dissimilarity is that promises normally relate two or more persons whereas a promise to oneself is a reflexive relation involving only one person. This fact alone, however, is no obstacle; for, while some relations are the sort one can have only with others (e.g., marrying, dragging by the heels), other relations are of a kind that one can have either with others or oneself (e.g.,

being attorney for, hitting). Such relations do not necessarily require a schizophrenic view of the person, dividing him, for example, into phenomenal and noumenal selves, actual and ideal selves, or present and futures selves. The question is whether the particular relation, promising, is a relation that one can have to oneself without troublesome metaphysical dualism. To determine this we need to look more specifically at the differences between promises to others and "promises to oneself."

One difference was noted in passing earlier: there can be no full-fledged lying promises to oneself. When one person makes a promise to another, the promisor can, with full intention and awareness, completely deceive the promisee about what he really intends to do, but no one can thoroughly deceive oneself in this way. Does this make a difference as to whether one could make a promise to oneself? Clearly not, for the practice of promising to others would in no way suffer if, miraculously, lying promises became impossible. Suppose, for example, that everyone's eyes would flash purple whenever they tried to make a deceptive promise and that, as a result, no one would take seriously such a would-be "promise." Then, because there would be no deceptive promises, the practice of promising would be even more useful than it is now. So the impossibility of making *lying* promises to oneself is hardly a reason for denying the possibility of *any* promises to oneself.

Controversies about private languages might call attention to the following apparent disanalogy: making promises to oneself is a language game that is parasitical on promising to others. The latter presupposes a public setting, a shared "form of life." If there were no such observable promises to others, we could not have learned to understand "promises to oneself."

This may well be true, but it does not show that promises to oneself are impossible. Presumably one could not talk to oneself if one had not first learned to talk to others, but this does not mean that one cannot talk to oneself. To say that promises to oneself are possible is not to say that one could make such promises if one grew up, like Wolf-boy, without human contact or in a society without public promises.

Another difference worth considering is this: when a person violates his promises to another to gain something for himself, typically the other person loses completely in the transaction; but when a person violates his promise to himself in order to get something for himself, no one loses completely.⁴ That is, assuming the original promise to himself continues to be in his best interest, when he breaks the promise he sacrifices long-term interest for some lesser (usually more immediate) gain; but the broken promise is not altogether opposed to his wants and interests. By contrast, a broken promise to another person may completely oppose that person's wants and interests.

⁴ This point was suggested to me by Morton Beckner.

In the absence of further information, we should no doubt prefer (a) a situation in which no one loses completely and someone gets something (though less than what is best for him) to (b) a transaction in which one person loses completely and another gains something (perhaps even what is best for himself). Typically broken promises to oneself exemplify (a) and broken promises to others exemplify (b). But the most we could infer from this is that, typically, breaking promises to others is more objectionable than breaking promises to oneself. Our question, however, is whether promises to oneself are impossible, and the disanalogy under consideration fails to show this. It is not essential to promising that the parties have radically diverse interests. Suppose, for example, that two persons, John and Mary, have interests so similar or complementary that (for a range of choices in question) whatever benefits John to some degree benefits or pleases Mary, and vice versa. Now if John, to pursue an immediate desire, breaks a promise to Mary, Mary is not completely the loser, though she may not get what she most prefers. But this does not mean that they cannot make promises or that they would have no reason to do so. Promises may help them to count on each other to carry out coordinated plans for their mutual benefit, and they will also enable Mary to be more assured that John will do what she most wants in certain cases even when he might prefer to do something else (and *vice versa*.) If we were fortunate enough to live in a world so empathetic that everyone shared in the joys and sufferings of everyone else, promises would still be possible and would have a point.

The last disanalogy suggests another. Promises between two (or more) persons are often useful because each needs to be assured how the other person will act when they are not together to consult and coordinate their activities. Suppose, for example, that A and B know that they will not be able to communicate for a while and that during this period it is mutually advantageous if both do X or both do Y but harmful to each if one person does X and the other Y. They might select a plan, e.g., that both do X, and then assure each other by promising that they will stick to the plan despite anticipated temptations to deviate from it. The promise should increase their motivation to carry out the plan because it makes them liable to complain and to feel guilty if they should fail. Promises to oneself, however, do not serve a strictly analogous function. Despite normal memory lapses, *etc.*, no sane person is as completely "out of touch" with himself as two persons can be with each other. The left hand always knows, to some extent, what the right hand is doing. So, after all, the person who makes a promise to himself is not like John and Mary in the previous example: they needed promises because they were not in constant communication but there is no similar communication gap to give a point to a promise to oneself.

This objection notes a genuine dissimilarity, but not one that shows promises to oneself to be impossible or pointless. The objection calls attention to one rationale for promises to others that seems inapplicable to promises to oneself,

but this rationale does not seem necessary to promises of any kind. Suppose John and Mary are handcuffed together and capable of reading each other's minds. Suppose further that their relevant wants and interests are not only similar in content but identical in intensity and importance: that is, the events that satisfy John satisfy Mary to the same degree, and *vice versa*.⁵ Now, as in the case of one person, there is no communication gap and no diversity of interests. Would they have any use for promises? Apparently so; for, assuming that making a promise increases the likelihood of performance, John has an interest in Mary's promising to do what is mutually advantageous, and *vice versa*. In fact in the special circumstances we have imagined, John also has a self-interested reason for making the promise, and so does Mary. The rationale is similar to that behind a promise to oneself, namely, by promising one makes it more likely that one will overcome temptations and do what, all considered, one sees as best.

I turn now to two influential objections succinctly stated by M. G. Singer.⁶ The first disanalogy Singer notes is that in the case of obligations to others the second party can release the first from obligation whereas, according to Singer, no one can release *himself* from an obligation. That is, if A has an obligation to B, B can release A from obligation; no one can release himself from an obligation; therefore no one can have an obligation to himself.

Some of Singer's critics understood him to mean by the first premise that if A has an obligation to B then B can *at any time* release A; but, so construed, the premise is open to numerous counterexamples, as they readily noted. Infants and deceased friends cannot release us from our obligations, and yet it seems that we still have obligations to them. We would not lose our obligations to a friend if he fell into a coma and so became temporarily unable to release us. Again, suppose you promise a friend to take her money or drugs in trust for an anticipated period of stress, with the understanding that you will refuse to listen to any of her requests for the money or drugs in that period. Then under stress, as anticipated, she asks for the money or drugs, professing to release you from your promise. It seems that, despite what she says, she cannot release you then.

In response to this sort of example Singer explained that what is in question is the logical possibility of release, not whether in fact the promisee is so situated that he can actually do it. Being in infancy, comatose, or even deceased, Singer suggests, can be regarded as contingencies that do not establish that it is *logically* impossible for the promisee to release the promisor. If the friend under stress were to prove herself thoroughly competent, she could release you from your promise.

5 This is no doubt possible only within a limited range of events, but we can imagine that the promises concern only such events. An example might be two persons playing a gambling machine which has equal payoffs and costs.

6 *Op. cit.* Singer's position is discussed by Warner Wick, Daniel Kading, Mary Mothersill, and Frank Knight in *Ethics*, 70, 71 (1962-63). Singer's reply is "Duties and Duties to Oneself," *Ethics* 73 (1962-63): 133-42.

The first premise, then, is that if A has an obligation to B, it is *logically possible*, under some conceivable circumstances, for B to release A from the obligation. For the conclusion to follow, the second premise must be similarly qualified. That is, this second premise must state that it is logically impossible for a person, under any circumstances, to release himself from an obligation. But given the understanding of "promises to oneself" I have hypothesized, this second qualified premise is not true. There is a way in which one could release oneself from an obligation to oneself: by reconsidering the promise to oneself in the light of unanticipated new information, in the absence of immediate pressures of anticipated temptations, judging that it is best, all considered, to cancel the promise, and then explicitly declaring the promise cancelled. Admittedly, there remains a difference between the two-person case and the one-person case; for when one person promises another to do something, the second can release the first *simply* by saying (seriously, with awareness), "Forget it." There need be no new, unanticipated information, careful reflection, absence of temptation, etc. But all that is required to nullify Singer's argument is that release from promises to oneself is possible, not that it is possible in all the same circumstances as release from promises to others.

Singer's insight, I think, is valuable but less troublesome than he supposes. To extend the metaphor, I cannot be literally *bound* by ropes if they are so loosely draped around me that I can "free" myself *at will*, but this does not mean that if I am bound there is no way I can release myself. I might with some effort cut or undo the knots if appropriate objects happen to be around. Similarly, I could hardly have an obligation to myself if I could release myself at will, e.g., by simply saying "I release myself." But I could be bound by a promise to myself if release was not easy or automatic but still possible in special circumstances.

We might also question whether Singer's first premise is logically necessary. Suppose A promises B to give B no hard drugs no matter what B says. A and B, let us say, agree that the promise is for all time, not just for a momentary period of stress. It seems that B cannot release A, under any circumstances, for if A takes what B says as a release A will violate the original promise. Singer counters, cleverly, with the suggestion that the original promise really consists of two promises:⁷ (a) to refuse to give B hard drugs and (b) to refuse to accept anything B says as release. B can release A from the first promise (a) by first releasing A from the second promise (b). If B tries to make release from (b) impossible by adding a further promise [e.g., (c) to refuse to accept a release from (b)], then this will obviously lead to an infinite regress.

Singer's way of construing the promise, however, is not obviously the only or best way. Unless one is already committed to releasability as a precondition of an obligation to someone, why not view the original promise as one promise

7 "Duties and Duties to Oneself," *Ethics* 73 (1962-63): 133-42.

with a special content, i.e., to refuse-to-give-B-drugs-no-matter-what-B-says? Perhaps we believe as a matter of moral judgment that circumstances could arise in which A is no longer under obligation, e.g., the drug B has foresworn proves to be less harmful than previously thought and is in fact necessary to B's survival. But then it is questionable whether B has released A or whether the obligation is simply cancelled by changed circumstances. In any case, the possibility of release seems more a matter of moral judgment than of conceptual necessity. If A believed that nothing B said would release him from his promise, would his mistake, if it is one, be a failure to understand the concept of an obligation to someone or would it rather be that he has unduly rigorous moral principles?

Singer's second argument calls attention to the apparent fact that making promises to others confers rights in a way that "promises to oneself" cannot. Briefly, he argues: if A has an obligation to B, then B has a right against A; it is absurd to suppose that a person could have a right against himself; and so no one can have an obligation to himself.

At least part of the force of this argument is borrowed from the previous one, for one function of saying that B has a right against A is to imply that B can release A (if anyone can). As this point has already been discussed, what we need now to consider is whether there is anything more to the argument.

To refute, or support, Singer's premises in the most thorough way, we would need a detailed analysis of "having a right against someone." But because necessary and sufficient conditions for having a right are so controversial, I suggest that instead we review features which are at least characteristic of the situation in which one person has a right against another. Whether the features are necessary becomes important only when it is clear that they are absent in the case of promises to oneself.

Characteristically, when B has a right against A (that A do X), the following is true: (1) If, without B's informed free consent, A fails to do X, then, barring special excuse or justification, A is wrong. (2) In case of such failure, B is regarded as an especially injured party (even if in fact others suffer as much). (3) Similarly, B is the person normally regarded in a special position to complain (though others may sometimes rightly chastise A for failing in an obligation to B). (4) B is normally the person in a special position to release A from the obligation, if anyone can. (5) B is justified, barring special circumstances, in demanding that A do X, whether A wants to or not. (6) B is justified, barring special circumstances, in demanding that A do X, whether or not X is in A's best interest. (7) B is justified to demand compensation or at least apology if A should fail to do X without B's free informed consent and in the absence of special excuse or justification.

Though admittedly it sounds odd for us to speak of "rights against oneself," at least the first five characteristic features of rights could conceivably be generated by "promises to oneself" as well as by promises to others. Imagine, as before, a

community that acknowledges and reinforces a practice of "promises to oneself" that is as closely analogous as possible to the common practice of promises to others. Suppose, then, that with these conventional understandings I make a serious "promise to myself," thereby assuming the roles of both A and B for purposes of the conditions (1)-(7). Now, by our previous discussion, the following should be at least logically possible. I make a promise to myself but I can, in the special circumstances described earlier, release myself from the promise (4); but if I fail to keep the promise to myself, without such a release or special excuse, I do something wrong (1); nevertheless, I am the injured party, the recipient and usually the intended beneficiary of the broken promise (2); so I am in the most appropriate position to complain and criticize (3); and if release is impossible, I am justified in demanding of myself that I fulfill the promise, whether at that time I want to or not (5).

It may seem more doubtful that I would be justified in demanding of myself that I keep the promise whether or not it serves my best interests (6). For, after all, the point of promises to oneself, as imagined here, has been to reinforce one's motivation to carry out one's best deliberative plans. But, on closer consideration, there is less disanalogy here than it seems.

First, some promises to oneself may be to do what one believes one ought to do, on independent grounds, even though it is not in one's self-interest. In this case, the fact that keeping such a promise turns out to be against my best interest is not a "new circumstance" permitting release, and so, believing in promises to oneself, I would be justified in demanding of myself that I carry out such a promise even against my interests. The point of this unselfish promise, after all, was just to strengthen my motivation in anticipated circumstances of this sort.

Second, consider promises to oneself designed to further one's own interest without harming others. If new information (even that one's desires changed in unanticipated ways) shows the previous deliberation mistaken, then, provided immediate anticipated temptations are absent, I can simply release myself from the promise. Because special circumstances permit release, I need not (even according to [6]) demand of myself that I keep the promise contrary to my best interests.

Third, consider a case in which the promise to oneself was self-interested but release is impossible because one is under the pressure of immediate, anticipated temptations. Then would I be justified in demanding of myself that I keep the promise contrary to self-interest? Two replies might be given. One might argue that, even from a self-interested point of view, the *policy* of sticking to one's promises to oneself until one is again in a pressure-free position to reassess the promise is best in the long run, even if on some particular occasions it leads one to sacrifice maximum self-interest. Alternatively, one could argue that the release conditions I described for promises to oneself are too stringent and should be loosened to allow release under temptation when new information makes the

promise *obviously* obsolete. In any case, there is no need to deny (6) for promises to oneself.

The final characteristic feature of rights, (7), seems at least partly inapplicable to promises to oneself. If the heart of an "apology" is a frank admission of wrongdoing together with implied recommitment to avoidance of similar behavior in the future, then one might at least do this much in the case of a broken promise to oneself even though one cannot, literally, "apologize" to oneself. To demand compensation would be pointless, for the same person who demands would be receiving the compensation. Payment would, so to speak, go from one pocket to the other. Perhaps because no meaningful compensation is possible, people who break their promises to themselves tend to feel a need to *punish* themselves. When we break promises to others we can sometimes "pay them back"; but we cannot do this with broken promises to ourselves.

To conclude this survey, it seems far from clear that most characteristic features of "having a right against someone" must be absent in the case of promises to oneself. Where the analogy is weakest (regarding [7]), it is not entirely obvious that the feature is a *necessary* condition of rights rather than merely a typical or characteristic one. The notions of self-criticism, making demands on oneself, releasing oneself, and the like, are admittedly in need of further analysis; but there seems little reason to suspect that they can be understood only if we presuppose an untenable metaphysical dualism of "selves." Again, I concede that talk of *violating* one's own rights, *pressing* one's rights against oneself, and the like, seems either metaphorical or out of place. Still, our imagined practice of promising to oneself mirrors enough of the characteristic features of rights that it seems quite misleading, if not false, to insist that promises to oneself are conceptually impossible.

III

The thesis that it is conceptually impossible for obligations to arise from "promises to oneself" is often accepted uncritically, I suspect, because most people have an intuitive sense that in fact we do not in general put ourselves under *moral* obligation when we say, "I promise myself." When, for example, we say, "I promise myself a treat tonight," we rarely, if ever, understand this in the serious way required to make such a "promise" analogous to promises to others. In particular, we do not really believe that we will be morally wrong not to give ourselves a treat even if no unanticipated special circumstances arise. Such "failures" do not typically arouse guilt feelings. Even if we *could* make binding promises to ourselves, then, we rarely do so because we do not meet the conditions that would give our words ("I promise myself") the appropriate seriousness.

But suppose a person did meet the required conditions. That is, imagine

both that there exists the conventional understanding of "promises to oneself" that I have described and that a particular person satisfies all the criteria for making a promise to herself. Would we then say that she had put herself under a genuine moral obligation? Not in general, I think. We might be prepared to grant that she has a personal obligation to herself, reminding ourselves that legitimate uses of "obligation" are not restricted to the moral sphere but also have a place in games, in the law, and elsewhere; but, except in special cases, most would rightly resist the suggestion that a moral obligation was generated.

Suppose, for example, a young woman solemnly promised herself that she would become rich by the age of forty. Satisfying our conditions for serious promises to oneself, she actually believes that if she does not do so, barring unanticipated special circumstances, she will be morally wrong. Again, imagine that she promised herself, in all seriousness, that she would taste all thirty-one flavors of Baskin-Robbins ice cream. She had always been inclined to order the same flavor but became convinced that she might be missing something better, and so, to resist the old habit and gain more pleasure in the long run, she makes the promise, not merely resolving to taste all the flavors but intending to make it wrong for her to fail to carry through. Suppose, further, that when she later lets ample opportunities pass and does not taste all the flavors or make much money, she feels guilty. Our intuitive response, I suspect, would be that her guilt feelings are misplaced, that she was not really under a moral obligation to become rich and taste the thirty-one flavors even though she thought she was.

The situation would be quite different, however, if she had made the promises to some other interested person. Suppose, for example, she had promised her aged grandfather that she would become rich in order to assure the old man that his descendents would be well off and she promised the Baskin-Robbins flavor master that she would taste all thirty-one flavors so she could give the company a comparative evaluation. Then, though the ensuing obligations may not be viewed as especially stringent, there seems little doubt that they qualify as moral ones.

The contrast with promises to oneself is evident when we consider how the promisor might rebut a claim that she ought to be doing something that conflicts with what she promised to do. Suppose, for example, someone tells the woman, "You should spend less time making money and more time with your grandfather" and "You should save your ice cream money for famine relief." She would offer at least partially counterbalancing moral considerations if she replied, "But I promised my grandfather that I would be rich by forty and the Baskin-Robbins flavor master that I would taste all thirty-one flavors." By contrast, to say "I promised *myself* that I would be rich and taste all thirty-one flavors" would not be accepted as a claim to competing moral obligations. It might be acknowledged

as a legitimate insistence on a *right* to pursue self-interest in these matters, but hardly as a claim that she is not morally free to do otherwise.⁸

One might object that in failing to keep her serious promises to herself the young woman is at least morally at fault for not living up to her own moral convictions for, by hypothesis, *she* believes that it is wrong not to fulfill the promises to herself. But this objection misses the mark. By not fulfilling what she mistakenly believes to be a moral obligation, she displays a defect of character; but this does not mean that she has really failed to meet her moral obligations. As some would say, what she does is objectively right and only subjectively wrong. The proper advice when the woman feels guilty for not keeping her promise *to herself* to become rich and taste all thirty-one flavors seems to be this: "Of course a good person will try to live by her honest moral convictions, but in this case you are mistaken about what your moral obligations are. Saying 'I promise myself,' however seriously, does not in general make a moral requirement out of what you were previously free to do or not to do." By contrast, if she had made her promises to her grandfather and the flavor master, that sort of response would be inappropriate.

To summarize, the conclusion suggested by our investigation so far is this. A practice of "promising to oneself" is logically possible, and might serve a useful purpose. By considering this in some detail we found that we can make sense of the expression that an obligation is "*to oneself*," even when treating the term as more or less parallel with "*obligation to another*." But, nevertheless, we do not in fact regard promises to oneself as morally binding in most cases.

If correct, this conclusion should reopen for further consideration the old question, why are promises *to others* morally binding? This is, of course, a large and controversial topic, too much to undertake here. But our reflections on the analogy between promises to oneself and promises to others at least have some implications for the larger question.

First, we cannot explain the fact that promises to others are in general morally binding whereas promises to oneself are not by dismissing the latter as incoherent.

Second, it would oversimplify the difference to say that moral obligations are requirements to benefit others whereas promises to oneself are merely devices to benefit oneself. For one could make a promise to oneself to underscore one's resolve and strengthen one's motivation to pursue altruistic ideals as well as self-interested goals. Moreover, one can make a binding promise to another (e.g., one's mother) that one will take care of oneself (e.g., quit smoking).

Third, there is reason to doubt that rule-utilitarian and contractualist theories adequately explain the difference, for a "practice" of promises to oneself is at least conceivable and could both promote general utility and appeal to rational

8 Singer, in "Duties to Oneself," suggests that the characteristic function of claims to "obligations to oneself" is to affirm a right to do what, misleadingly, one says one has an obligation to do.

self-interested "contractors."⁹ Whether the practice would in fact have these advantages would depend, of course, not only on its benefits but also on the costs of "making a moral issue" out of what most now regard as morally indifferent. But this is an empirical question, and we may well wonder whether the moral status of promises to oneself really depends on this empirical question. Suppose that, on balance, a practice of promises to oneself (taught, sanctioned by peer pressure, etc.) promotes general utility and appeals to rational self-interested (e.g., Rawlsian) contractors simply *because* it serves the interests of those who use it or, more generally, because it is a device (like "behavior modification" therapy) that enables them more effectively to pursue their goals, whatever these may be. Then rule-utilitarians, it seems, would have to conclude that promises to oneself would be morally binding. But would they be? Would they, special cases aside, give us anything more than hypothetical imperatives?

These doubts naturally raise further doubts about an underlying assumption of the rule-utilitarian and (Rawlsian) contractualist treatment of promises, an idea also taken for granted in my discussion here: namely, that promising should be construed as a "practice," defined by conventional rules that we have some independent moral ground to accept. If we viewed the personal commitment expressed in promising in a different way, the comparison of promises to others and promises to oneself might yield quite different results.

Turning now to special cases, recall that so far I have only argued that we do not regard persons as *in general* morally bound by their promises to themselves. But there are certain cases in which it is hard to deny the moral significance of such promises. Moreover, I suspect that the temptation to view promises to oneself as morally binding stems from concentration on these special cases, just as the facile dismissal of promises to oneself results from focus on a different range of examples (e.g., promising myself a treat). To illustrate, consider the following. The once honored sheriff, humiliated by one act of cowardice, has become the town drunk, despising himself and ridiculed by all. When he hears that the outlaws from whom he ran are now coming back to town, he "promises himself" most solemnly to dry out and face them with courage. He will not do it for others, he thinks, but for himself, as one last chance to be true to himself. He deliberately places his last shred of self-respect on the line, fully accepting that if he makes and then breaks this "promise" he will be rightly condemned by himself and others more than if he simply skips town at once.

This story is a fiction from a romantic Western film, but similar scenarios closer to home are not hard to imagine. In these cases there is some morally

9 For rule-utilitarianism see John Rawls' "Two Concepts of Rules," *Philosophical Review* 64 (1955): 3-32 and Richard Brandt, "Towards a Credible Form of Rule-Utilitarianism," in George Nakhnikian and H. Castañeda, eds., *Morality and the Language of Conduct* (Detroit, Mich.: Wayne State University Press, 1963) pp. 102-43. A contractualist position on promising is represented by John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971) pp. 342ff.