

## Building Trust in Divided Societies\*

DANIEL WEINSTOCK  
*Philosophy, Université de Montréal*

**P**OLITICAL philosophers in the liberal tradition have, at least since the publication of Rawls's *Theory of Justice*, tended to focus on how social institutions might best embody and promote *justice*. But they have in general not considered how institutions might, within the limits set by liberal justice, contribute to the *unity* of modern societies. Yet disunity is a considerable problem for a number of tolerably just societies in the modern world. A number of separate factors contribute to the disunity of contemporary liberal democracies. Immigration has considerably altered the ethno-cultural fabric of those liberal democracies which take in large numbers of immigrants (such as Canada, the United States and Australia), and even of those states with more restrictive immigration and naturalization policies, such as Germany. Few modern democracies possess the kind of cultural homogeneity which John Stuart Mill had seen as a condition of the viability of representative government.<sup>1</sup> Moreover, the freedom of thought, assembly and conscience which is distinctive of liberal democracy has given rise to a great diversity of ways of construing the good life and the good society.<sup>2</sup> And finally, it is possible that liberal justice itself generates centrifugal tendencies. That is, it is possible that the norms of liberal justice consistently applied would involve acceding to the demands for autonomy, limited self-government, and exemption from prevailing legal norms of a wide range of social groups, and not only of those ethno-cultural groups which "multiculturalist liberals" like Will Kymlicka have emphasized.<sup>3</sup>

\*Thanks are due to the following people for helpful comments: Simone Chambers, André du Toit, Claire Finkelstein, Ruth Gavison, Amy Gutmann, Alison Jagger, David Kahane, Helen Nissenbaum, Josh Ober, Connie Rosati, Barry Strauss, Jim Tully and Melissa Williams. Thanks are due to Rainer Bauböck, Dick Moran and an anonymous referee for this journal for extensive written comments. Thanks, finally to Sayumi Takahashi for very able research assistance. The paper was written while I was a Fellow at the Laurence S. Rockefeller Center for Human Values, Princeton University, to which thanks are owed for material support.

<sup>1</sup>John Stuart Mill, *Considerations on Representative Government* (Amherst, N.Y.: Prometheus, 1991), pp.308–19.

<sup>2</sup>It is for this reason that John Rawls believes that the "fact of reasonable pluralism" is a permanent feature of modern liberal democracies. See his *Political Liberalism* (New York: Columbia University Press, 1993), p. xxiv.

<sup>3</sup>See Kymlicka's *Multicultural Citizenship* (Oxford: Oxford University Press, 1995). I have defended the claim that Kymlicka's argument has a broader application than he allows in my "How can collective rights and liberalism be reconciled?", *Blurred Boundaries: Migration, Ethnicity, Citizenship*, ed. R. Bauböck and J. Rundell (Aldershot: Ashgate, 1998).

It is plausible to suppose that *some* degree of social unity is an empirical condition for the viability of liberal democracies' main institutions. Citizens must to some extent view themselves as *united* under the law rather than as simply happening to stand there side by side. Liberal philosophers have been surprisingly silent about what might be done by liberal regimes to foster unity.<sup>4</sup> Those normative social theorists who have focussed on the conditions of social unity have usually viewed liberalism, or at least some of liberalism's central normative commitments, as part of the problem, and have thus recommended that certain aspects of liberal political morality be rejected. Thus, for example, Mary Ann Glendon has pointed out the litigiousness of liberal societies in which individual rights are placed at the summit of the normative pyramid.<sup>5</sup> And Charles Taylor has argued that the liberal commitment to neutrality precludes liberal societies from being able full-bloodedly to endorse the kind of patriotism sometimes required to prop up liberal institutions.<sup>6</sup>

Those thinkers who feel that some form of liberal-democratic theory provides the most appealing vision of *justice* will be unhappy with the abandonment, in the name of unity, of central elements of liberal justice. My aim in this essay is to argue that considerations of unity can figure, alongside and limited by considerations of justice, as normative constraints upon the design of liberal democratic institutions, and to determine, at least in their broadest outlines, what these normative constraints might look like.

I will proceed as follows: first, I will distinguish unity from other desirable properties which societies can possess, and in terms of which some theorists have sought to account for unity. I will then describe the *kind* of unity which liberal democracies ought to pursue. In section I, I will argue that while unity in and of itself is not a morally worthwhile goal, it becomes such a goal when it is sustained by *trust* among citizens. Next, in section II, I will offer a characterization of trust which will attempt to bring out its distinctiveness with respect to those properties with which it is sometimes conflated. I will then, in section III, consider certain obvious objections which one might want to make to the relevance of trust as described to an understanding of *political* unity. This will involve refining the account of trust given in section II so as to bring to the fore the specific aspects of trust-relations in which we are interested. Specifically, we will focus on trust between citizens *qua* bearers of politically salient identities or roles. My characterization of trust will then allow me to identify, in section IV, those aspects of the trust relationship which might be affected by judicious institutional design. I will end by considering, in section V, the objection that

<sup>4</sup>An exception is Will Kymlicka, "Social unity in the liberal state," *Social Philosophy and Policy*, 13 (1996), 105–36. Kymlicka argues that shared identity rather than, say, shared values, account for social unity. For reasons which will be made clear below, I think identity ought to be seen as a dependent, rather than as an independent variable.

<sup>5</sup>Mary Ann Glendon, *Rights Talk* (New York: Free Press, 1991).

<sup>6</sup>Charles Taylor, "Cross-Purposes: The Liberal-Communitarian Debate," *Liberalism and the Good Life*, ed. Nancy Rosenblum (Cambridge, Mass.: Harvard University Press, 1989).

institutionalizing trust in the way suggested in the previous section is incompatible with the very nature of trust, and that it in fact risks replacing trust with (mere) reliance.

### I.

Discussion of the question of social unity in political philosophy has been plagued by insufficient care in distinguishing between (at least) four desirable properties which societies can have, and which all seem relevant at first glance to the task of accounting for the basis of social unity. *Unity*, as I understand it, has to do with the continuing desire on the part of a population to continue living under the same political institutions, or, perhaps more precisely, with the *absence* of any desire to sever the existing bonds of political association. *Stability* has to do with the members of a society sharing a conception of justice which can reliably secure the consent of successive generations of citizens. *Cohesion* has to do with the extent to which a society's members share broader moral values or a religious, cultural and/or ethnic identity above and beyond their espousal of more specifically political principles. And *cooperation* denotes the extent to which a society's members are disposed to identify their individual interest with the common good, and, accordingly, to engage in joint undertakings, be they political or economic.

These properties clearly belong to the same general family. Let me for the sake of brevity refer to the set of property-names which includes the aforementioned properties as *C-properties*. They are however conceptually distinct. There is no entailment between any of them and, most important for our purposes, no entailment holds between unity and any of the other C-properties. Now, of course, there are contingent relations, and then again there are contingent relations. It could be the case that the possible worlds in which the instantiation of one of the C-properties fails to give rise to unity are very remote indeed. But this does not seem to be true of the C-properties mentioned above. It is actually quite easy to imagine scenarios under which stability and cooperation do not give rise to unity. (The problem with cohesion is somewhat different, and will be spelled out below.) A cursory overview of recent attempts to show how other C-properties might bring about unity is instructive in this regard.

For example, David Miller and Stephen Macedo have argued that a reinvigorated practice of citizenship might have unity-enhancing consequences, as the role of "citizen" is one that we can all share despite our other differences.<sup>7</sup> In terms of the typology suggested above, unity is here being conflated with cooperation or, more charitably, cooperation is seen as contributing reliably to unity. But it is just as easy to imagine increased practices of citizenship having a

<sup>7</sup>David Miller, "Citizenship and pluralism," *Political Studies*, 43 (1995), 432–50; Stephen Macedo, "Community, diversity and civic education: toward a liberal political science of group life," *Social Philosophy & Policy*, 13 (1996), 240–68.

deleterious effect on unity. Citizenship in mass societies such as ours is bound to be centered on the smaller-scale associations that make up civil society rather than on legislative institutions that encompass the whole society. And practices of citizenship centered around such associations risk awakening or exacerbating social divisions, or making participants attribute disproportionate importance to the limited good of their association as against the larger public good.<sup>8</sup> A very similar tale can be told about the relation between a shared conception of justice and of the good society: as Wayne Norman has pointed out, history gives us no evidence to support the thought that people typically seek to unite politically with others with whom they share values.<sup>9</sup>

Some philosophers claim that in order for the requisite level of unity to be guaranteed, citizens living under common political institutions must take themselves to share a common national identity.<sup>10</sup> Societies which contain some significant degree of ethno-cultural diversity, and whose members can therefore not be expected spontaneously to converge around a shared national identity, must on this view undergo a process of *nation-building*, whereby the state takes on the task of forging a new unified national identity which would supplant the diverse cultural identities already present on a given territory. The normative constraints posed by a (broadly) liberal political morality preclude the goal of a shared national identity being attained through illiberal practices aimed at stamping out sub-national languages and cultures (as was done by many nineteenth century European nation-builders), or through the kind of “ethnic cleansing” that has marked the former Yugoslavia. Nation-building projects which are concerned with staying within the bounds of some kind of liberal political morality will naturally have to avoid such excesses, and will therefore have to adapt themselves to the presence within their midsts of minority cultures which must somehow be included rather than destroyed. Thus the “liberal nationalist” and “multiculturalist liberal” theories which have in recent years been defended by Yael Tamir, David Miller and Will Kymlicka.<sup>11</sup> Tamir, Kymlicka and Miller, though in different ways, have defended the idea that though a common national identity is crucial to the viability of multicultural societies, a liberal nationalist will have to accept that the creation of a common national identity will have to result from the mutual accommodation of majority

<sup>8</sup>On this point, see Will Kymlicka and Wayne Norman, “The return of the citizen,” *Theorizing Citizenship*, ed. R. Beiner (Albany, N.Y.: SUNY Press, 1995).

<sup>9</sup>For instances of what Norman calls “the ideology of shared values,” see for example the essays collected in Jacques Lenoble and Nicole Dewandre, eds, *L’Europe au soir du siècle* (Paris: Le Seuil, 1992). See also Lawrence E. Mitchell, “Trust and the overlapping consensus,” *Columbia Law Review*, 94 (1994), 1918–35. A devastating critique of such views is provided in Wayne Norman, “The ideology of shared values,” *Is Quebec Nationalism Just?*, ed. Joe Carens (Montreal: McGill-Queen’s Press, 1995).

<sup>10</sup>See, for example, Will Kymlicka, “Social unity in a liberal state,” *Social Philosophy and Policy*, 13 (1996), 105–36. He writes that “people decide who they want to share a country with by asking who they identify with, who they feel solidarity with” (p. 131).

<sup>11</sup>Yael Tamir, *Liberal Nationalism* (Princeton, N.J.: Princeton University Press, 1993); and David Miller, *On Nationality* (Oxford: Oxford University Press, 1995). Kymlicka, *Multicultural Citizenship*.

and minority identities. I have argued elsewhere that this is a very difficult task which can very easily backfire.<sup>12</sup> The “watering down” of the majority ethno-national culture that will be required is likely to be resented by many nationalists, and the imposition of certain aspects of the majority culture is likely to be viewed as unduly impositional by some members of minority cultures. In terms of my typology, liberal nationalists would have the state engineer cohesion, and while cohesion might be a sufficient condition for achieving unity, it is, as I will suggest below, not necessary. Theorists and politicians seeking to secure the conditions for unity need not set themselves as exigent a task.

The principal attempts by philosophers in recent years to give an account of what might contribute to the unity especially of multicultural societies have failed, in large measure because they have tended either to conflate unity with another C-property, or because they have thought, far too optimistically, that the instantiation of one such property would reliably lead to the attainment of unity. The relationship between these other C-properties and unity is contingent, and it is fairly easy to imagine scenarios under which they obtain while unity remains unrealized. The relation of cohesion to unity is slightly different. While it is difficult to see how a society could be cohesive without being united, cohesion is a very demanding goal which is not required in order to secure unity. Cohesion accounts at least point us in the right direction, however. That is, it is a social property which, when it is instantiated, would seem reliably to bring unity in its train. That is, it is difficult (though perhaps not impossible) to imagine a cohesive society that was not also united. What we are looking for is a social property that is causally connected in this manner with unity, but which does not, as it were, overshoot the mark.

These questions might be raised at this juncture: why are we looking for an account of unity *in terms of* some other social property? Why do I think it important that unity be achieved *through* the instantiation of some other value? In other words, if unity is so great, why not aim for it directly?

The answer is that unity, in and of itself, is a morally ambiguous property. By this I mean that it need not always embody values which a liberal democratic society will want to affirm. On the one hand, social unity would seem in all cases to have the desirable feature of allowing a society to withstand the kinds of crises which are almost bound to occur from time to time in the life of any polity. Societies are often rocked by crises during which their fundamental principles and self-understandings are put in question. Such crises will characteristically give rise to deep divisions in society, and such divisions will be particularly unity-imperilling when they overlap with the kind of ethno-cultural split which most often<sup>13</sup> occasions secessions. My claim is that it is, all other things being equal, a

<sup>12</sup>Daniel M. Weinstock, “Is there a moral case for nationalism?,” *Journal of Applied Philosophy*, 13 (1996), 87–100.

<sup>13</sup>Most often, but not always. It seems possible, for example, to account for Padanian autonomism entirely through economic considerations.

virtue of societies that they be able to undertake the often acrimonious debates and social deliberations which are needed to overcome such crises without putting the “we” in question, that is without abandoning the sense that this is *our* problem, and that we must thus arrive at a common solution to it.

But on the other hand, social unity can also have unattractive features. Social unity might simply be the result of a political culture having failed to breed any capacity for independent critical thought in its citizenry, and having instead successfully inculcated in them the belief that they are as individuals indissolubly and organically linked with the whole. Short of such brainwashing, social unity can, albeit unstably, be achieved through terror and oppression, which might simply raise the cost of seceding to an unacceptable level.

So we want the goods that unity makes possible, but not at any cost. For unity to be an attractive goal to pursue, it has to be secured in the right way. What we want is, in Peter Railton’s helpful term, a vindicating explanation of unity, that is, an account of how it can be causally generated which can be fully affirmed by those individuals bound together by it.<sup>14</sup>

Our account of unity will therefore have to satisfy two desiderata. First, we want to relate unity to some property with which it is in close causal connection. That is, we want a property which reliably generates unity, that is, which is such that it is difficult to imagine real-world conditions in which it was instantiated in a society without being accompanied by unity. And second, we want the account to *vindicate* unity. We want it to be the case that, when people realize that unity has been achieved by the causal operation of the property in question, this will not detract from, and will if anything contribute to, the appeal of unity.

My suggestion is that the property of *social trust* is particularly well suited to the role just defined by these desiderata. It is difficult to see how people who trust one another in the sense which I will be defining below can fail to be united. And it would seem at first glance that when unity is underpinned by trust, the realization of this fact will not detract from the value attached to unity by citizens, as opposed to the effect which the realization that unity was being sustained by terror or brainwashing would tend to do.

What do I mean by trust? How does it differ from the other C-properties listed above? What is distinctive about the *value* that it embodies, as compared to the values embodied in cooperative relations, or in social relations marked by cohesion or stability? It is to these issues that I now turn.

## II.

Accounts of social unity according to which unity requires either cohesion, cooperation or stability all assume that (to denote the core idea by means of a

<sup>14</sup>See Peter Railton, “Naturalism and prescriptivity,” *Social Philosophy and Policy*, 7 (1990), 151–74 at pp. 166–7. I thank Connie Rosati for having helped me to see the relevance of this idea for my purposes.

slogan), *unity requires sharing*. Unless we are joined together by some shared national identity, or by some thick set of values, or at least by a conception of social justice, then we are not really joined together at all.

The relation of trust is distinct however in that it does not require sharing in this way. Although I will be complicating this skeletal schema below, the distinctiveness of trust can already be brought out by reflecting upon its most basic elements. Following Annette Baier and Russell Hardin, among others, we can begin by noting that trust is a three-place relation: “A trusts B to do,” or “A trusts B with valued thing C.”<sup>15</sup> Those aspects of the relation which I want to stress in the present context in order to distinguish it from relations which require sharing are, first, that it is an *asymmetrical* relation in which an individual is making herself vulnerable to another by entrusting some aspect of her good to another person. The asymmetry derives from the fact that it is part of what trust involves that the trusted does not have a self-interested reason to care for the good in question in the manner required by the truster. I can, other things being equal, rely upon another person acting on her self-interest. *Trust* requires that the person being trusted act because of how she stands to me rather than of how she stands to that which I am entrusting to her. This brings us to a second important aspect of trust: it is a *moral* relation, in that it is related to one’s beliefs or attitudes concerning another person’s motivations toward one. I am not *trusting* you if my expectation that you will do X or care for C is based on the fact that I know that a device has been implanted in your brain that will *cause* you to, or that I know that your *only* motivation is avoiding being shot by some third party who has credibly threatened to do so if you do not (though, as will be made clear below, trust is compatible with *mixed* motivations). Unlike some authors, however, I think it is sufficient that my trust be based on the belief that the trusted is not ill-disposed toward me, rather than that he be positively well-disposed toward me. It is sufficient for the trust I vest in the total stranger on the street from whom I ask directions to be justified that he not have any positive reason to steer me wrong. It seems odd to say of him that he must be *positively* well-disposed toward me, or that his motivations toward me be benevolent, as a condition of meriting my trust.

If I am right in the above characterization, there is a moral relation we stand in to others in some circumstances that does not depend upon our sharing either a set of values, a collaborative end or a (national, ethnic, linguistic, religious) identity. Indeed, it might be seen as depending upon our *not* sharing one or the other of these, as they may as it were *overdetermine* others’ behaving toward the object of our concern as they would were they simply acting in a *trustworthy* manner.

<sup>15</sup>This much unites authors who have otherwise very different conceptions of trust. See for example Annette Baier, “Trust and antitrust,” *Moral Prejudices* (Cambridge, Mass.: Harvard University Press, 1994), pp. 95–129 at p. 101, and Russell Hardin, “The street-level epistemology of trust,” *Politics and Society*, 21 (1993), 505–29 at p. 505.

What is the value embodied in the kind of relation I have just described? It might be argued at this point that trust as I have just characterized it is, morally, an excessively thin relationship. A society marked *only* by trust as I have characterized it would be a depressing and uninspiring one, hardly the kind of thing capable of attracting and sustaining the allegiance of a citizenry.

A clarification is in order: I have not wanted to claim that it is desirable for citizens to be linked only by trust as I have characterized it. My intention is to claim that one can conceive of a relation that binds citizens that is not one in which identity, value or cooperation are shared, and that attention must be paid to the conditions which must be in place to ensure that this relation will not be damaged, as it is at the core of the problem of social unity. It is moreover on the basis of a fund of trust that a more morally exigent and rewarding relation can be built. The overtures which people make to one another on the road to friendship, for example, would be inconceivable were it not for some degree of initial, default trust.

Despite the fact that trust is a minimal condition of adequacy of social relations, one which, it is hoped, will constitute the groundwork for more full-blooded relations, I still want to maintain that there is both instrumental and intrinsic value to social trust which cannot simply be accounted for in terms of the value realized by these more full-blooded relations.

The instrumental value of a relation of trust among citizens should be clear by now. First, as we have seen, it makes possible those more full-blooded relations (friendship, cooperation) through which truly communal goods are made possible. And second, even in the absence of such shared goods, it makes possible the attainment of individual ends which, at least in a social setting, depend, if not on the active cooperation, at least on the forbearance of those who could in theory pose a threat to the individual's attainment of her goals.

Might there be intrinsic value in a relationship apparently as thin as trust? In order to answer this question, we must attempt an answer to the more basic question of what it might mean for a relationship to have *intrinsic* value. I take it that in order for a relationship to have intrinsic value, there must be some good which it realizes which is independent of the good which accrues to the truster (and perhaps also to the trusted) *as a result* of taking part in a trusting relationship. There must be some good *in* relations of trust which does not simply reduce to the good which is created *as a result* of taking part in such relationships.

I think that there is such value, and would like briefly to attempt to characterize it. The point is a recognizably Kantian one: relations of bare trust as I have described them are the only relationships in which we can express respect for others simply as human. "Thicker" relations, be they the intimate relations of friendship, love or kinship or the somewhat more instrumental ones of cooperation in its many guises, are ones in which our benevolence toward others is, as it were, overdetermined. There are a multiplicity of motives which



might in such relations move us to assist others in the attainment of their ends, or at least to forbear from getting in their way. It might be that we have a selfish interest in seeing to it that another's end is reached, or that we identify with another (because they are our spouse, child, compatriot, etc.) to the extent of seeing their ends as in a sense *our* ends. A relationship of bare trust is one in which none of these motives is present. In it we find ourselves like the Job-like figure from *Groundwork I*, his mind "overclouded by sorrows of his own," acting well despite the fact of being bereft of any motive *but* that of respect for others to act benevolently, or at least non-malevolently.<sup>16</sup> Our reacting positively to the trust placed in us by a stranger expresses the respect we have for others *qua* distinct setters and pursuers of ends, a respect which takes the form of taking another's ends and his pursuit of those ends as a reason to assist or at least to forbear from posing an obstacle to the other's pursuit even in situations where one derives no individual benefit from assisting, or in which one might even have countervailing reasons *not* to assist or forbear.

Looking at the relationship from the point of view of the truster, moreover, the action of placing one's trust in others expresses one's belief that the other is a being *capable* of acting in a manner respectful of the personhood of others. Even putting aside the question of whether acting in a trusting manner might elicit trustworthiness from others who may not have been antecedently disposed to act in a trustworthy manner,<sup>17</sup> I would claim that there is value in what the fact of placing trust in others expresses about the beliefs about others upon which we are prepared to act (though in cases in which that trust is betrayed, that value may very well be overridden by the disvalue incurred by the truster because of her decision to act in a trusting manner).

So I think that trusting relationships among strangers have both instrumental and intrinsic value which cannot simply be seen as a pale or incomplete version of value more fully realized in cooperation, friendship, love, or what have you. The ends we pursue individually, and for which we require the assistance or forbearance of others who do not share our ends, are simply not the same as the shared ends which characterize our cooperative and intimate relations. And the respect we betoken in our relations to others is simply not the same as the benevolence we are disposed to observe in our relationships with friends, spouses, associates, coreligionists, etc., in which our motivations are more complex, and our benevolent actions perhaps overdetermined by self-interest (understood in the broadest sense as including the identification with the interests of those others with whom we are engaged in some kind of intimate relationship).

<sup>16</sup>Immanuel Kant, *Groundwork of the Metaphysic of Morals*, trans. H. J. Paton (New York: Harper, 1953), p. 11 (AK. 398–9).

<sup>17</sup>On this point, see Philip Pettit, "The cunning of trust," *Philosophy & Public Affairs*, 24 (1995), 202–25 and "Republican theory and political trust," *Trust and Governance*, ed. Valerie Braithwaite and Margaret Levi (New York: Russell Sage Foundation, 1998), pp. 295–314 at pp. 304–8.

## III.

I hope that I have established in the foregoing section that trust is a property of social relations which is distinct from the other C-properties listed in my typology, and that it embodies independently attractive values. It would thus seem to satisfy *one* of the two desiderata described above. That is, were it the case that trust as described reliably generated unity, the value embodied in trust would give a weighty reason to *affirm* unity achieved in this manner. The question that must now be faced is what relevance this account could possibly have for the question of *political* unity. Trust characterizes the relation of strangers in civil society, but surely what is needed in the political context is something quite different, such as trust between political *elites*, or citizens' trust in political leaders and institutions.<sup>18</sup>

I clearly do not want to deny that the smooth functioning of political institutions depends upon citizens trusting their leaders and believing in the "point" of their institutions. Nor do I think it controversial that trust among elites is crucial in helping societies navigate successfully through periods of crisis. What I do want to deny is that trust among elites and institutional trust render trust among citizens superfluous. There is, first of all, an aspect of institutional trust that is irreducibly interpersonal. Part of what it means to trust institutions is to trust that the people who occupy positions of authority within them will interpret and apply institutional rules in a trustworthy manner, that they will not use the authority vested in them for self-serving reasons, etc.<sup>19</sup> Trust in institutions and trust among citizens also interact in another way in the context of divided societies. A member of group A's belief that his fellow citizens of group B are not trustworthy might contribute to his sense that their common political institutions disproportionately serve the interests of B's members. The causal arrows can also flow in the other direction (or indeed in both): one's mistrust of government as serving the interests of others to too great a degree might give rise to the belief that those others are inherently untrustworthy.

I would argue moreover that, in divided societies, trust among elites and trust among citizens are related to one another in numerous complex ways, one of which is the following: distrust among the political elites of different groups is in part a function of the distrust that characterizes the relations between members of these groups. In other words, it will be much more difficult for elites to justify acting in ways which betoken distrust when there is no distrust among the members of groups.

There is thus no *prima facie* reason to think that understanding the possible bases of trust among citizens is not relevant to the task of understanding how unity could come to obtain in divided societies. Clearly, though, we must

<sup>18</sup>This worry has been pressed in different ways by Rainer Bauböck and André du Toit.

<sup>19</sup>Cf. Claus Offe, "How can we trust our fellow citizens?" (unpublished paper presented at Columbia Law School, New York, 25 February 1999), pp. 36–9.

complicate the skeletal account of trust offered below in order for the relevance of social trust to the task of achieving unity in divided societies to become clearer. And we must complicate it in the following way: citizens in civil society encounter one another as bearers of different roles. Some of these roles are such that we will be inclined to trust their bearers to a significant degree, to give them, as it were, the benefit of the doubt. Other roles have the opposite effect: knowing that a person occupies a given role might make us adopt a default attitude of *distrust* toward them. Societies are divided in unity-threatening ways when people distrust one another *qua* bearers of politically salient properties such as race, ethnicity, language, nationality, and the like. And I would argue that unity is threatened not only by cases in which secession is a likely consequence of continued distrust because the groups in question are territorially concentrated (Quebec, Catalonia), but also by cases in which distrust exists among territorially dispersed groups where continued distrust threatens to lead to forms of “internal exile” (think of calls for self-reliance and “Black separatism” among African-Americans in the United States). So clearly we must complicate our account of trust relations by bringing to the fore those aspects of our relations to others in which we are characterized by our membership in politically salient groups such as the ones listed above, as well as those goods which are principally at stake when we concentrate on this dimension of our social relations.

What would the needed refinement of the skeletal account look like? Let me begin with the bare, three-place, picture of trust I traced above.<sup>20</sup> It must be augmented in order to serve as a model of the way in which we usually interact with others. For we do not only encounter one another as mere individuals, as simple As and Bs, as it were (though we often do, as my example of asking for directions bears out). Rather, we encounter one another as the bearers of different roles. Thus, at a minimum, the standard picture must be fleshed out in the following way: “A trusted B, as an X, to,” where X ranges over the roles which individuals can occupy, but knowledge of which are not such as to transform the relation between A and B into one of shared identity or cooperation. (The latter restriction is required in order to keep the focus trained on trust-relationships.) Obviously the values of X that will be of interest to us are those that range over the subset of what I have called “politically salient” identities.

A second addition to the initial schema must also be made. Though the *policy* of trusting does not admit of degrees (one either places one’s trust in another or one does not), the *attitude* of trust does.<sup>21</sup> One’s attitude of trust can range from

<sup>20</sup>Cf. Simon Blackburn, “Trust, cooperation and human psychology,” *Trust and Governance*, ed. Braithwaite and Levi, pp. 28–45 at p. 30: “[A]lthough we start with the simple tripartite form, we should approach it with more than half an eye on the various expansions we might want to accommodate”.

<sup>21</sup>The distinction is also present in Pettit’s work on trust, though he employs it to different ends. See also John Braithwaite, “Institutionalizing distrust, enculturating trust,” *Trust & Governance*; Braithwaite and Levi, pp. 343–75, and Barbara Misztal, *Trust in Modern Societies* (Oxford: Blackwell, 1996), pp. 95–101.

distrust to unquestioning trust, with all shades of grey along the way. Whether, on the basis of a given degree of attitudinal trust, one decides to adopt a policy of trust, depends upon a range of factors, including the stakes involved and the options available.<sup>22</sup> At any rate, it is essential that one keep attitudinal and policy trust distinct, if for no other reason than that only by so doing will we be able to evaluate the *rationality* of a given instance of *deciding* to trust.<sup>23</sup>

So if we accept that trust as an attitude admits of degrees, we get a further amendment to the initial schema. We might state the revised schema as follows:

A trusts B *n*ly as an X to  $\emptyset$ .

(We can imagine *n* as ranging between 0 and 1, where 0 means complete distrust and 1 is absolute trust).

The trust relation as I have fleshed it out now possesses three variables. These variables interact in all sorts of ways. Fixing one will have an effect on the values which the other variables can take on. But there are no algorithms allowing us to divine how this will happen. There will be much cultural and social relativity here. For example, finding out that a person is a lawyer or a politician in a situation in which one knows nothing else about her will affect the value of *n* differently in different cultures. Again, fixing X in a situation in which one knows nothing else about a person will affect the value of  $\emptyset$ . For example, knowing that someone is a plumber makes it antecedently more likely that one will entrust her with the fixing of one's pipes than with the care of one's child. Relevant for our purposes is the fact that focussing on politically salient identities will also permit us to limit the relevant values of  $\emptyset$ . Indeed, we will be particularly interested in those interests which characterize us as bearers of politically salient identities.

The schema I have proposed is probably still far too simple. But we can begin to see ways in which it can be put to use in order to answer the question of what governments might do in order to sustain or increase social trust. It gives us a clearer sense of the causal nodes at which third parties, including governments, can intervene so as to increase the likelihood that trust-relationships will hold. I will develop this point at length in the next section.

Let me summarize what I take this section to have established. First, we cannot simply assume that unity will be secured if we ensure "impersonal" institutional trust and trust among elites. And second, the relevance of trust for the task of understanding the grounds of a morally worthy unity becomes more apparent when we focus on trust among citizens *qua* bearers of politically salient identities or roles, and on the interests which they possess in those capacities. Briefly stated, the claim is that when citizens trust one another *qua* members of politically salient groups *with* those interests which characterize them as members of such groups, this will reliably generate a morally worthy form of unity.

<sup>22</sup>These issues have been discussed in Hardin, "The street-level epistemology of trust."

<sup>23</sup>That we can decide to trust seems clear when we consider trust as a policy. It has been denied however, e.g., by Karen Jones, in "Trust as an affective attitude," *Ethics*, 107 (1996), 4–25.

## IV.

Before embarking upon an examination of the various strategies which can be pursued in order to increase trust in divided societies, let me make three plausible assumptions, and strike one methodological warning bell.

The two first assumptions are borrowed from Annette Baier's work. First, let me assume that trust is *malleable*, by which I mean that, *ceteris paribus*, if A trusts B to do a specific thing, and if B responds to A's trust in a trustworthy manner, A's trust will tend to extend to objects not covered by the initial act of trust. Conversely, if B's trust is betrayed, this will have a knock-on effect on other areas of potential trust. Thus, Baier writes that "trust comes in webs, not in single strands, and disrupting one strand often rips apart whole webs."<sup>24</sup> Call this the *malleability assumption*. Second, trust cannot be created *ex nihilo*. In order for actions designed to increase trust to have any chance of success, the web of trust must be at least partially intact. Thus Baier: "Only if trust is already there in some form can we increase it by using what there is to contrive conditions in which it can spread to new areas."<sup>25</sup> Call this the *principle of realism*.<sup>26</sup>

Third, trust and distrust are both to some degree *evidence-resistant*. If I trust you to some significant degree, then I will have a tendency to "read" your behavior in a way which tends to confirm my trust. This might have the effect of making me much more charitable in my assessment of your trustworthiness following an isolated episode of betrayal than I might otherwise be. Arguably, this evidence-resistance is to a point a desirable aspect of relationships of trust: we want such relationships to be resistant to counter-evidence at least to some degree. We might want to question whether someone who comes to distrust a person at the slightest sign of untrustworthy behavior has ever really trusted that person at all. Trust involves a hopeful attitude regarding other people's motives, and as in all cases of hope, this involves some degree of running ahead of the evidence.<sup>27</sup> While this is, within certain bounds, a virtue of relationships of trust, it becomes a crippling aspect of relationships marked by significant *distrust*. If I distrust you, I will tend to interpret your behavior with suspicion, overemphasizing untrustworthy behavior in my overall assessment of you, and placing a negative "spin" even on behavior which should normally count in my favour. The evidence-resistance of distrust makes it the case that it will be far more difficult to repair damaged trust than it would have been to initiate it among strangers, for whom the value of *n* (in the terms of the schema proposed above), might be seen as hovering around 0.5 (or even above 0.5, if we assume,

<sup>24</sup>Annette Baier, "Trust and its vulnerabilities," *Moral Prejudices*, p. 149.

<sup>25</sup>Annette Baier, "Sustaining trust," *Moral Prejudices*, p. 177.

<sup>26</sup>In "Trust and the Rationality of Toleration," *Noûs*, 32 (1998), 82–98, Richard Dees argues that establishing trust between, e.g., warring religious groups requires a kind of arational leap of faith or conversion. This may very well be. I would argue, though, that the type of case he considers simply falls foul of the principle of realism.

<sup>27</sup>On this issue, see Judith Baker, "Trust and rationality," *Pacific Philosophical Quarterly*, 68 (1987), 1–13 at p. 3; see also Jones, pp. 11–14.

like some authors, that evolution has set us with a more positive default setting).<sup>28</sup> Now, as Claus Offe points out, we often form hypotheses about people's trustworthiness on the basis of the *roles* they occupy. Offe emphasizes the importance for the purpose of establishing "horizontal" trust among citizens in a large anonymous society of there being certain roles which are accompanied by positive presumptions of trustworthiness.<sup>29</sup> However, group identities are often in divided societies markers of *untrustworthiness*. When distrust among groups A and B has set in, members of group A will typically see a person's membership in group B as grounding a strong presumption of untrustworthiness. Divided societies therefore pose a particular problem for would-be trust-builders, as group membership will tend to substitute for behavioral evidence when it comes to assessing an individual's trustworthiness. Let me call this the problem of *evidence-resistance*.

The methodological warning bell is this: in what follows, I will be describing ways in which institutional design might offset certain tendencies toward distrust in divided societies. It is important that the reader keep in mind that what are being presented here are suggested normative constraints for *institutional design* rather than directly for *policy-making*. Institutions structure citizens' relations within civil society in a much more foundational and subterranean view than do policies. Fostering trust through institutional design might thus give rise to stabler and more long-lasting results. What is more, they will not risk being as counterproductive as might policies which are more explicitly geared to this task, and toward which members of a divided society might be more wary and cynical.

With these prefatory matters out of the way, we can now frontally address the question: what might trust-enhancing norms governing institutional design look like? Reverting to the slightly more ornate five-place schema of trust suggested above (A trusts B *nly* as an X to  $\emptyset$ ), we can identify the causal nodes on which pressure might be placed by judicious institutional design. First, it can affect the stakes and the options which are involved for A in trusting B, so as to make it more rational for A to *decide* to trust B, even in circumstances in which A's attitudinal trust is low. Let me call such types of strategies *truster-directed*. Or, second, it can attempt to affect A's level of attitudinal trust by eliciting behavior on the part of B that might count as evidence for A of B's trustworthiness. Let me call such strategies *trusted-directed*. Third, it can take steps to sever the link which exists in the eyes of A between the attainment of and the behavior of B. Let me call such strategies *object-directed*. Let us now return to the types of strategies mooted above.

Let me begin with what I have called object-directed strategies. In such strategies, governments will attempt (to revert to Baier's metaphor of trust as a

<sup>28</sup>See, for example, Baier, "Trust and anti-trust," p. 107. For a more sceptical note, see Russell Hardin, "Trusting persons, trusting institutions," *Strategy and Choice*, ed. R. Zeckhauser (Cambridge, Mass.: The MIT Press, 1991), pp. 185–209 at p. 195.

<sup>29</sup>Offe, "How can we trust our fellow citizens?," pp. 33–6.

web which has been partially damaged) to eliminate the strand between A and B which links them around the realization of the interest of A's concerning which distrust reigns. This can take one of two forms, which I would call *containment* and *replacement*. Containment involves severing the link between A's ability to achieve a goal and B's ability to act so as to inhibit A's attainment of the goal in question, either by ensuring that A possesses all the political levers required to attain the goal in question, or by placing obstacles in the way of B's being able to affect A's pursuit of the goal. This will involve a fair degree of devolution of powers. Replacement involves the state itself taking up the causal role in the attainment of A's end which would be played by B were B considered trustworthy. It thus involves some degree of *centralization*. There are advantages and disadvantages to both variants of this strategy. Containment has the disadvantage of leaving A's attainment of the goal in question more vulnerable than it need be, since the state presumably commands greater power and resources than does A. Replacement has the disadvantage however of opening the state to the charge that it is unduly and unfairly biased toward A. To the extent that it also in principle involves some degree of state centralization, it also runs the risk of giving rise to distrust directed toward it. Group-specific interests often involve the wish not just that an interest be satisfied, but that it be satisfied by the members of the group themselves.

A third variant on the goal-directed strategy is for the state somehow to reduce the salience or importance of the goal in question in the eyes of members of A. Since group-specific goals are often conceptually rather than contingently related to the group-identity in question (that is, it is because it is part of what it means for me to be a member of the group that I have the interest in question, rather than its just happening to be the case that the interest is disproportionately present among members of my group), one way to do this is by what one might call a strategy of identity-promotion, through which the state can attempt to make group-identities themselves less salient, for example, by trying to inculcate a shared, national identity, or by identifying and attempting to activate group-identities which cut across the unity-threatening ones. In terms of the schema I have been working with, this would involve in effect not trying to affect the values of the variables which determine whether A will trust B, but rather short-circuiting the problem by undercutting A- and B-identities. This amounts to attempting to manufacture a national identity from above, a task to which liberal democracies are, for reasons already noted, ill-suited. This is not to say that pursuing *other* trust-promoting strategies might not give rise to new identities. I want to resist the view that identities are primeval givens. History teaches us that identities often crystallize the results of political and economic struggles. But it is to say that the creation of new identities is best pursued by indirection.

Another way of making goals and interests over which distrust has set in less salient, and thus less likely to affect the overall level of trust between groups is through what one might call "trust-spreading." In keeping with what I have

termed the “malleability assumption,” institutions can be designed in such a way as to encourage inter-group collaboration in those policy areas unrelated to the interests which people have *qua* members of politically salient groups, and concerning which significant inter-group distrust exists or is most likely to exist, in the hope of loosening the grip of trust-inhibiting evidence-resistance and of “spreading” trust even to more sensitive areas. If the mechanism operates reliably, then it seems sensible to multiply initiatives bringing the groups in question together in those areas, in the hope that trust will seep through to the areas over which there is significant distrust. Again, these kinds of initiatives must be deployed with a keen eye trained on risks of perverse consequences. Joint initiatives forced upon groups in too heavy-handed a manner can generate counterproductive backlashes. The incentives and sanctions deployed in order to bring the members of groups among which distrust has set in into common political ventures must be calibrated with an eye to this fact.

Let me now move on to what I have called truster-directed strategies. Let me begin by re-emphasizing that we must distinguish between attitudinal trust and the decision to trust. Attitudinal trust registers one’s assessment of the trustworthiness of another person, while the decision to trust is, as it were, an all-things-considered judgment. Given one’s level of attitudinal trust, but given also the options to trusting and the possible benefits and harms that might accrue to one if one decides to trust, it might be rational to decide to trust even if one’s attitudinal trust is quite low. If you tell me to wait at a street corner in a safe neighborhood on a warm and sunny day while you go and fetch a million dollars which you will then hand to me as a gift, it might be rational for me to wait even if I believe that the chances of your being true to your word are vanishingly small.

So truster-directed strategies can be divided between those that seek to alter A’s attitudes, and those that are directed toward changing A’s assessment of the rationality of *acting* on those attitudes. How can the first kind of measure be put into place without directly acting on B’s motives? One possibility, which might seem paradoxical at first glance, is to increase A’s opportunities to monitor B’s behavior. Though we have seen that distrust is stubbornly evidence-resistant, it is possible that the simple fact that institutions make available abundant information about the behavior of others will in and of itself have a trust-enhancing effect. The intuitive point is quite easy to grasp: if the institutions we share with others make it easy for us to observe their behavior, and the way in which that behavior impacts upon us, we will, all other things equal, be less likely to distrust, as whatever the objective grounds we may have to distrust will not be compounded by concerns about what might lie behind lack of publicity and secrecy.<sup>30</sup>

<sup>30</sup>There are obvious similarities here with suggestions made by Braithwaite and Pettit in their essays in *Trust and Governance*, ed. Braithwaite and Levi.



In the second general kind of truster-directed strategy, a third party can affect the stakes and options which, along with attitudinal trust, are relevant to the task of deciding whether or not to trust. Take the most extreme case where a relevant option is secession: a state can tilt the balance in favour of trusting by placing significant obstacles on the road to A's secession. A full discussion of the compatibility of various types and degrees of substantive and procedural obstacles to secession which might be compatible with a liberal democratic political morality falls well outside the purview of this paper. It seems obvious however that, regardless of normative considerations, making secession too difficult can have perverse consequences from the point of view of the pragmatic task of restoring or maintaining trust. What is required are procedures and substantive criteria which will not prevent a group from seceding when distrust is high and pervasive, but which might tilt the balance in cases of more ambivalent groups. A properly framed constitutional right to secession can somewhat paradoxically *enhance* trust by giving groups reason to believe that they are not forever trapped against their wills in an unhappy union, but by defining the procedure which must be followed in order to trigger lawful secession in a way which only makes it worth a group's while in extreme cases.

States can also increase the likelihood that groups will decide to trust by rewarding that decision in various ways. In its crassest manifestation, this will mean tilting the distributive balance in favour of the group in question. Now, unless it is already subject to discriminatory distribution, in which case equalization would be a requirement of justice as well as one of unity, such strategies will quickly run up against problems from the point of view of distributive justice. Melissa Williams has argued that bonds of trust between majorities and historically disenfranchised minorities can be repaired by ensuring the presence of the members of the minority in the society's representative institutions, thus ensuring that they have a robust stake in the decision-making process.<sup>31</sup> Here again, however, care will have to be taken not to grant the alienated group unfair advantages. But there may be some indeterminacy to what fair representation requires which might allow some purchase to considerations of unity. For example, some analysts as well as some political actors in Quebec have argued that, in virtue of being the home of one of the founding nations of Canada, constitutional provisions should be in place which would guarantee the province a certain minimum representation both in the House of Commons and in the make-up of the Supreme Court, regardless of demographic trends. Since first-past-the-post electoral systems already attempt to achieve balance between sheer numerical representation and equitable representation of regions, it does not seem on the face of it unjust to seek to restore trust in this manner.

The third type of strategy which states might put in place in order to restore trust is directed toward the (potentially) trusted rather than the truster. It

<sup>31</sup>Melissa Williams, *Voice, Trust and Memory* (Princeton, N.J.: Princeton University Press, 1998).

involves affecting the motives and/or the behavior of the (partially) distrusted group. The most obvious way of achieving this end is, as it were, to *rig the incentives*, that is, to structure the context of the group's decision to act in a trustworthy manner or not with appropriately targeted sanctions and rewards. I do not agree with Russell Hardin that trust can entirely be thought of as "encapsulated interest," as this would break down the distinction between trust and reliance.<sup>32</sup> I nevertheless do think that the use of incentive-rigging can reinforce existing dispositions toward trustworthiness and disable more narrowly self-interested ones in ways which need not undercut a relationship's status as a *trust-relationship*.

Incentive-rigging can take the form either of sanctions, whereby penalties are attached to behavior likely to fuel the distrust of members of A, or positive inducements, whereby trust-enhancing behavior is rewarded.

The following general concerns must be kept in mind when recourse is made to incentive-rigging. First, as I mentioned above, distrust, like trust, is stubbornly evidence-resistant. Eliciting trustworthy behavior from the distrusted group might because of this not have the desired effect. There is still the risk that distrust runs so deep that objectively trustworthy behavior will be interpreted with suspicion by the distrusting group. Second, partly as a result of this evidence-resistance, group A's distrust of group B might not always be justified. That is, the mere fact that the members of group A feel distrust toward the members of group B over some issue-area is not in and of itself a sure sign that lack of trust is, objectively speaking, warranted. Government action seeking to secure trust must always be balanced with considerations of fairness and reasonableness. Sanctions directed against Bs and aimed at restoring A's trust might end up alienating Bs if it felt that they go well beyond what is required to repair the relation of trust. The stubbornness of A's distrust must not become a reason to make governments hostages to the cause of restoring their trust.

Second, incentives and sanctions aimed at restoring trust must be designed in such a way as to *support* rather than to *supplant* trust-sustaining motives. The governing idea is that in divided societies which have not yet descended into group hatred or civil war, we are possessed of mixed motives toward other groups. The task of trust-enhancing institutions will be, as it were, to work with the trust-supporting motives that are already in place, so as to neutralize the motives that would tend to make one act in an untrustworthy manner. What is to be avoided is providing people with coarse-grained motives which simply get them, through the operation of sanctions and rewards, to *act* in the same way as they would have done had they been motivated by the trust-supporting motives. We do not want, that is, to replace trust with mere reliance. How might

<sup>32</sup>The following quote captures Hardin's view well: "Most of us are somewhat like bank tellers: we are secured in our normal honesty by institutional arrangements that make significant dishonesty risky, even difficult. Much of what looks like honesty is essentially self-interest at work." It is taken from "Trusting persons, trusting institutions," pp. 203-4.

institutions operate so as to avoid this? Assume that people have *higher-order motives*, that is motives regarding which *first-order* motives they would want to see operative in their actions.<sup>33</sup> What we want is for institutional arrangements to make it more difficult or costly to be motivated to act in accordance with trust-eroding than with trust-promoting motives. What I have in mind here are the kinds of reputational sanctions described by republican theorists such as Philip Pettit. The main idea behind Pettit's argument<sup>34</sup> is that trustworthiness can be elicited from others by acting in a trusting manner toward them, thus raising the cost of acting in a non-trustworthy manner to include losses to reputation and self-image. People by and large attach value to being well thought of, and disvalue to the sanction of loss of honour or reputation. Institutions which bring to the fore the costs in terms of self-image of untrustworthy behavior will thus be geared in the requisite fine-grained manner toward getting people to act in a trustworthy, and not simply in a reliable, manner. It might be that in order to elicit cooperative behavior from people possessed of mixed motives toward members of other groups, we must project back to them through institutional arrangements an image of themselves as acting on the higher-order motives which are part of the self-image which is required to promote trust. This would militate for sanctions geared toward higher order motives, and against sanctions geared to uncooperative *behavior*.

I have provided a brisk overview of some considerations compatible with a broadly liberal political morality which might inform institutional design in divided societies. Not all of these norms are compatible with one another, and specific contexts will dictate which particular mix of institutional devices might be most appropriate to specific cases. The latter are offered more as an invitation to philosophers to take up the challenge of imagining creative institutional solutions to problems of disunity, rather than as any kind of a final word.

## V.

One final question remains to be examined. The foregoing section has examined various ways in which governments can design institutions so as to create an incentive structure which is best suited to fostering trust in societies in which it has been badly frayed in a manner likely to threaten unity. We have been looking for ways in which the state might, as it were, "stage-manage" trust. An obvious question which arises from this enterprise is the following: are we still talking about trust at all? Is a person who has to rely upon sticks and carrots in order to act in a trustworthy manner really trustworthy at all? Is it not the case that in

<sup>33</sup>The idea of a higher-order motive is drawn from highly influential writings by Gerald Dworkin, Harry Frankfurt and Charles Taylor.

<sup>34</sup>The argument is developed principally in Pettit, "The cunning of trust."

rigging the incentive structure in a sufficiently clever way, we have simply succeeded in making him *reliable* rather than *trustworthy*?<sup>35</sup>

Some philosophers who have written about trust clearly see it that way. For example, D. O. Thomas has argued that “where I trust someone I do not resort to any inducements or threats to ensure that he does it. I depend upon his choosing freely to do what I trust him to do.... [T]rusting someone to do something does, logically, exclude sanctions.”<sup>36</sup> I believe that this is overstated. The argument rests on a false dichotomy: either we act in a trustworthy manner freely and without the support of any institutional sanctions or rewards, or we are simply being coerced to act in a way in which, absent the rewards and/or sanctions, we would have had no inclination to act. There is an intermediary case, one which, I think, corresponds much more closely than either of these extremes to real-world relations of trust. We have mixed motives in most situations. In interpersonal relations, only angels are devoid of self-serving, uncooperative motives. We often want to act on the “better” motive, but the tug of the more suspect sometimes impresses itself urgently. In such situations, it is a sign of the high regard which we have for the reasons which the better motive would have us act on that we put ourselves in situations which disable the more suspect motive, or which make it more costly to act upon. The point is familiar from Jon Elster’s work on “self-binding mechanisms”<sup>37</sup>: we want to want to act on X, but we are afraid that when the time comes, we will find ourselves succumbing to the temptation to act on Y. We therefore take steps now to ensure that, at that future point, Y will be difficult or impossible to act on. We can also take steps to make it more likely that our higher-order motive only to act in trust-supporting and cooperative ways will be operative. To the extent that we can non-self-deceptively see the self-binding mechanism as a product of our own agency, it is a manifestation of our devotion to the reasons X tracks, rather than a revelation of the fact that we do not take these reasons seriously at all, since we must be dragged to realize them kicking and screaming.

The institutional devices which make relations of trust more likely to be sustained or restored can be seen in a similar fashion. To the extent that we can view them as emanating from our own political agency, they bear witness to the seriousness in which we hold the goal of sustaining and strengthening relations of trust with our *concitoyens*. That seriousness is marked by the fact that we want to ensure that we will act in a trustworthy manner toward one another, rather than allowing the quality of our relations to become a hostage to fortune. Provided that they not be the result of pure imposition from on high, the institutions, laws and policies we devise in order to promote trust by “rigging the

<sup>35</sup>For the difference between reliability and trustworthiness, see Richard Holton, “Deciding to trust, coming to believe,” *Australasian Journal of Philosophy*, 72 (1994), 63–76.

<sup>36</sup>D. O. Thomas, “The duty to trust,” *Proceedings of the Aristotelian Society*, 79 (1978–79), 89–101 at p. 91.

<sup>37</sup>Jon Elster, *Ulysses and the Sirens*, rev. edn (Cambridge: Cambridge University Press, 1984), ch. 2.

incentives," far from detracting from the ensuing social relations being perceived as trusting ones, are actually expressive of our allegiance to the goal of trust.

## VI.

I have tried in this paper to argue that trust, rather than cohesion, shared identity or cooperation, is the core constituent of social cement. Especially in the case of societies for which secession is a real option, it is trust which in my view prevents societies from falling apart in times of crisis when what many people might feel the society's *raison d'être* is called into question. I have also attempted to bring out the independent value of relations of trust. Though it is hoped that in most societies, trust will simply lay the groundwork for more full-blooded types of relationships, and thus in a sense disappear from view, it is also true that there is a value to trust which cannot be reduced to the value of any of these more robust relationships. To wax post-modern (or Levinassian), it is in the relations held together only by trust (rather than by shared identities, values or ends) that we bear witness to the respect that is due to the radically other. And the dimension of trust that persists even in our more full-blooded relations should not be underestimated: it is important that our being well disposed and respectful of other people not be *entirely* due to the fact that we share some interest, identity or way of seeing the world.

What is true of interpersonal relations is also true of politics: though polities marked by a robust shared identity, by large-scale cooperative enterprises and/or by a shared conception of the good life are admirable and enviable things, so are societies capable of remaining united despite deep differences. Though the deep differences dividing some societies might be more robust and difficult to dislodge than those marking others, there can very well come a time in the life of *any* political association when its survival requires that it be able to fall back on a sufficient fund of trust. I hope in this paper to have contributed to the task of understanding what trust is, how it can break down, and what might be done to restore it.