



---

Agreements, Coercion, and Obligation

Author(s): Margaret Gilbert

Source: *Ethics*, Vol. 103, No. 4 (Jul., 1993), pp. 679-706

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/2381633>

Accessed: 23/11/2010 18:48

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ucpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

# Agreements, Coercion, and Obligation\*

Margaret Gilbert

## I. INTRODUCTION

Agreements and kindred phenomena are ubiquitous in human life.<sup>1</sup> They are a central source of obligation. Obligation is a forceful species of reason for action. It is therefore important to be clear about the nature of agreements and the obligations of agreement.

These matters have not as yet been well understood. Or so a number of common pronouncements suggest. Among these are some prevalent claims about the impact of coercion on agreements. Philosophers standardly claim that coerced agreements do not impose obligations to fulfill the agreement: they are not binding.<sup>2</sup> Many assert that

\* The first draft of this article was written when I was a Visiting Fellow of Wolfson College, Oxford University (1989–90). That work was supported by an American Council of Learned Societies Fellowship for Research and by a sabbatical grant from the University of Connecticut (1989–90). Versions of this material have been presented at the Political Theory conference held at New College, Oxford, January 1990, and at York University; the University of Connecticut, Storrs; Queens College, CUNY; M.I.T.; and Princeton University. I am grateful for all of the comments I have received. Special thanks to John Horton, Chris Megone, John Troyer, and Frank Stewart for lengthy comments on written material, and to John Deigh, Christine Korsgaard, and Saul Kripke for extended discussion. Responsibility for the ideas expressed here is mine alone.

1. I argue that paradigmatic social phenomena such as social conventions have an underlying structure analogous to that of agreements in *On Social Facts* (London: Routledge, 1989; Princeton, N.J.: Princeton University Press, 1992). See also my papers “Modelling Collective Belief,” *Synthese* 73 (1987): 185–201, “Walking Together: A Paradigmatic Social Phenomenon,” in *The Philosophy of the Human Sciences*, ed. T. Uehling, Midwest Studies in Philosophy, vol. 25 (Notre Dame, Ind.: University of Notre Dame Press, 1990), pp. 1–14, and elsewhere.

2. See, e.g., L. C. Becker, “Hard Choices Are Not Enough,” *Virginia Law Review* 67 (1981): 97–102, p. 100; H. Beran, “In Defense of the Consent Theory of Political Obligation and Authority,” *Ethics* 87 (1977): 267; G. Kavka, *Hobbes* (Princeton, N.J.: Princeton University Press, 1986), p. 396 (not ‘morally binding’); A. J. Simmons, *Moral Principles and Political Obligations* (Princeton, N.J.: Princeton University Press, 1979), p. 82 (with specific reference to ‘consent’), and “Consent, Free Choice, and Democratic Government,” *Georgia Law Review* 18 (1984): 791–819, pp. 812, 816; M. Walzer, *Obligations: Essays on Disobedience, War, and Citizenship* (Cambridge, Mass.: Harvard University Press, 1970), pp. xii–xiv; A. D. Woozley, *Law and Obedience: The Arguments of Plato’s Crito* (Chapel Hill: University of North Carolina Press, 1979), p. 104.

*Ethics* 103 (July 1993): 679–706

© 1993 by The University of Chicago. All rights reserved. 0014-1704/93/0304-8946\$01.00

coerced "agreements" are not agreements at all.<sup>3</sup> In my view both of these claims are false.

This article offers an account of agreements and the obligations that flow from them. It brings this account to bear on the question of the possibility and bindingness of coerced agreements.

Section II is a largely critical section that addresses two common arguments for the impossibility of coerced agreements, questioning them from a pretheoretical standpoint. In Section III, I present my positive thesis on the nature of agreements and their obligations. Although I believe that the impact of coercion on a promise is essentially the same as its impact on an agreement, I reject the standard view of an agreement as an exchange of promises. After explaining why I reject that view, I propose that typical everyday agreements are "joint decisions." I articulate the joint decision model of agreements in detail, arguing that obligations of a distinctive sort flow from every joint decision. This section of the article is essentially self-contained with respect to the preceding and following sections. Working with the joint decision model of agreements, I argue in Section IV that coerced agreements are possible and that they all carry with them obligations of conformity. A number of implications of this claim are explored.

My conclusions bear on the interpretation and assessment of "actual contract" theories of political obligation, that is, theories supposing there to be widespread "political obligations" that derive from actual (as opposed to hypothetical) agreements. I touch briefly on the topic of political obligation in Sections II and IV.

I should stress that I am here concerned with the notion of agreement expressed in informal everyday transactions, as opposed to any concept of contract in law.<sup>4</sup> Accordingly, the type of obligation at issue in this article is something other than legal obligation.

## II. TWO NO-AGREEMENT ARGUMENTS

### *The Obligation Argument*

The claim that coerced agreements are impossible is often defended with something like the following 'obligation argument':

1. Any genuine agreement generates a moral obligation to abide by it: this is a conceptual matter.
2. If an apparent agreement is made in the face of coercion, there is then no moral obligation to abide by it.

3. See, e.g., Kavka; J. Raz, "Authority and Consent," *Virginia Law Review* 67 (1981): 103-31, p. 126 (with specific reference to 'consent'); Simmons, *Moral Principles and Political Obligations*, p. 82 ('consent'), and "Consent, Free Choice, and Democratic Government," p. 816.

4. That is not to say that legal judgments have no bearing on the discussion, or vice versa. Some legal judgments are referred to below, but I attempt no systematic consideration of legal discussions of contract in this article.

Therefore:

3. A genuine agreement cannot be made in the face of coercion.<sup>5</sup>

In well-known writings on political obligation A. J. Simmons has suggested that the conclusion of the obligation argument can be extended. Consider those who appear to agree to remain in a state and accept burdensome duties of citizenship (such as military service) when their only option is emigration. Perhaps such people cannot be said to be *coerced* into agreeing. However, any so-called agreements exacted in this context are *unconscionable*, and so not morally binding. They are, indeed, not agreements at all.<sup>6</sup>

This is a strong conclusion in relation to actual contract theories of political obligation. According to Simmons, it would make no difference if agreements of the relevant kind appeared to be widespread. No such agreements are possible in standard political circumstances.

I shall not consider here whether Simmons has properly extended the conclusion of the obligation argument. Suffice it to say that the argument may extend beyond the context of coercion strictly speaking.

#### *Interpreting the Obligation Argument*

The argument is not entirely perspicuous as presented above. Let us say that one performs an “act of agreement” when what one does is sufficient from a behavioral point of view to constitute entry into an agreement, provided that any other relevant conditions are satisfied. Thus one might nod one’s head emphatically, say “Sure” or “Fine” or “I will indeed,” and so on, in response to some proposal by one or more others. Let us say that an “apparent agreement” occurs when two or more people perform matching acts of agreement, and this is common knowledge. (‘Common knowledge’ is a frequently used technical term, coined by David Lewis. When something is common knowledge between persons A and B, then, roughly, it is “out in the open” between them.)<sup>7</sup> I take it that all genuine agreements are at the same time apparent agreements in the sense just defined.

Let us now construe the argument’s first premise as follows:

- 1A. An apparent agreement is a genuine agreement only if, when it is taken together with its background circumstances, there is a moral obligation to abide by it: this is a conceptual matter.

5. There is a similar common argument about promises. I shall not explore this directly here.

6. Simmons, “Consent, Free Choice, and Democratic Government,” esp. pp. 809–17, and also *Moral Principles and Political Obligations*, pp. 99–100.

7. D. K. Lewis, *Convention* (Cambridge, Mass.: Harvard University Press, 1969). There is no need to fix on a precise definition for ‘common knowledge’ here. I discuss this matter further in *On Social Facts* (see esp. pp. 188–95).

The rest of the argument proceeds as before:

2. If an apparent agreement is made in the face of coercion there is then no moral obligation to abide by it.

Therefore:

3. An apparent agreement cannot be a genuine agreement if it is made in the face of coercion.

Or, in short:

A genuine agreement cannot be made in the face of coercion.

Given the above construal of the first premise, the obligation argument appears to be valid. It is nonetheless quite problematic, as I shall now argue. A number of other construals fare no better (see the section on revisiting the obligation argument in Sec. III below).

#### *The Knowledge of Agreements Assumption*

In advance of theorizing about agreements, something like the following 'knowledge of agreements' assumption seems plausible: if someone enters into an agreement, then she knows that she does. I should stress that this assumption concerns informal agreements as these are conceived of in everyday life, as opposed to contract in some particular legal system.<sup>8</sup> Without considering it further for now, let us consider its consequences for the obligation argument.

According to the first premise of that argument, whether or not one has entered into an agreement is a matter of the moral impact of one's circumstances. The argument assumes that coercion is *one* circumstance that prevents an apparent agreement from having the appropriate moral impact.

If we endorse the premises of the obligation argument, then, we apparently have to accept that in order to know that one is entering into an agreement, one must make what may in its context be quite a delicate judgment about the moral impact of one's circumstances. Thus one must know whether one is faced with the moral equivalent of coercion or, rather, with a form of strong pressure with different moral consequences. Given the knowledge of agreements assumption, we must accept that in order merely to enter into an agreement one must always accurately judge the moral impact of one's circumstances, however delicate a judgment that is in its context.

I find this implausible. Entering into an agreement and knowing that one does would seem to be a relatively simple business, possible

8. Thus claims such as the following would not contradict the assumption: "By getting on a bus one concludes, whether or not one knows it, a legally binding contract" (J. Finnis, *Natural Law and Natural Rights* [Oxford: Clarendon, 1980], p. 299).

even for creatures unable to make delicate judgments about the overall moral impact of their circumstances.

*Agreements: A Rough Sketch*

How firm is the knowledge of agreements assumption? Can we give it an articulate basis?

Consider the formation of a typical informal agreement. Mike says to Jane: "What do you say to my taking Fido for a walk at three and your feeding him at five?" Jane replies: "That's fine."

On the face of it, something like this seems to have happened.

CONDITION 1. Mike intentionally expressed to Jane his willingness to accept that he is to walk Fido at three and she is to feed him at five.

CONDITION 2. Jane did likewise.

These may be thought of as the "core" conditions. We may also add:

CONDITION 3. The fulfillment of the above conditions was common knowledge between Jane and Mike.

I shall introduce some important refinements in this rough account later.

If these or similar conditions capture the essence of this case, that would help to explain any initial plausibility of the knowledge of agreements assumption. It is already part of the full set of conditions that the fulfillment of the core conditions is common knowledge between Jane and Mike: it is out in the open between them. If something is out in the open between two people, then we can expect that each of them will know this is so.

The conditions roughly sketched also support the idea that agreement makers need have no understanding of the moral impact of various circumstances such as coercion. As to what it is intentionally to express one's *willingness* to accept something, I use this term to indicate only that a state of the agent's will is at issue or, more specifically, a positive inclination of the agent's will. The terms 'readiness' or 'preparedness' might have been used instead. Most to the present point, I take the state in question to be something of which the agent can be fully aware at the time he expresses it without reference to the circumstances leading up to it (such as coercion) or their moral consequences.

In sum: the premises of the obligation argument coupled with the knowledge of agreements assumption imply that some agreement makers must be able to make moral judgments of considerable sophistication. A rough sketch of the conditions that appear to produce a typical informal agreement both supports the knowledge of agreements assumption and confirms the independently plausible idea that

agreement makers never require the capacity at issue. These considerations suggest that something is wrong with the premises of the obligation argument.

### *The Implications of Practice*

The conclusion of the obligation argument would hardly matter if attempts to force another person to agree were never made. However, the existence of any such attempts makes that conclusion suspect. Further, we seem able to allow that such attempts succeed: certainly someone could say "He forced me to agree" without linguistic oddity.

Proponents of the obligation argument might contend that the sentence cited must always mean something like "He forced me to go through the motions of agreement," with the implication that "I did not agree in fact."

It is not clear how they could back this up. Consider the following dialogue. "Why on earth did you agree to do such a thing?" "He forced me to—I had no choice." It seems most natural to read *both* parts of this dialogue as presupposing that a genuine agreement took place.

Of course the second speaker could be misguided. It may be that her interlocutor could truthfully reply, "Then you didn't agree at all! (Thank goodness!)" But that the second speaker would be misguided cannot be assumed here: it is what is at issue.

Our usage is *prima facie* evidence against the obligation argument. In our pretheoretical judgments we do seem to allow the conceptual possibility of coerced agreements. Some argument must be found to show that, if we do this, we are somehow confused or misguided. And the argument must be a powerful one.

### *The Voluntariness Argument*

What of the quite common voluntariness argument? This runs as follows:

1. Agreements are by definition voluntary: they cannot come about against the will of either party.
2. Coercion does not allow an agreement to be voluntary.
3. Therefore, coercion rules out the very possibility of agreement.

This argument may gain a spurious plausibility from an important ambiguity in the notion of voluntariness or of acting in conformity with one's will. Suppose that Betty has decided not to buy a certain house. However, a crafty trickster gets her to sign her name at the bottom of an agreement to purchase it without realizing what she is doing. Perhaps he covers the agreement in such a way that she thinks she is simply giving him her autograph. We might say that she signed the agreement against her will in what I shall call the "decision-for"

sense: she never made any decision in favor of signing the agreement in question.<sup>9</sup>

In contrast, if someone has a gun at Betty's head she may well decide in favor of signing the agreement. She may not only go through the motions of signing it but also have the intention of signing it. Then, though her signing it may be against her will in an important sense, or not be voluntary in an important sense, it will not be against her will in the decision-for sense. It seems, however, that only if Betty goes through the motions of signing against her will in the decision-for sense could she be said not to have signed the agreement.

Let us return to the voluntariness argument. The first premise appears to be uncontentious if and only if 'against the will of either party' is interpreted in the decision-for sense. The rough conditions on agreement proposed earlier suggest that one who agrees must indeed decide in favor of something, in the foregoing technical sense. Let us ascribe that interpretation to the first premise, then.

The second premise appears to be false when it is interpreted accordingly. It seems to be false both if construed as a logical claim (as is presumably intended) and if construed as a claim of fact. As the above example indicates, one's rational choice in the face of coercion may well be to decide in favor of a course of action, or to agree to something.<sup>10</sup> Nor is it plausible to suppose that everyone becomes so unnerved in the face of coercion that they are incapable of making up their minds at all, that they are, so to speak, rendered witless. Some people may be so affected, but it is too bold to presume that all will be. There is surely no conceptual warrant for such a presumption. Compare Lord Scarman: "The classic case of duress is . . . not the lack of will to submit but the victim's intentional submission arising from the realization that there is no other practical choice open to him."<sup>11</sup>

9. The phrases 'decision for' and 'decision in favor' may strike a more deliberative note than is apposite to make my point. I am assuming that if one acts with the intention of doing X, then this incorporates a "decision-for" X in the relevant technical sense: it incorporates a positive inclination of the will toward X. With this caveat, I shall stick with the phrase 'decision-for', which I feel emphasizes the involvement of the person's will in the process, whereas 'intention' seems rather to emphasize the understanding or cognitive component.

10. It is sometimes rational to do more than *feign* agreement, given that such feigning is possible. For example, the other party may have a nose for deceptions. See also n. 27 below.

11. Quoted in A. G. Guest, ed., *Anson's Law of Contract*, 26th ed. (Oxford: Clarendon, 1984), p. 242. Some earlier judgments, condemned by Anson/Guest as "fallacious," argued that in "a successful plea of duress . . . the party affected must show that the compulsion was such as to vitiate his consent, to deprive him of any *animus contrahendi*" (ibid.). In these legal discussions, then, we seem to find an oscillation between a sense that a successful plea of duress is actually a plea of "nonagreement" and



Interpreted as I have suggested, then, the voluntariness argument must be rejected. The first premise uses the decision-for sense of voluntary. If this is the sense used by the second premise, that premise is false. Otherwise, the argument trades unacceptably on an ambiguity.

In the paper cited earlier, Simmons in effect proposes the following argument: agreements must be voluntary in order to be morally binding; agreements are necessarily morally binding; so agreements must be voluntary.<sup>12</sup> He believes that coercion rules out the type of voluntariness at issue. If he is right, then this type of voluntariness is not voluntariness in the decision-for sense.

If, in fact, agreements cannot be morally binding unless they are voluntary in some sense other than the decision-for sense, then it seems that—*pace* Simmons and many others—agreements do not have to be morally binding. For on the face of it there is nothing about agreements that entails that they must be voluntary in any sense other than the decision-for sense. Some argument must be given for the claim that they *must* be voluntary in some other sense. I know of no such argument.

### III. AGREEMENT AND OBLIGATION

#### *Obligation in General*

Now the common assumption that there is a conceptual connection between agreements and obligation of some kind seems to me correct: there *is* an important sense in which agreements generate obligations of conformity. As I shall argue, the obligations in question can be incurred in the face of coercion. I must first say more about agreements and their obligations.

Following certain pointers from common usage, let me start with some firm points about obligation in general.

It is clear that if someone, S, has an obligation to do A, then (at the least) S has a reason to do A. More precisely, an obligation is a *reason for acting* in the following sense: once you have such a reason, rationality requires that you act in accordance with it, all else being equal.

Now normally we do not count someone as having an obligation whenever they have a reason for action. Here are two cases in point.

First case: Alice is trying to win a race, and the only way she can do this is by running faster. She would now appear to have a reason

---

a sense that even if there is an agreement under duress, it should be “revocable” in law (Scarman). The main issue there, of course, is the determination of proper legal judgments. It may well be that what is an agreement intuitively speaking should sometimes be considered void, or at least voidable or revocable in law. The law has its own reasons, which have to do, among other things, with fairness and right (as well as history and tradition).

12. See Simmons, “Consent, Free Choice, and Democratic Government,” p. 810.

to run faster. She may well say to herself, "I've *got* to run faster!" At the same time, it does not seem right to say that she now has an obligation to run faster. Second case: Molly decides to eat lunch in the factory cafeteria and has not at this point reconsidered and changed her mind. She would now appear to have a reason to eat lunch in the cafeteria. At the same time, it does not seem right to say that she now has an obligation to eat there.<sup>13</sup>

If an obligation is more than a reason for action, what more might it be? The foregoing discussion suggests a relatively precise negative characterization.

There is something special about a person's aims and decisions. All that is needed for their creation is the relevant act of that person's mind or will: the adopting of an aim or the making of a decision. This, too, is all that is needed to destroy them. The formation of an aim or decision may be preceded by deliberation, but it may not. It may have reasons behind it, or it may not. Nonetheless, once someone has taken up an aim or made a decision, she has a reason for acting. That reason disappears if she gives up the aim or rescinds the decision.

Let us say that S has a reason for acting that is *the creature of her own will* if the reason has this property: all that was needed for its creation and all that is needed for its destruction is an arbitrary act of will on S's part. It seems that we do not call a reason an obligation if it is the creature of a person's will in this sense.

This thought is sustained when we look at a typical example of the use of the term 'obligation'. Suppose that Sally is out for a walk when a child trips in front of her and breaks its arm. There is no one else in sight. It may well be said that in these circumstances Sally has an obligation immediately to help the child.<sup>14</sup> Precisely what the implied ground of obligation is in this case is a moot question among theorists.

13. That a decision provides a reason at all may appear to need arguing. It is implicit, however, in the way we think about decisions. Suppose that, having decided to eat in the factory cafeteria, and without first changing her mind, Molly starts in the direction of a local restaurant. We would surely judge that (all else being equal) something is amiss. More precisely, we would judge that Molly had a reason not to act in this way, a reason her decision provided. It may well be that there are importantly different types of reason, such that decisions and the like generate reasons of a special type. In *Intention, Plans, and Practical Reason* (Cambridge, Mass.: Harvard University Press, 1987), Michael Bratman appears to argue along these lines, decisions and the like being what he calls 'framework reasons'. For present purposes, such distinctions need not concern us. The main point here is that from an intuitive point of view we regard decisions as reasons for acting in the broad sense at issue here. (I say more about some aspects of decisions below.)

14. Some philosophers have asserted that use of the term 'obligation' in this type of case is improperly loose. However, the term is often so used in both everyday and philosophical talk. See R. B. Brandt, "The Concepts of Obligation and Duty," *Mind* 73 (1965): 374–93, p. 387.

But it seems clear enough that the obligation is not perceived as the creature of Sally's will in the sense defined above.

We have arrived at an apparent necessary condition on obligation that is relatively stringent: *S has an obligation to do A* only if S has a reason to do A that is not the creature of her own will (where this is understood in accordance with the above discussion). This gives us a partial characterization of obligation in general that accords with the plausible thought that the obligations that I have are not had solely "at my pleasure."

Now, decisions do not generate obligations, but they do generate reasons. It will be useful at this point briefly to consider how this can be so.

Decisions appear to generate reasons in a way which is independent of their content, their context, and their consequences. It doesn't matter what my decision is, or why it was made. As long as it is my decision it generates a reason for me.

How precisely it does this depends on what precisely a decision is and, indeed, on what a reason is. These are delicate and important matters which must to a large extent be waived here. However there is at least one thing that can always be said in favor of the act in question, once the decision is made, and this may be crucial to the understanding that decisions generate reasons.

We understand that whatever else is true of it the act decided upon at least conforms to the agent's decision. It can then be argued that simply by deciding, I set up a situation in which a particular act is rendered the act it is appropriate for me to perform, all else being equal. It will also be inappropriate for me to accept new aims which conflict with my decisions, insofar as I have not yet changed my mind and rescinded my decision. The "appropriateness" in question is the appropriateness of simply being consistent: if my acts match my decisions I am at least being consistent. Such consistency has something "good" about it intuitively. This type of goodness is, equally evidently, only one aspect of goodness in general.

In my view agreements are importantly analogous to decisions: they too provide reasons that are independent of their content, context, and consequences. Unlike decisions, however, they generate obligations. To argue this I must say more about what I take an agreement to be.

### *Approaching Agreement*

The standard proposal about agreements is that an agreement is, in effect, an exchange of promises.<sup>15</sup> In my view this is not so. I discuss

15. A similar view is found in accounts of contract (or certain types of contract) in common law. That does not of course clinch anything with respect to the proper analysis of vernacular notions.

this in detail elsewhere.<sup>16</sup> It will be useful to look at it briefly here before presenting an alternative account of agreement.

Let us focus on Mike and Jane's agreement that he will walk Fido at three and she will feed him at five. Mike says "What do you say to my taking Fido for a walk at three and your feeding him at five?" and Jane replies, "That's fine." I shall consider for the sake of argument the simple suggestion that an agreement will have taken place if and only if the words of Mike and Jane were implicitly understood as a promise exchange as follows: Mike: "I promise to walk Fido at three." Jane (in reply): "I promise to feed him at five." Here each person promises the other to perform some act. I shall be taking these promises quite literally, in particular, each will be construed as making an unconditional or absolute promise. This is simply to fix the nature of the proposal I am focusing on.

There are two aspects of this exchange which make it inadequate as a representation of the agreement. First, as soon as Mike has promised to walk Fido at three, he has an obligation to do so. Jane has not yet incurred an obligation to feed Fido. This seems wrong: the parties' obligations surely arise simultaneously from their agreement.

More crucially, once both promises are made, Mike's obligation is logically independent of Jane's obligation, and vice versa. Even if Mike breaks his promise, Jane's promise—and hence her obligation to feed Fido at five—remains intact.<sup>17</sup> This too seems wrong: the obligations of the agreement appear to be (as I shall put it) *interdependent*. Among other things this means that if the agreement is broken by either party, neither obligation remains. Putting it with maximum generality: interdependent obligations cannot exist one without the other. Their independent existence is logically impossible.

We must therefore reject the suggestion that when we say a person 'breaks' an agreement we are supposing that an agreement is a pair of promises, such that if one of the pair is broken the agreement as a whole is said to be broken. To clarify the problem, let us first define

16. In "Is an Agreement an Exchange of Promises?" (presented as part of longer talks at the Political Theory conference held at New College, Oxford, January 1990; at the Philosophical Society, Oxford University, May 1990; at Massachusetts Institute of Technology, December 4, 1992; at Princeton University, December 11, 1992; and elsewhere). Here and in that paper my discussion expands on some comments in *On Social Facts* (see p. 380).

17. In general terms, if a person X promises to do a certain act A on a certain occasion, no one else can break X's promise. X alone can break it, by not doing A when the occasion arises. The same goes for conditional promises. Suppose X promises that he will do A on a certain occasion, provided only that a second person Y performs act B on some prior occasion. Y still cannot break X's promise. If Y fails to do B when it is called for, X no longer has any option of performing his promise. But it has not been broken. Indeed, it never can be broken, now, since X will never have the opportunity not to perform when his performance is called for.

a new term, 'shmagreement', as follows: when promises are exchanged there is a shmagreement. Let us say, further, that if one of the promises is broken the shmagreement is "broken." Let us now suppose that Mike and Jane have exchanged their promises and, hence, have entered a schmagreement. Mike breaks his promise, and hence the schmagreement is broken (in our stipulative sense). Both Mike's promise, and the schmagreement between Mike and Jane, are at an end. Yet Jane's *promise* is still unbroken. So she still has an obligation to act as she promised. I take it that if we are to capture the structure of our illustrative agreement, we cannot be left with such a dangling obligation. It seems, then, when we speak of the breaking of this agreement we mean something other than the breaking of one of an "exchanged" pair of promises.

Might one say that the agreement is a pair of promises such that we understand that if one of the promises is broken then (a) the agreement is broken, and (b) the other promise is broken too? What kind of pair of promises, though, is this? A promise which is broken given that another promise is broken is not so much a promise as . . . something else. (See the next section.)

In sum, I take it that the obligations arising from Mike and Jane's agreement are simultaneous and interdependent. I argue elsewhere that no possible exchange of promises by one person to another, including exchanges involving conditional promises, can meet both criteria.<sup>18</sup>

In the next section I present a new model of agreements which does meet these criteria. I shall here waive the question whether some forms of interaction which do not meet the criteria can still appropriately be referred to as 'agreements'. In particular I shall waive this question with respect to exchanges that are strictly and literally exchanges of promises. I shall go on to argue that agreements according

18. The notion of a "conditional" promise is not unproblematic. The "condition" may be read as internal or as external to the promise. Here is a pair of "promises," one of which is (externally) conditional. Mike: "On condition that you promise to do A, I promise this: I will do B." Jane: "I promise to do A." Depending on how this is construed, simultaneous obligations may ensue. These obligations are independent, however. Thus the interdependence criterion is not met. Given our example agreement both parties have unconditional obligations to perform the acts promised. An exchange of (internally) conditional promises of the following form does not, therefore, capture what transpires: Mike: "I promise this: I will do B, if you do A." Jane: "I promise this: I will do A, if you do B." This exchange results in each promisor's having an (independent) obligation to perform a certain action provided that the other performs a certain act. Suppose it is common knowledge that this is the situation. So far neither promisor has any reason actually to *do* anything. Each will only have reason to act given a reason to believe the other will act. Nothing about what each knows so far gives either one such a reason. In contrast, given unconditional obligations, each immediately has a reason to act.

to the model I present are possible in the face of coercion and that any such coerced agreements will be binding. I believe that analogous arguments can be made for single promises and thus for promises that are part of an 'exchange'. I shall not attempt to argue that here. I should stress, however, that the discussion in this section was in no way intended to imply that if the exchange of promise model of agreements were accepted the conclusion that coerced agreements are impossible or nonbinding would be justified. It was intended only to show that the standard model of agreements is not adequate to capture the structure of obligations that characterizes typical everyday agreements and, thus, to prepare the way for the introduction of an alternative model.

### *Agreements as Joint Decisions*

I shall now sketch a different model of agreements, developing the idea adumbrated in *On Social Facts* that an agreement is, in effect, a joint decision.<sup>19</sup> That is, the achievement of those who enter into an agreement is analogous to that of a single person who makes a personal decision; in this case, however, the achievement is a joint one. Precisely how this joint achievement is constituted needs some explanation.

If someone comes to a personal decision on a course of action, then, roughly, she comes personally to accept that she is to do such and such. As I understand it, the parties to a joint decision "jointly accept" that such and such is to be done by one or more of them. Joint acceptance, in the relevant sense, requires what I call a 'joint commitment'.<sup>20</sup> There must be a joint commitment jointly to accept that such and such or, alternatively, to accept that such and such as a body. When there is a joint commitment between two or more parties, there is what I call a 'plural subject' or a (collective) 'we'.<sup>21</sup>

Before saying more about joint commitment, I address a query about describing an agreement as (in effect) a joint decision. Evidently our language allows us to speak both of "our agreement" and of "our decision." Given my analysis, the phenomenon referred to by these different phrases is essentially the same. Why, then, would we allow ourselves the two ways of talking when one might have sufficed?

There could well be a core phenomenon, a particular achievement, that is identical in each case. Meanwhile, we may speak of our decision when we mean to suggest the existence of one particular type of background circumstance, and we may speak of our agreement

19. *On Social Facts*, pp. 380–81.

20. See *ibid.*, chap. 4 and *passim*, esp. pp. 198, 382, and elsewhere.

21. For more on plural subjects, see *ibid.*, esp. chap. 4 and subsequent chapters, as well as elsewhere. I introduced the relevant notion of joint acceptance in "Modelling Collective Belief."

when we mean to suggest another. (Whether or not this has to do with the semantics of the terms, or something else, need not concern us here.)

I conjecture, more specifically, that we speak of a decision when we wish to suggest that before entering the decision-making process the parties already formed a we or plural subject. Perhaps they were committee members, or the members of a walking party. This usage may indeed suggest that in arriving at their decision, the parties were engaged not so much as separate individuals but, rather, as members of some relevant group or collective, with a single, jointly accepted agenda. We decide collectively. We speak of an agreement, meanwhile, when we wish to make no such suggestion. The parties to an agreement are viewed as separate individuals with (possibly distinct) personal agendas. We individually agree. Nonetheless it is true that, whatever our previous relationship, by entering an agreement we thereby constitute ourselves the members of a (collective) we.<sup>22</sup> In other words, we achieve the same immediate result as those who collectively decide: we come jointly to accept that such and such is to be done by particular people.<sup>23</sup>

I now return to the concept of joint commitment. In *On Social Facts* I argued that this concept lies at the core of a family of everyday concepts, our "collectivity" concepts. These include the concepts of a collective goal, a collective belief, and a social convention. I shall summarize those aspects of my discussion that are relevant here.

In order for two or more people to become parties to a joint commitment each must express to the other his or her willingness to do so, in conditions of common knowledge.<sup>24</sup> As all understand, the

22. Is it possible in principle for people to make an agreement without previously forming a plural subject? Insofar as "mutual recognition" on meeting in person itself creates a plural subject (see *On Social Facts*, pp. 217–18), this may be difficult. Perhaps this is an example however: using electronic mail, I send out my part of an agreement to someone I know only by name. "Would you answer all my e-mail messages if I answered all of yours?" Were she to reply "Yes," we could have an agreement, it seems, without any prior 'we' formation between us. Our only communication would be the communication of each one's individual willingness to form the agreement.

23. The "individualistic" nature of our focus in talking of agreement may be what has led many to assume that nothing of a contractual nature could lie at the basis of society. What those who agree *de novo* need, however, is only the power jointly to commit themselves in the area in question. They need in no other way be "asocial" or "atomistically conceived" either by themselves or anyone else. Those sympathetic to a contractual account of social groups are presumably sensitive to the "holistic" nature of what is achieved by agreements (i.e., of a joint commitment), whether or not they have clearly grasped what this achievement amounts to. See *On Social Facts*, chap. 7, and "Walking Together," for some references to contract theory, Rousseau, and Hobbes in this connection. I hope to treat this matter at greater length elsewhere.

24. Compare *On Social Facts*, pp. 197–98, 408–9, and elsewhere.

joint commitment is in place when and only when all of the necessary expressions have been made.<sup>25</sup>

Of course the parties will not generally have the (technical) phrase 'joint commitment' at their disposal. Nor need they be able to give an articulated analysis of the concept of joint commitment. They must possess that concept, however, in order to be parties to a joint commitment.

As the parties to a joint commitment understand, they are individually committed in the sense that each individually has a commitment. Nonetheless, these commitments are seen to flow from the joint commitment. In this respect the joint commitment of two people may be likened to a string that they hold taut between them. Neither can hold a taut string by this method unless the other does: neither is or can be individually committed through the joint commitment unless the other is.

A joint commitment is "joint" in a strong sense: no individual is committed until all the others are; it is impossible to rescind the commitment for one party without rescinding it for the others; if the commitment is broken, it is broken for all: it does not remain to require any action of anyone. In short, it cannot exist to affect one party unless it also affects the others. The parties' 'individual' commitments are therefore interdependent with respect to their generation and their persistence.

Mutual consent is required to rescind a joint commitment. The commitment produces, in effect, a single subject with its own commitment: we are committed unless we (together) change our mind. It may be true that one party can break a joint commitment with the result that it is henceforth no longer there to impose constraints on anyone. But then the constraints it creates exist in full force right up to the break (a break requires something to be broken).

Suppose, then, that Jack is party to a joint commitment. As he understands, the source of his own ensuing commitment is not an act of his own mind alone and he is not in a position to rescind that commitment unilaterally.

Jack's participation in the joint commitment immediately gives him reasons for acting. *Any* commitment on someone's part sets up a

25. What precisely needs to be expressed is a somewhat delicate matter. Each party must evidently be "ready for joint commitment" as far as his will is concerned; yet (as he understands) only once a certain condition is fulfilled can he be committed as a function of the intended joint commitment. Thus one who says "Shall we dance?" indicates that, in effect, he will be committed to dancing if (and only if) his question is answered appropriately. This commitment to dance will be a function of the joint commitment then (and only then) created. One relatively crude way of describing what needs to be expressed is (as I have sometimes put it) a "conditional commitment." This description needs to be understood as indicated here.



situation in which certain acts are appropriate for him and others are not. For some acts conform to his commitment and others do not.

Jack's participation in the joint commitment immediately gives him reasons for acting of a special kind, a kind that I previously had occasion to characterize in discussing the nature of obligation. These reasons are not the "creatures of Jack's own will" like the reasons that flow from his personal decisions or aims. In other words, unlike the reasons flowing from his personal decisions, these fulfill the necessary condition on obligation arrived at above.

As we have seen, even Jack's own "part" of the joint commitment required the act of another's will to bring it into being. As he understands, he is committed through it, unless and until the other parties concur in its dissolution.

Thus the reasons for action that a joint commitment produces for each participant cannot be created or destroyed by an act of that participant's will alone, as is the case with the reasons flowing from personal aims and decisions. The wills of one or more others are an essential part of the picture.

Insofar as a personal decision locks you into a course of action, you yourself have the sole key needed to turn the lock. In order to unlock yourself all you need to do is to change your mind: to rescind your decision. In contrast, insofar as a joint commitment locks you into a course of action, at least two keys are required to turn the lock. You have only one of these keys. Each of the other parties has another. Changing your own mind is not enough; all must concur.<sup>26</sup>

Changing the metaphor, one can say that the parties to a joint commitment are tied to one another or bound together. One who violates the commitment by acting inappropriately breaks that bond. Personal decisions as such involve no such interpersonal bond. One can act against one's own unrescinded decision without ignoring another's legitimate interest in the matter. What is at issue in the case of conformity to a *joint* commitment is not simply self-consistency. It is the maintenance of an acknowledged interpersonal tie or bond.

I propose that it is appropriate to speak of joint commitments as producing not just reasons for action but obligations. The word 'obligation' comes, after all, from the Latin *ligare*, to bind.

More precisely, I suggest that a central use of the term 'obligation' is to refer to a reason for acting such that he who has it (whom we may call the 'obligee') is beholden to another or others (the 'obligors') for eradication of that reason by virtue of the participation of both obligee and obligors in a joint commitment. If this is right, then any attempt to give a generic account of the everyday concept of obligation

26. John Deigh suggested the metaphor of lock and key(s) (personal communication, December 1992).

must take account of what I shall henceforth refer to as the obligations of joint commitment. I say more about an apparently distinct class of obligations shortly. But first I must come back to our everyday concept of an agreement.

I conjecture that the intuitive judgment that agreements generate obligations or are binding is responsive to the ineradicable presence, given the agreement, of obligations of joint commitment. A joint decision involves a joint commitment: a commitment to accept as a body that such and such is to be done by one or more of the parties. The parties to the decision are thus obligated to uphold it: in particular, they must act as it dictates. In addition, they must not act to impede either one's conformity, for instance. These obligations have a single source: the joint commitment. They are arrived at simultaneously when and only when the joint commitment is in place. They are interdependent: they both stand as long as the joint commitment stands; they cease if it is dissolved or broken. Any joint decision, then, involves both simultaneous and interdependent obligations, key features of an agreement.

I can now revise my original sketch of the conditions that produced Mike and Jane's agreement as follows:

CONDITION 1. Mike intentionally expresses to Jane his willingness jointly to accept with Jane that Mike is to walk Fido at three and Jane is to feed him at five.

CONDITION 2. Jane does likewise.

CONDITION 3. The fulfillment of the first two conditions is common knowledge between Mike and Jane.

Given these conditions, Mike and Jane are jointly committed to accepting as a body that Mike is to walk Fido at three and Jane is to feed him at five. As a result, Mike's obligation to walk Fido and Jane's obligation to feed Fido will be simultaneous and interdependent.<sup>27</sup>

These conditions do not refer to a background social institution or practice of agreeing.<sup>28</sup> Nor are the obligations of agreement a func-

27. What of people who feign willingness to enter the relevant joint commitment? Can they avoid incurring the obligations of agreement by such a pretense? (Compare the idea of 'mental reservation' to avoid lying.) Two brief comments must suffice here. First, there is some question whether the very idea of setting out not to incur the relevant commitment is intelligible, at least in the context of standard conventions. One cannot appropriately form just any intention one pleases, regardless of one's own other beliefs and commitments. (I thank Catherine Elgin for relevant discussion [January 1993].) Second, suppose that somehow Adam contrives to have Eve reasonably presume that they have made an agreement. Eve then relies on the presumed agreement between them, to her own detriment. Whether or not he has obligations under an agreement, Adam is *morally* responsible for whatever damage his manipulations produced, all else being equal. See below, in the text, for some discussion of the relevant contrast.

28. I am not alone in thinking that the obligatoriness of agreements can be understood without reference to social practices. For a recent discussion along these lines

tion of such "external" phenomena as expectations of conformity or reliance.<sup>29</sup> On this account, obligation is intrinsic to the given particular agreement—something I take to confirm the account from a pre-theoretical point of view.<sup>30</sup> That it is plausible to see the obligations of agreement as having such a "thin" source as that proposed here is supported by the fact that even the most trivial of agreements will appear to bind the participants in a way outside their personal control.<sup>31</sup>

For present purposes there is no need to investigate the nature of agreements in more detail. It is time to explore some ramifications of the joint decision model.

### *The Obligations of Agreement*

Given the joint decision model of agreements, the obligations of agreement are of the type that derives from joint commitment in general.<sup>32</sup> As I have emphasized, obligation of this sort involves a relationship between persons of a precisely specifiable type. To the person with the obligation (the obligee) there corresponds an obligor or obligors who are instrumental in his having a certain reason for acting and whose concurrence is both necessary and sufficient (given his own concurrence) for the eradication of that reason. The obligors have

---

(though along different lines from this article), see Thomas Scanlon, "Promises and Practices," *Philosophy and Public Affairs* 19 (1990): 199–226, and further references therein.

29. By an "external" phenomenon here I mean one that is not part of or logically derivable from the "conceptual essence" of agreement. It may be unusual, but as long as it is logically possible for there to be an agreement unaccompanied by the phenomenon in question, then it is "external" in the sense at issue.

30. In this article I appeal to several judgments I suspect people would make before accepting any particular articulated theory. There is no space here to argue at any length for the independent plausibility of each of these judgments from a pre-theoretical point of view. I hope to do so in another place. For further comments on these appeals, see Sec. V below.

31. Even a silly agreement fabricated to make this point will have that quality. Try agreeing with a friend that after she leaves the room she will say "Bip" and you will say "Bop." Presumably neither of you particularly cares to have this happen, and neither will base any future action on the other's performance. It is an entirely capricious and whimsical agreement. Nonetheless it seems to have the binding qualities any agreement has.

32. Given that an agreement is as described here, and that it therefore has an ensuing obligation of the kind described, is there something about conforming to any agreement that grounds another obligation to conform, an obligation of the kind, perhaps, that is at issue in the injured child case (see below)? This seems doubtful on the face of it. I shall not attempt to consider it further here. Any such obligation is likely to be a function of the existence of the first kind of obligation—the kind intrinsic to agreements as such.

their status by virtue in their participation with the obligee in an appropriate joint commitment.

The obligation here appears, then, to be of a type different from that in the injured child case. Though Sally's obligation has to do with the child, they do not stand in the precise obligee-obligor relationship referred to above.

That there is an apparently different kind of obligation involved here would perhaps be shown most clearly if we could find a case where we felt someone had an obligation of the relevant kind that had nothing to do with any other person. Possibly we are all obligated to do what we can to preserve endangered species of tree. There are no clear candidate "obligors" of the relevant kind here. Even if there is a supreme being, one might argue that He or She could not release us from *this* requirement in the relevant way.

If the obligations of agreement are the obligations of joint commitment, does this mean that they are not moral obligations? That depends, of course, on what one means by 'moral obligation'. More grandly, it depends on the nature of morality.

It is standardly assumed that the obligation of agreement *is* moral obligation. Usually, however, this is just an unargued assumption.

In this connection it is worth pointing out that a claim may have moral significance, in the sense of helping to decide what you ought to do morally speaking, without itself being a claim of "moral fact."

Thus consider the claim, "Your guests will all go blind if they eat that cheese." Taken at face value, this is a plain assertion of nonmoral fact, a fact about the predictable effects of certain behavior on the body. Yet it clearly has moral significance. Given its truth, you are presumably morally obliged not to serve the cheese to your guests. You could not be counted a virtuous person if you had a tendency to ignore such facts.

This could be the structure of the situation regarding agreements and their obligations. To say that the obligations of agreement were not moral obligations would not be to deny them moral relevance. The above argument presumes, of course, that it is not the case that any fact which is morally significant is therefore itself a moral fact.

R. M. Hare has suggested that to presume a form of obligation not to be moral obligation may be to grant it a status that could lead to uncritical compliance.<sup>33</sup> What seems to be the case with the obligations of agreement? Are they 'absolute' in any sense?

Let us first consider how things seem pretheoretically. On the one hand, we normally allow that it is permissible to break an agreement

33. See R. M. Hare, "Political Obligation," in his *Essays in Political Morality* (Oxford: Oxford University Press, 1989), pp. 8–9.

in certain circumstances, all things considered. For instance, I may break an agreement to attend a party in order to save a life.<sup>34</sup> In describing how the permissible contexts of violation are determined we sometimes refer to our exercise of “judgment.” On the other hand, what our judgment tells us is precisely when we may *violate* or *break* an agreement.

This suggests that the obligation of agreement is viewed as absolute in the following sense. Once it is in place through the existence of an agreement, it remains in full force whatever the further circumstances—only provided that the agreement remains in place.

Thus, if Mike and Jane agree that she is to feed Fido at five, etc., she then has an obligation to feed Fido, as long as the proviso is satisfied. We assume that even when the proviso is satisfied it may still be permissible for Jane not to feed Fido at five, all things considered. We understand, however, that her agreement will then still be “binding” in a central sense: she will have to break the bond. Even where it is true that she has an obligation to break it, all things considered, what she has is an obligation to break an agreement, thereby violating one obligation in the service of another.

Compare this with the following version of the case of the injured child. Sally is out on her walk. At four-thirty P.M., the child falls down close to where she is walking and breaks its arm. Let us assume that (as many would suppose) Sally now has an obligation of some kind immediately to help the child. Now imagine that before she can do this, at four-thirty-one P.M., an elderly person collapses with a heart attack in a nearby driveway. I take it that at this point Sally has an obligation of some similar kind immediately to go to the aid of the elderly person, who may otherwise soon die. Does she still have the original obligation immediately to help the child? Surely not. She certainly does not have an obligation immediately to help the child, all things considered. If we artificially limit the considerations available (in particular omitting any reference to the heart attack) we can say that she has an obligation to help the child given these considerations only. But if we expand those considerations (bringing in the heart attack), surely any obligation immediately to help the child simply disappears.

Someone may doubt this. Why not say that the obligation remains but is discounted or overridden by a more compelling obligation? One

34. The case is similar to that of lying. Some may be inclined to think that a lie is never permissible morally. Some may feel the same about breaking an agreement or a promise. But the majority view goes the other way. Thus the view that one should not break an agreement even to avoid a great evil tends to be seen as absurd. (Possibly it stems from the perception that there is something “absolute” about the obligations of agreement. See below on the “absoluteness” of the obligation of agreement.)

reason is this. We surely do not want to say that Sally violated an obligation to help the child if she rushes to the heart attack victim. This can be disputed if one insists that an obligation immediately to help the child “remains.” But I see no reason for such insistence. Though the *basis* for such an obligation remains, in the actual surrounding circumstances no obligation gets *generated*.<sup>35</sup> In order for the obligation to be generated, the right surrounding circumstances must obtain.

Pretheoretically, then, it appears that there are two quite different types of obligation: a type that is context sensitive in a certain way, and a type that is not. The first type is such that even when an obligation of this type is present in one context, it may in principle disappear if the context is enlarged and an obligation of the same type but with a different content stands in its stead. The second type is such that if an obligation of this type is present in one context (e.g., an agreement stands) then it does not disappear if the context is enlarged (e.g., the agreement still stands, but one can save someone’s life by violating it). It may be discounted in the light of the additional considerations, but it does not disappear.

The joint decision model of agreements provides an explanation of how the obligations of agreement can persist in the way that they are taken to. The obligations of agreement stand as long as the corresponding joint commitment does. Whether or not the commitment stands is a matter for the participants to decide. They are in a position to keep it in existence through whatever changes in the circumstances.

The obligations of agreement are, then, absolute in a sense: they do not go away unless the agreement is rescinded. At the same time, they do not necessarily determine what must be done, all things considered. Once we understand this we see that the type of absoluteness in question here need not lead to uncritical compliance.

The obligation in the case of the injured child would normally be referred to as ‘moral’ obligation. Though it may not matter for some purposes whether we call both of the types of obligation discussed here “moral”—though that may even have something to recommend it—it is important to note that our obligations differ widely in logical character. The obligation of agreement is of one type, which is in a sense absolute, or lacking in overall context sensitivity. It has its source (I propose) in an existing joint commitment which stands unless and until those who participate in it concur in its dissolution. Obligation of the sort in the injured child case is of the other, context-sensitive

35. As Sam Wheeler has pointed out to me, there is some concordant discussion in Donald Davidson’s “How Is Weakness of the Will Possible?” in *Moral Concepts*, ed. Joel Feinberg (London: Oxford University Press, 1969), pp. 108–9.

type. The nature of its presumed source is a controversial matter that I am waiving here.<sup>36</sup>

Whether we call them 'moral obligations' or not, agreements generate obligations, obligations which only disappear when the agreement is at an end. One who manifests a simple "take it or leave it" attitude to agreements has failed to understand what an agreement is.<sup>37</sup> Precisely when an agreement may reasonably be violated appears to be a matter of judgment that goes beyond our understanding of the logic of agreement itself.<sup>38</sup>

### *Revisiting the Obligation Argument*

Let us now briefly look back at the original version of the obligation argument (in the first part of Sec. I). Let us refer to the obligation of agreement as I have so far characterized it as 'obligation(ABS)'. The most clearly correct version of the first premise is this:

- 1B. Any genuine agreement generates an obligation(ABS) to abide by it: this is a conceptual matter.

Taking 'obligation(CON)' to be a (completely) context-sensitive obligation, and hence an obligation of a distinct type from obligation (ABS), let us interpret the second premise of the obligation argument as follows:

- 2B. If an apparent agreement is made in the face of coercion there is then no obligation(CON) to abide by it.

Even if 2B is true, when coupled with 1B it does not give us the desired conclusion:

- 3B. A genuine agreement cannot be made in the face of coercion.

36. If this sort of obligation is based on "where the most value lies," as some theorists would have it, one can see why it has the context sensitivity it does: enlarging the context at least potentially changes the value of any act. Perhaps some contexts are such that the context-sensitive obligation that arises within them will stand through however many possible enlargements of the context. The persistence of the obligation is then still a function of the character of the enlarged context.

37. Compare Michael Robins, *Promising, Intending, and Moral Autonomy* (Cambridge: Cambridge University Press, 1984), p. 109, who—in contrast to Hare—suggests that unless we think of promissory obligations as moral obligations we will not take them seriously.

38. In *On Social Facts* and elsewhere I have suggested that the obligations of joint commitment are not moral obligations. I am inclined to think that this is a useful way of looking at things. Throughout the present article I attempt to be as agnostic as possible with respect to the definition of moral obligation. Everything of substance that I have to say about coerced agreements can be said without deciding that issue. The considerations raised in this section, meanwhile, are clearly relevant to that decision.

To think so would be to treat obligation(ABS) and obligation(CON) as the same. In other words, it would be to commit a fallacy of equivocation.

I conjecture that the truth of 1B illegitimately gives rise to the common assumption that a similar premise about moral obligation is definitionally true, moral obligation being viewed as a form of obligation(CON). If I am right about agreements and their obligation, however, no putative facts about the disappearance of any obligation (CON) in the face of coercion could show anything about the impact of coercion on the genuineness of an agreement or the persistence of its obligation. They are simply beside the point.<sup>39</sup>

Suppose that we start with 1B: we assume that the 'obligation' referred to in the first premise of the obligation argument is 'absolute' given the agreement from which it derives. Then the second premise, to be relevant, must claim that if an apparent agreement is made under coercive conditions, there can be no agreement and, hence, no (ensuing) obligation. Why not? The argument seems to collapse into the bald assertion that coercion prevents an agreement from taking place.

#### IV. AGREEMENTS, COERCION, AND OBLIGATION

##### *Coerced Agreements*

In direct discussion of the relation between coercion and agreement I shall assume that agreements are joint decisions, though many of the following points should be independently plausible pretheoretically.

Nothing that has been said so far entails that coerced agreements are impossible. It is true that if two people have agreed, each is to that extent obligated to act in conformity with their agreement. Once we properly understand the nature of the obligations in this case, we should be less inclined to see this apparently logical consequence of agreeing as a ground for denying the possibility of a coerced agreement. The parties know they have these obligations. They have intentionally assumed them.

39. As I have already indicated, changing 1B by replacing "an obligation(ABS)" with "an obligation(CON)" does not give us a good argument. (i) If the new version of 1B refers to an alleged obligation(CON) that flows from an apparent agreement considered in isolation from its circumstances (of which coercion could be one), then the presumed fact that coercion left one with no moral obligation to conform to one's apparent agreement would not show that there was no agreement at all. For it would have no bearing on the question of the implications of an apparent agreement considered in isolation. That is why in Sec. II of this article I began by construing the first premise in terms of an obligation that stood in the light of all background circumstances. (ii) Suppose we now explicitly interpret the new version of 1B in terms of an alleged obligation(CON) that stands in the light of all background circumstances. As we have seen, though the desired conclusion 3B follows when the first premise is so construed, the premise is then implausible, at least given the limited understandings that seem to be necessary to enter an agreement.



What could be the point of coercing agreement? Clearly there is a possible point. Namely, to introduce an extra motivational force into the situation: the force precisely of perceived obligation. Those who agree know that they are under an obligation. They may or may not understand that this leaves open the question whether they must fulfill it, all things considered. Simply knowing what an agreement is would not seem to give one this understanding. Nor is the outcome of a particular person's deliberations on the merits of keeping a particular agreement always obvious in advance. So getting someone to agree may, at the end of the day, amount to getting them to act, even when good judgment would not allow that the act was mandatory, all things considered.

Let us consider a gunman case: "Yes, I agree to marry Emily," says Ben, staring down the barrel of Emily's father's shotgun. Having agreed, he may or may not marry Emily in the end. If he does, it could be because he has agreed to do so.

"I don't owe him anything," Ben might say of Emily's father, "after all, he forced me to agree." We can construe this in ways conformable to my analysis so far. 'Owe' here could allude to a context-sensitive type of obligation, distinct from the obligation of agreement. According to at least some conceptions, it could allude, that is, to moral obligation or an 'owing' based on all morally relevant considerations.

It may well be that (in some appropriate sense) Ben's coercer is not morally entitled to his performance, does not morally deserve his performance, and will have no basis for complaint against Ben's moral character if he breaks the agreement, given what he himself did to bring the agreement about. It may be that Ben has no moral obligation to perform, that Ben owes his coercer nothing, morally speaking. However, all this can be true without it being clear what Ben should do all things considered. One thing to consider, or so it seems, is the fact that he agreed and has the corresponding obligation to conform.

Does the fact that Ben was coerced into agreeing in itself show that there was something to be said against marrying Emily from Ben's point of view? I have made no attempt here to define coercion or duress, preferring to operate on the basis of tacit understandings and cases assumed to be clear-cut. Nonetheless, our understanding of these terms is relevant here, and so must be (all too briefly) considered.

A standard way to define duress in the law is exemplified in the following statement: "Duress may consist of actual or threatened violence or imprisonment . . . [which threat] contributed to the decision of the person threatened to enter into the contract . . . even though he might well have entered into it all the same if no threats had been made."<sup>40</sup>

40. Guest, ed., p. 240.

It is explicitly not made part of this account of duress that the person was pressured into entering an agreement he would not have entered except under such pressure. It is only that the pressure must in fact have motivated him in the situation under consideration. If we define a coerced agreement in terms of a threat of violence that in fact *motivated* a person to agree, the definition does not tell us that there are independent reasons *against* keeping the agreement once it is made.

Whether the agreement should in fact be kept will be a matter of judgment, in the light of all the relevant facts. Constrained by Emily's father's gun to agree to marry Emily, Ben might decide, after a thoroughly enjoyable evening with her, that he would very much like to marry her. There will then be no conflict between the agreement and his evaluation of the courses of action open to him. The agreement may well be instrumental in motivating his performance. He may be an indecisive person. His obligation under the agreement may be what moves him.

In contrast, constrained by an intruder's knife to agree to sit still while he rampages through the house, Emily may perceive that she has a moral duty to escape as soon as he leaves the room. In other words, she has a moral duty not to conform to the agreement. In such cases the cards are stacked against the coerced agreement, or so most people would hold.

It is evident that in some cases the coerced person agrees to do something she finds wholly repugnant, though she deems death to be worse. It may seem dangerous to argue that in such cases, as in others, there is an obligation to perform. This is surely not so, however, if it can also be argued that the other party has no moral right to performance, the coerced person is morally entitled to renege, and it only remains to consider whether she has adequate reasons for failing to conform. If to conform would be to bring about a great evil, perhaps her own terrible pain and anguish, then these reasons are presumably more than adequate. All things considered, she *should* violate the agreement if she can.<sup>41</sup>

### *Coercion and the Extent of Political Obligation*

If I am right about agreements, there is no need to accept Simmons's claim that few people are in a position to agree to fulfill the duties of citizenship. Simmons argues that even in modern democracies the

41. It is of some practical importance to stress that the fact that one agreed does not show that one was not coerced into agreeing. Judges have been known to rule against a verdict of rape on the grounds of the victim's consent to sexual intercourse. Surely the judgment of rape should focus on the presence and influence of coercion, not on whether there was acquiescence or some type of agreement. There is no space here to attempt a careful discussion of this important issue.

circumstances of the proposed agreement are generally analogous to coercion. If my argument here succeeds, it robs this argument of all force, for I have argued that agreements are possible in the face of outright coercion. I have also argued that whenever there is an agreement, there are corresponding obligations to conform to it. Thus the potential scope of an "actual contract" theory of political obligation is wider than Simmons's claim suggests.

It is standard to formulate the philosophical problem of political obligation as a problem about moral obligation. Whether or not the obligation of agreement is best characterized as a form of moral obligation, the obligations of agreement are surely highly relevant to the traditional questions concerning political obligation.

First, as far as the literature goes it appears that what is understood to make an obligation "political" is not so much the type of obligation but its content. Political obligations are obligations to obey particular laws, support particular political institutions, and so on.

Second, the political obligations of interest have been contrasted with the "positional" or "institutional" obligations that some group or individual associates with a particular social position (such as the position of citizen). People may come by "positional obligations" in this sense without themselves having a new reason for acting.<sup>42</sup> One's participation in an agreement generates obligations that are reasons for acting. They are not merely positional obligations.

Third, contracts or agreements have always been considered important sources of political obligation in political philosophy. It appears that the situation has generally been this: the obligations intrinsic to agreement have not so much been ignored as misperceived, as in the obligation argument that takes coercion to rule these obligations out.

I argue elsewhere that we can give a plausible account of a central class of perceived political bonds in terms of obligations of the type agreement produces—in terms, more specifically, of the obligations of joint commitment. This surely has to count as a theory of at least one central kind of political obligation.<sup>43</sup>

There are a number of importantly different ways of modeling the psychological reality that underlies the persistence of political regimes of all kinds, including oppressive dictatorships. Sometimes obedience may be wholly motivated by fear. At other times conformity may be at least partly motivated by the sense of participation in an

42. See Simmons on "positional" obligations, *Moral Principles and Political Obligations*, pp. 16 ff.

43. In *ibid.*, Simmons takes as the primary datum for the theorist of political obligation an inclination to see ourselves as tied to our government and our country's institutions by "political bonds." I argue in "Group Membership and Political Obligation" (*Monist*, vol. 76 [1993], in press) that the datum he focuses on can plausibly be accounted for in terms of the obligations of joint commitment.

agreement or something like it. If people perceive themselves to be parties to an agreement to conform, this will provide an additional pull in the direction of conformity, a pull that works at the level of reason.<sup>44</sup>

I cannot explore the actual extent of agreements and agreement-like phenomena here. As I have explained elsewhere, I believe it to be quite wide.<sup>45</sup>

In sum, unless political theorists are clear about the nature of agreements, they are likely to misunderstand the scope of actual contract theory, to underestimate the potential extent of the obligations that derive from agreements and the like, to miss a crucial basis of perceived “political bonds,” and to overlook a central type of political motivation.

## V. CONCLUSION

In the course of this article I have appealed to several judgments I suspect people would make pretheoretically: not thoughtlessly, but in advance of accepting any particular articulated theory of agreements. These include the assumptions that those who enter into agreements know that they do, that there is a conceptual connection between agreement and obligation, that the obligations of agreement makers are interdependent, and that it is possible to force someone to agree.

There has not been space to argue for the independent plausibility of each of these judgments from a pretheoretical point of view, though such arguments can be given. In any case, the proposal about agreements and their obligations presented here provides indirect support for each of these judgments individually: the mere fact that it provides a single articulated basis for them all encourages the idea that each one is indeed an acceptable ‘ground-level’ judgment. At the same time this systemizing role helps to confirm the acceptability of the account of agreements itself, given that there is at least some plausibility in the allegedly ground-level judgments. I take it that the account also has some independent plausibility of its own. Thus theory and ground-

44. Thus David Schmidtz is overgeneral when he writes: “The idea of a conqueror becoming justified by forcing his captives to pledge allegiance as the price of escaping with their lives is hardly plausible. It think it is more charitable to Hobbes to read his discussion as a purely descriptive account of the possible ways in which sovereigns can actually emerge, with no normative implications intended” (“Justifying the State,” *Ethics* 101 [1990]: 98). My dispute is with ‘no normative implications’. No *moral* implications of a certain sort, perhaps. But if captives enter an agreement or make a similar commitment, then this does have normative implications of an important kind. That is, it bears on what the captives have reason to do. Any genuine pledge or agreement has its own normative weight—as the conquering sovereign may well discern.

45. See *On Social Facts*, and elsewhere. See also M. Walzer, *Obligations* (Cambridge, Mass.: Harvard University Press, 1970), introduction, pp. ix–xvi.

level judgments are (to a degree at least) in "reflective equilibrium," to borrow Rawls's phrase.

Given a certain picture of agreements, and a particular understanding of obligation, coerced agreements are possible and they all generate obligations of conformity. There is much that is speculative in this picture. At the same time, I am not aware of any more intuitively satisfactory picture of the way things are in our everyday conceptual scheme of agreements, coercion, and obligation.