

Building Trust in Divided Societies

By
Daniel Weinstock

Political philosophy in the (very broadly understood) liberal tradition has, at least since the publication of *A Theory of Justice*, focussed overwhelmingly on the principles and institutions which must be put in place in order for *justice* to be achieved. Scant attention has been paid, however, to the conditions underpinning the *unity* of modern societies, in particular of culturally divided societies. Yet events of the last 10 years or so, both in the former Soviet Empire and in the West, make it clear that the problem of unity should occupy as much of a place on the political philosopher's agenda as that of justice. Countries east of the Rhine have been falling apart at a vertiginous rate, giving rise to unspeakable human misery and suffering. And the bonds which hold together traditionally more peaceful and just countries such as Canada, Belgium, Spain and Great Britain have been severely frayed in recent years. In this context, it makes sense to explore the resources which philosophy might possess to help uncover what the source of social unity in divided societies might be. I believe that philosophers can be useful in this field of inquiry in two ways. First, they can contribute conceptual clarification which, as I will show in the next section, is much needed in this area. And second, they can contribute to the normative task of determining what states can do to promote social unity without overstepping the normative bounds laid out for them by their favored conception of justice. (I will assume for the purposes of this essay that this will mean a broadly liberal political morality). This essay makes a tentative first stab in both these directions.

Some Conceptual Clarification

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 are all relevant to the overall task of discovering and analysing the nature of what Jon Elster has called the "cement of society".¹ It is in my view essential in the present context not to conflate four descriptive notions which I will label unity, stability, cohesion and cooperation. *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

¹ Jon Elster, *The Cement of Society. A Study of Social Order*, (Cambridge: Cambridge University Press, 1989).

any desire to sever the existing bonds of political association. *Stability* has to do with a society's ability to ensure that the principles of justice which structure its main institutions will reliably secure the consent of subsequent generations of its citizens. *Cohesion* has to do with the extent to which a society's members share moral, religious, cultural values 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. They can be empirically related, but in all sorts of complicated ways which put paid to any attempt at formulating general laws governing their interactions. Particular circumstances have a tremendous impact on the ways in which they relate to one another. In some cases, practices of cooperation can go some way in explaining the resilient unity of a society. But in other sets of circumstances, unity might depend upon at least temporary *lack* of cooperation. Where group enmities within a society run deep, for example, unity might, at least during certain periods of a society's history, require that members of different groups *not* be pressed into common enterprises beyond what is required for the viability of the society's main institutions. In some societies, cohesion might contribute significantly to stability. One can imagine a society in which shared political values flow directly from -- or are already included within -- an overall shared conception of the good. In others, the relationship between general conceptions of the good life and principles of political morality might be much more tenuous. For much of its modern history, Italians shared a thick religious conception, but it has seemed compatible with quite a broad range of political forms.² And so on. The general point is that there don't seem to be very many interesting generalizations to make as to the empirical relations between these various properties. Understanding how their instantiations (or lack thereof) will interact causally will always be a case by case affair, and the understanding of any individual case will depend upon retaining a keen sense of the separateness of these properties.

There is moreover a practical political reason to ensure that conceptual clarity reigns here. As we know from the field of economics, conceptual error can in some circumstances become self-fulfilling prophecy. Thus, for example, though there is no theory-independent reason to suppose that the conception of human rationality favoured by economists and game theorists exhausts the concept of human rationality, there is the risk that it will be built into our institutions and our popular culture to such a degree that people will begin to act *as if* it did. They might for example

² On the great variety of political culture both historically and at present in Italy, see Robert Putnam, *Making Democracy Work*, (Princeton: Princeton University Press, 1993).

make themselves unable to access the mutually beneficial outcome in PD-like situations which may have been pre-theoretically accessible to them.³

Similarly, confusion between the properties I have sketched can have damaging results. For example, thinking that (what I have called) unity requires that (one of) the more exigent properties such as cohesion, collaboration or stability be instantiated in a society might give rise to the sense among members of the society in question that the unity which their association does in fact instantiate is a kind of sham, second-best unity, that it is merely a pale copy of the much more robust original, rather than possessing a property independently desirable in its own right. I will return to this point below.

The confusion which in my view permeates discussions of this issue can be seen in the fact that the principal strategies which those political philosophers who have to this day devoted some attention to the problem of unity have deployed in order to give an account of the grounds of social unity have all sought to account for unity in terms of one of the three other C-properties. Let me call these strategies, for the sake of easy reference, participatory, nationalist, and constitutional.

The participatory route to social unity for societies divided along the type of ethnic and cultural lines most likely to lead to autonomist or separatist pressures is based on a rather simple and apparently plausible claim. The general idea is the following:⁴ in order to counteract the centrifugal force of cultural, ethnic, religious and other identities, contemporary pluralistic societies must promote the practices associated with the one role which we can all share despite our differences, that is, the role of citizen. According to this view, the view of citizenship espoused by liberals has been too passive to ground any strong identification. Citizenship has been defined primarily in terms of the rights citizens hold against one another and against the state, rather than in terms of shared roles and practices which might constitute a thicker identity, one more likely to at least counteract more divisive ethnic and cultural elements of our identities. Different authors have in this context focussed on different kinds of activities related to a normative conception of citizenship of this kind, ranging from deliberation on matters relating to the public good to public service in the pursuit of that good as contributing to the requisite “social cement”.

In terms of the typology developed above, unity is here being conflated with cooperation, or more charitably, cooperation is seen as contributing reliably to unity. Yet there are good reasons to suppose that the mechanisms hypothesized by Miller and others to underpin this connection are no more plausible than the contrary hypothesis.⁵ Given that we live in mass societies, the practice

³ On this point, see Martin Hollis, *Trust within Reason*, Cambridge: Cambridge University Press, 1998, pp. 154-159.

⁴ A clear recent statement of this view can be found in David Miller, “Citizenship and Pluralism”, in *Political Studies*, vol. 43 (1995).

of citizenship can, it seems, take one of two forms. Either it is focussed upon questions of the public good of the whole society, in which case it is difficult to see how participation (which, realistically, will for most people in this context take the shape of voting) can be anything more than the pallid (which is not to say unimportant), non-identity-engaging activity republicans and communitarians have always suspected liberal accounts of citizenship to be reducible to. Or it is focussed on activities and practices centred around the kind of more limited shared good that one finds at the heart of the associations that make up civil society. And here, it seems antecedently just as likely that participation in such associations will constitute a “school for democracy” which will have positive knock-on effects for the unity of the broader society, as that such participation will have deleterious effects from the point of view of social unity, for example by awakening or exacerbating social divisions, or by making participants in such associations attribute disproportionate importance to the limited good of their association as against the larger public good.⁶

The nationalist strategy claims 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.⁷ This can be achieved either through *nation-building*, whereby the state takes on the task of forging a new unified national identity which would supplant the diverse national identities already present on a given territory, or through *separatism*, whereby multinational states are broken up so as to have the boundaries of the state correspond more closely to the boundaries of the nation. Despite the fact that they adopt what at first glance appear to be quite different routes in order to address the problem of unity, separatists and nation-builders are faced with a common problem. Unless they are willing to engage in the kind of ethnic cleansing that has sadly been seen all too often in recent years, separatists will find themselves replicating the problem they had sought to avoid in that they will inevitably find ethnic and cultural minorities in the midst of the smaller political entity they will have created. Thus, separatists wanting to generate a shared national identity will also have to be nation-builders.

⁵My understanding of the notion of a “mechanism” as an explanatory device occupying an intermediary between pure description and the kinds of iron-clad, deductive-nomological laws which logical positivist philosophers of social science had hoped to discover is derived from the work of Jon Elster. See particularly his *Political Psychology*, (Cambridge: Cambridge University Press, 1993); and “A Plea for Mechanisms”, in Peter Hedstrom and Richard Swedberg (eds.), *Social Mechanisms*, (Cambridge: Cambridge University Press, 1998). One way to understand the distinction between laws and mechanisms is that mechanism-based explanation eschews the aspiration to prediction.

⁶For the optimistic view, see Stephen Macedo, [check ref]. For the less sanguine view, see Will Kymlicka and Wayne Norman, “The Return of the Citizen”, in R. Beiner (ed.), *Theorizing Citizenship*, (Albany: Suny Press, 1995).

⁷See, for example, Will Kymlicka, “Social Unity in a Liberal State” , in *Social Philosophy and Policy*, 13 (1996). He writes (p. 131) that “people decide who they want to share a country with by asking who they identify with, who they feel solidarity with”.

Nation-building can be carried out more or less liberally. The kind of nation-building which gave rise to modern European nations such as France, Germany and Italy was carried out quite illiberally, riding roughshod over local cultures and languages, often making the teaching of non-metropolitan languages illegal. 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” theories which have in recent years been defended by Yael Tamir, David Miller (though he does not explicitly call himself a liberal), and Will Kymlicka (though he does not explicitly call himself a nationalist).⁸ 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 and minority identities. Miller for example describes a process in which a majority national group “waters down” the culture of the majority cultural group so as to allow immigrants and national minorities to take up the identity in question and to contribute to its future development. Again, though one can imagine mechanisms whereby this process might succeed in creating a unity-promoting shared identity, one can also observe that the process has historically failed, for reasons which can readily be explained. Indeed, the “watering-down” process I have just sketched will typically involve the stripping from the majority national identity of what, for want of a better term, I would call its “experiential” dimension. I refer here to those aspects of national identity which, beyond the values at the basis of social institutions, the main lines of the nation’s history and its language, which can be explicitly stated and taught, conatively anchor these elements of an identity in an individual’s psychological make-up. These are the kinds of unformulatable latent understandings to which communitarians refer when criticizing putative liberal views of the self, and which are inculcated not through any explicit teaching, but through the rituals and practices which make up a “thick” community.

Constructing a national identity capable of unifying a culturally diverse citizenry will often mean ridding it of its experiential dimension. The attempt at creating a “civic” nationalism in Quebec for example has centered around the attempt at delineating a “shared public culture” which might be taken up at will by immigrants and members of cultural minorities.⁹The problem with such an

⁸Yael Tamir, *Liberal Nationalism*, (Princeton: Princeton University Press, 1993); and David Miller, *On Nationality* (Oxford: Oxford University Press, 1995).

attempt from the point of view of the problem of social unity is that, on the one hand, the elimination from a national culture of its deeply particularistic and unshareable aspects is more difficult than Miller and Quebec nationalists have supposed, and its imposition will therefore often be met with hostility, rather than embraced, by members of cultural minorities; and on the other hand, the thinning out of a national culture beyond a certain point is seen by some of the members of the majority culture as undercutting the very *raison d'être* of nationalist politics, and so risks generating a Maurrasian backlash on the part of those who would put forward a much thicker, and thus more exclusivistic, version of the majority national identity.¹⁰

The problem here again has to do with the conflation of two of my C-properties, or perhaps, with the overly optimistic view that the instantiation of one, in this case cohesion, will reliably contribute to another, unity. Here the problem in my view is not so much, as in the case of collaboration and unity, that the mechanism connecting them is not as reliable as some theorists have claimed them to be, but rather that the achievement of the kind of cohesion which nationalists aspire to is much more difficult (for the reasons just adduced) than is the attainment of unity, and though the achievement of a shared national identity is probably a sufficient condition of unity obtaining, it is (or so I will argue) certainly not necessary. The practical risk involved in not seeing this has already been alluded to above: it is that people will come to believe that they are insufficiently united as a society if they are not cohesive, and that this false belief will become self-fulfilling.¹¹

The final, “constitutional” strategy for unification which I want to mention can be seen as a special case of the nationalist strategy. The attempt here is not to attempt, as it were, to “compete” with the thick cultural and ethnic identities of a multicultural society’s constituent communities, but rather to subordinate them to a “principled” identity centered around abstract moral and political principles which might be included in a constitution. This strategy has been central to pro-European European intellectuals such as Jean-Marc Ferry and Jürgen Habermas, though it recurs among some Anglo-American philosophers who have bemoaned the turn which “political liberals” have

⁹For a discussion of this phenomenon in the context of recent political debates in Quebec, see my "Le problème de la culture publique commune", in F. Blais et al. (dirs.), *Libéralismes et nationalismes*, (Ste-Foy, Presses de l'Université Laval, 1995).

¹⁰I have taken these points up at greater length in my "Is there a moral case for nationalism?", *Journal of Applied Philosophy*, 13, 1996.

¹¹I would argue that Canada’s current unity woes are due to Pierre Elliott Trudeau’s belief that Canada needed to be cohesive in order to be united, and to the policies he put into place to attempt to achieve cohesion. These measures led to regionalist backlashes in Quebec and in the Western provinces, which are still being acutely felt to the present day. For Trudeau’s writings, see *Federalism and the French Canadians* (Toronto: MacMillan, 1968). For important interpretations of Canadian political history along the lines I have suggested, see Guy Laforest, *Trudeau and the End of a Canadian Dream*, and Kenneth McRoberts, *Misconceiving Canada*.

taken away for the need for shared values.¹² The hope is that allegiance to universalistic values will attract a stable identification capable of counteracting the divisive impact of nationalisms and localisms.

The idea that shared political values might reliably contribute to or constitute unity is however an instance of what Wayne Norman has aptly termed the “ideology of shared values”.¹³ Let a Canadian example serve to illustrate the difficulty: with the waning of the influence of the Catholic Church in Quebec, the values of English-Canadians and Quebecers have arguably never been closer than they are today. The Charters of Rights entrenched in both provincial and federal Constitutions differ in detail only.¹⁴ Yet this has obviously not correlated with increased unity. And there is evidence to suggest that the same is true of the relation between Catalans and other Spaniards, and between Flemish and Walloons.¹⁵ There is just no evidence to suggest that people form new political associations because of a sense of shared values, or that secessions are best explained because of a *lack* of shared values.

We have thus seen that 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.

Perhaps these philosophical attempts at accounting for social unity are at root all manifestations of the same flawed assumption: it consists in the idea that what might in principle motivate individuals or groups to contract into a new political association also explains their remaining together in existing political associations. Let me refer to this as the *symmetry assumption*, to evoke the idea that according to this assumption, the reasons for opting in are the same as the reasons for not opting out. I believe that this assumption is mistaken. Briefly, though it is clear that people would have to be presented with reasons of a certain kind in order to view as legitimate their integration into a larger political unit (increased economic cooperation and greater

¹²See for example the essays collected in Jacques Lenoble and Nicole Dewandre (eds.), *L'Europe au soir du siècle* (check ref.). See also Lawrence E. Mitchell, “Trust and the Overlapping Consensus”, in *Columbia Law Review*, 94 (1994).

¹³Wayne Norman, “The Ideology of Shared Values”, in Joe Carens (ed.), *Is Quebec Nationalism Just?*, (Montreal: McGill-Queens Press, 1995).

¹⁴They also differ in that the Canadian Constitution includes provisions which were explicitly framed to render some of Quebec’s traditional jurisdictional claims, having to do particularly with language, unconstitutional.

prosperity, national reunification, shared values, and the like), it is unclear that what accounts for people choosing to remain within existing political associations, or for their not putting their political association in question, is that the reasons which may have explained its coming into being at some point fail to obtain. In other words, people do not secede because what may have been the original *raison d'être* of the political association has somehow lapsed. The creation of common institutions, and the habits and practices born of sharing a life under such institutions, generate patterns of expectations and interest unrelated to the initial *raison d'être* of the political association, a kind of inertial momentum which make it the case that the reasons people have to continue living under shared institutions might end up being quite different from those which lay at the basis of their union. The fact that they do not question the very existence of the political union need not be related to the fact that the reasons they may have had for joining are still in place. There will be occasions in the life of any collectivity when it will not fully live up to the principles which for some people might constitute its fundamental rationale. But I would claim that the viability of a political association depends upon its generating endogenous causal mechanisms that have a tendency to relegate foundational questions into the background. Political associations probably require for their tolerably smooth functioning that citizens not be inclined constantly to check to see whether they continue to instantiate their initial *raison d'être*. And I would argue that, in most circumstances, this is a good thing point, not because of *Realpolitik* concerns with international order and stability, but because (or so I would want to claim) there is some positive *value* to a society's being sufficiently united that, even when there is debate over fundamental issues, or when crises occur which place some element of the society's *raison d'être* in question, that (provided that the crisis in question lies within certain bounds) this not lead to people immediately deciding to quit the association in question.

Let me put the point in another, starker way. Imagine two political associations, A and B. Both A and B have been formed, or are understood by their members as having been formed, in order to pursue some good or to instantiate some valued property of political systems as a whole. Imagine now that both societies are beset by a crisis during which the society's ability to realize this foundational good is called into question. We can imagine that it fails to achieve promised levels of economic prosperity, or that it fails to instantiate the lofty moral values many of its members, or that it fails to achieve promised national greatness, or whatever. Let's stipulate, moreover, that secession is a real option in both cases, that is, that the associations result to begin with from the merging of two or more previously self-governing societies, or that divisions within the society occur along regional lines which might form the bases of *future* self-governing societies (call this the *Padanian* scenario). Whereas A weathers the storm and remains united, B's citizens become

¹⁵See Will Kymlicka, *Finding Our Way. Rethinking Ethnocultural Relations in Canada*, (Oxford: Oxford University Press, 1998).

disposed to quit the association. I want to claim that one possible explanation of this lies in the fact that A possesses a property which B does not possess, and which we might call *social trust*. I want in what follows to make four claims about trust as I understand it.

- 1) It is importantly distinct from the kinds of properties that ground cooperation, cohesion or stability.
- 2) It has independent value. That is, its value cannot simply be expressed in terms of the value of other C-properties.
- 3) There are steps which governments can take in order to sustain or generate trust which are not identical to the measures they can take to sustain or generate any of the other C-properties.
- 4) Relations of trust which sometimes have to be sustained by the kinds of institutional mechanisms which states are best situated to put into place are not intrinsically less valuable than trust-relationships in which trust is sustained, as it were, spontaneously.

I will address these points in turn.

The Distinctiveness of Trust

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 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 however does not require sharing in this way. The basic structure of the trust-relation brings this out clearly. It is common ground among philosophers writing about trust that trust is a three-part relation: “A trusts B to do ”, or “A trusts B with valued thing C”.¹⁶The central aspects of this relation are, in my view, the following:

- i) it is an asymmetrical relation in that the truster is rendering himself vulnerable to the trusted in circumstances where it is uncertain that the trusted will not exploit the truster. (It seems odd to say that I *trust* you to do X if I know with certainty that you will do X).
- ii) 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 don’t (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

¹⁶This much unites authors who have otherwise very different conceptions of trust. See for example Annette Baier, “Trust and Antitrust” in *Moral Prejudices*, (Cambridge, Mass.: Harvard University Press, 1994), p. 101, and Russell Hardin, “The Street-Level Epistemology of Trust”, in *Politics and Society*, 21 (1993), p. 505.

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 to whom I ask for 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.

iii) the truster and the trusted are differentially situated with respect to the action X or the valued thing C. Though this aspect of the trust relation has not been much commented on, I think it is required to maintain the distinguishing properties i) and ii), and thus to ensure that the relation in question is in fact one of trust. Indeed, were it the case that B had as much self-interested reason to do X or to care for C as A has to want X to be done or to have C cared for, then the asymmetry in the relation would be significantly diminished, as A and B would both have as much of an interest in X occurring or C being cared for, though residual asymmetry might persist in that B might have greater *ability* to bring the desired end about. And were the interests of A and B not at stake to different degrees, then B might be undertaking the required action for entirely self-interested reasons.

It is important to note that these three aspects of trust relationships are connected. My vulnerability toward the person to whose care I am entrusting something stems from the fact that I require that she aid or forbear not because she has an independent interest in the object of my care, but because, in some minimal sense, she is morally related to *me* rather than to that which I entrust.

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, and which indeed might be seen as depending upon our *not* sharing one or the other of these. And I would argue that it is this kind of relation which binds us in our everyday dealings with those fellow citizens with whom we are not bound as well by some thicker tie, be it familial, professional, religious, etc. In these latter cases, it is entirely likely that one of the characteristics i)-iii) will fail to obtain, or will obtain only imperfectly, so that our relation with them will not be one of unalloyed trust. The relation in which I stand to the stranger in the street who gives me directions, or to the fellow library-user who I encounter in the darkened stacks and who I trust not to harm me, seems adequately characterized by i)-iii), and certainly does not require anything as lofty as value or identity-sharing or as cooperative interaction.

My claim is that a minimally well-functioning society is one in which citizens' (as it were) default attitude toward one another is one of minimal trust in this sense. When, in a minimally healthy society, we are brought into interaction with our fellow-citizens in such a way as to put an interest at stake, we trust them in the minimal sense described above. Our trust not being betrayed might lead to an enriching of our relationship (cooperation, friendship, and the like). But it is a

mistake to think that fellow-citizens must be linked at the outset by more than trust in order to be able to say of them that they form an adequately united society.

The Independent Value of Trust

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 (in terms of the example given above, it is the presence or absence to a sufficient degree of trust in this sense that determines whether we find ourselves in society A or society B), and it is on the basis of a fund of trust that more morally exigent and rewarding relation can be built. The overtures which people make to one another on the road to friendship 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 forebear 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.¹⁷ 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 forebear 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 forebear.

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,¹⁸ I would claim that there is value to 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

¹⁷Immanuel Kant, *Groundwork of the Metaphysics of Morals*, (trans. H.J. Paton), (New York: Harper Torchbooks, 1953), p. 11 (AK. 398-399).

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).

Social Trust as a Goal of Political Action

I have argued thus far that trust among citizens is a relationship distinct from those which are born of a shared identity, of cooperative enterprises or of a common allegiance to moral and political principles. Whereas these latter require that values, ends and/or identities be shared by citizens, trust does not. It is thus perhaps more useful as an account of what holds already existing societies together even when crises occur which imperil these “thicker” components of social cement. I have also argued that there is independent moral value to trusting relationships. I have claimed, moreover, that it is essential in political contexts that trust not be conflated with the other C-properties. This is because promoting one of them in the belief that one is thereby promoting trust can have perverse consequences. Attempting to instill cooperation, or to impose a shared identity, upon a society in which trust has been damaged or is in insufficient supply can exacerbate divisions by placing pressure upon social relations that they cannot support, and/or by contributing to the belief that a political association which is not characterized by cohesion or cooperation is not worth having at all.

The upshot of this for political philosophy is that the sustaining or fostering of trust is, all other things equal, a valuable goal for political actors to strive for, one which cannot simply be equated with the goal of promoting cohesion, stability or cooperation. The question that now faces us is the following: is there anything that governments can do, within the normative bounds set by political morality, to contribute to trust, and in particular, to repair trust where it has been damaged?

Before answering this question, I need to make two prefatory points. First, I must anticipate an obvious objection. Surely, it might be claimed, I have in moving from the interpersonal to the political context, illicitly changed the subject. The trust that characterizes our more micro-level interactions is, it might be argued, simply different from that which in well-functioning societies characterizes the citizen’s relation to the state and to his fellow-citizens taken as a whole. To revert to a distinction made in much of the sociological literature on trust, the latter is

¹⁸On this point, see Philip Pettit, “The Cunning of Trust”, in *Philosophy and Public Affairs*, vol. 24 (1995); and Philip Pettit, “Republican Theory and Political Trust”, in Valerie Braithwaite and Margaret Levi (eds.), *Trust and Governance*, (New York: Russell Sage Foundation, 1998), pp. 304-308.

on the order of “generalized” trust the object of which is an entire institutional structure, while the former is a kind of familiarity characteristic of smaller groups.¹⁹

I want to claim that the distinction upon which this objection would be based is overdrawn. It may look, from the sociologist’s external observational standpoint, as if societies have emergent or supervenient systemic properties which cannot simply be reduced to the actions of individual agents. From this it is but one step to the conclusion that the object which citizens take their trust to have, when they claim that they trust or distrust their government or their fellow-citizens, has these properties as its object. But I think that this would be to conflate the participant and the observer perspectives. Regardless of whether the sociologist is right in thinking of society as a system with non-reducible properties, individual agents view their trust as applying to other individuals. When an individual claims that she trusts her government, she means that she believes that, when and where her interests come to depend upon some sub-set of agents that make up the government, they will, inkeeping with their specific roles within government, act benevolently, or at least non-malevolently, toward them; when she claims that she trusts her fellow-citizens, she means that, antecedently of any particular knowledge she may have of any one of them, she expects them, taken individually, not to be ill-disposed toward her in their dealings.

Another way of making the same point is to say that individuals are never simply determined in their every action by systemic imperatives, or more specifically by the roles they occupy within a social system. Whatever the constraints and incentives which such roles set up, there is always room for them to fulfill them self-servingly, maliciously, benevolently, responsibly, etc. Trust in government and trust in fellow-citizens apply not to roles conceived of as impersonal parts of a system, but to the manner in which these roles are filled by individuals who always retain some room for maneuver within any set of role-constraints.²⁰

What’s more, I would argue that, though it may be important to keep them distinct for analytic purposes, what I have called trust in government and trust among citizens interact importantly. 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. This is arguably the kind of mechanism which has given rise to “Padanian” nationalism in Italy.

¹⁹See, e.g., Adam Seligman, *The Problem of Trust*, (Princeton: Princeton University Press, 1997), pp. 120-123. The idea of trust in the modern world as applying to social systems as a whole also marks the work of Niklas Luhmann. See his *Trust and Power*, (New York: John Wiley and Sons, 1979).

The second preliminary remark I want to make is the following: Remember that we are primarily²¹ concerned with societies in which secession is a real possibility, in which, therefore, political unity as I have described it is not simply a given. Our focus must therefore be upon tears in the fabric of trust which overlap with those salient lines on the political landscape most likely to threaten unity. Note however that those lines are not always pre-existing ones. Distrust within a society can give rise to new territorially based political identities. It is a mistake to see political identity as in all cases an independent variable. The Northern Italian proto-nationalism of the Lombard League is only the most obvious recent case of a political identity having been created of whole cloth as a result of distrust born of economic factors. Quebec nationalism is also at least in part a political creation rather than a primeval, unanalysable force. Indeed, Quebecois identity, based on territorial belonging rather than language, religion and ancestry, is a fairly recent arrival on the political scene. The French-Canadian identity that preceded it was much more territorially dispersed, and encompassed many French-speaking Canadians who did not reside on the territory of Quebec, while excluding many non-French-speaking residents of the province. It was only when distrust between French and English political elites reached a certain point that the former thought it necessary to give a territorial -- and thus more easily politically acted upon -- basis. This had a deeply transformative effect upon the traditional aspirations of French-Canada, which I cannot elaborate here. The lesson to be learned is that we can not always easily guess where unity-threatening rifts will occur. They are not “naturally” given, and can emerge in surprising places when trust breaks down to a sufficient degree.

Let me begin to answer the question of what governments might do in order to sustain or repair trust by adding to the very simple, three-place, picture of trust I traced above.²² In line with authors as different as Hardin and Baier, I said that trust is a three-place relation which can be schematized as “A trusts B to ”. Though this is the schema which best describes what I have been

²⁰The point is similar to one made by Pettit in “Republican Theory and Political Trust”, pp. 296-299. See also in this connection Annette Baier’s “Trusting People”, in her *Moral Prejudices*, p. 193.

²¹Primarily rather than exclusively, because trust is clearly of great importance even in societies not threatened by break-up. The poignancy of the situation of mistreated but territorially dispersed minorities such as African-Americans stems from the fact that many of them rightly distrust the American political system but that they have no realistic exit options. For an illuminating account of the relation of white and black Americans seen through the lense of the issue of trust, see Melissa Williams, *Voice, Trust and Memory*, (Princeton: Princeton University Press, ch. 5. Perhaps we must be slightly more expansive than I have been in our characterization of unity than I have been thus far in this paper. We might want to allow that the kind and level of political disaffection and demobilization which characterizes African-Americans is, absent the real possibility of secession, a sufficient marker of fractured unity.

²²Cf. Simon Blackburn, “Trust, Cooperation, and Human Psychology”, in Braithwaite and Levi (eds.), *Trust and Governance*, 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

calling “bare” trust, 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 examples of asking for directions and encountering a stranger in a library stack bear 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 \dots ”, 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).

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.²³ One's attitude of trust can range from 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.²⁴ 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.²⁵

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 \dots .

(We can imagine *n* as ranging between -1 and 1, where -1 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

²³The 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”, in Braithwaite and Levi (eds.), *Trust & Governance*; and Barbara Misztal, *Trust in Modern Societies*, (Oxford: Basil Blackwell, 1996), pp. 95-101.

²⁴These issues have been discussed by Russell Hardin, “The Street-Level Epistemology of Trust”, in *Politics and Society* 21 (1993).

about a person will affect the value of n . 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.

The schema I have proposed is probably still far too simple. For example, one would want to work past experiences of entrusting things to a person into the schema. Past trust that has been betrayed will reduce the value of n , and perhaps also narrow the range of n . 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 points at which third parties, including governments, can intervene so as to increase the likelihood that trust-relationships will hold. Very roughly speaking, 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 it can attempt to affect A's level of attitudinal trust by eliciting behaviour on the part of B that might count as evidence for A of B's trustworthiness. Let me call such strategies *trusted-directed*. Or, again, it can take steps to sever the link which exists in the eyes of As between the attainment of n and the behaviour of B's. Let me call such strategies *object-directed*.

Let me make two preliminary points, borrowed from Annette Baier, before going into more detail about these different possible modes of trust-enhancing government action. First, let me assume that trust is *non-specific*, by which I mean that, *ceteris paribus*, if A trusts B to do a specific thing, and if B responds to B'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".²⁶ Call this the *contagion principle*. 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".²⁷ Call this the *principle of realism*.

I would claim that unity-imperilling distrust occurs when groups within a society come to coalesce in such a way that members of group A come to form the belief that members of group B are not trustworthy relative to the care of some good that members of A feel central to their well-

²⁵That 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", in *Ethics*, 107 (1996).

²⁶Annette Baier, "Trust and its Vulnerabilities", in *Moral Prejudices*, p. 149.

²⁷Annette Baier, "Sustaining Trust", in *Moral Prejudices*, p. 177.

being. In other words, when the group-identity of an individual, rather than her actual behaviour, begins in the mind of a member of group A to count, in and of itself, as evidence of the untrustworthiness of the individual in question, then distrust has set in. People who view others' group identities as robust evidence of their degree of trustworthiness will tend to display confirmation biases which will make the restoration of trust difficult. That is, they will tend to interpret all new evidence as supporting the judgment they will have made about a person on the basis of her status as a member of a distrusted group. 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 behaviour 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.²⁸ While this is, within certain bounds, a virtue of relationships of trust, it becomes a crippling aspect of relationships marked by significant *distrust*. 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, (or even above 0, if we assume, like some authors, that evolution has set us with a more positive default setting).²⁹

So on my account, unity-imperilling distrust occurs i) when group identities come to the fore, such that, for a typical member of group A, the fact that an individual belongs to group B gives rise, for a certain value of n to which A attaches great importance, to n being set below 0; ii) when distrust becomes robustly resistant to contrary evidence; and iii) when the lines along which these group enmities form overlap to at least some degree with territory, that is, when the groups in question are territorially concentrated. Situations in which iii) fails to hold might be less unity-imperilling than situations in which it does, which is not to say that distrust in these situations can't poison the body politic in other ways (witness race relations in the United States), or that group enmities cannot deteriorate to such a degree that some might be tempted to bring about the movement or decimation of populations required to *make* iii) obtain (as in the republics of former Yugoslavia).

²⁸On this issue, see Judith Baker, "Trust and Rationality", in *Pacific Philosophical Quarterly*, 68 (1987), p. 3; see also Jones, pp. 11-14.

²⁹See, for example, Baier, "Trust and Anti-Trust", p. 107. For a more sceptical note, see Russell Hardin, "Trusting Persons, Trusting Institutions", in R. Zeckhauser (ed.), *Strategy and Choice*, (Cambridge, Mass.: The MIT Press, 1991), p. 195.

Let us now return to the types of strategies mooted above.³⁰ 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 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 possess all the political levers required to attain the goal in question, or by placing obstacles in the way of B being able to affect A's pursuit of the goal. 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. 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. The Canadian government's oscillation over the course of the past 30 years or so between containment and replacement in the support of Quebec's pursuit of cultural survival illustrates both the attractions and the perils of these two variants on the goal-directed strategy.

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 interests 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

³⁰ A final prefatory remark: the measures I will be describing are aimed at restoring bonds of *trust*. While ruling out measures that obviously offend against basic norms of liberal-democratic political morality, I have by and large not sought systematically to assess the compatibility between the trust-promoting measures and requirements of justice. My rationale for this is that philosophers have for too long considered the requirements of justice without devoting any sustained attention to the effects which their view of justice would have on social unity. I felt it might be worth devoting some independent attention to the requirements of unity while holding matters of justice at arm's length (though never completely out of view). My view is that, on the one hand, a globally satisfactory political theory might end up being one which accepts some policies that are second- or third-best from the point of view of justice in order to realize other values, such as trust; but that on the other hand, governments which stray too far from justice in the pursuit of social trust will inevitably end up courting the distrust of its citizens. In other words, the cost of unifying a polity in unjust ways might end up being that citizens are united *in* distrust of the government.

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.

Identity-promotion however strikes me as a non-starter, at least for liberal democracies. The manufacturing of identity from above is fraught with peril. In cases in which it has been somewhat successful (say, in the creation of French, German or Italian identities), it has required that local identities and languages be subjugated and destroyed in ways which no modern liberal democracy could tolerate. In cases in which liberal democratic strictures have largely been respected, the failure to eradicate the pre-existing local identities has led, predictably, to backlash. The Canadian case represents a cautionary tale for all would-be national identity builders. The attempt by the government of Pierre Elliott Trudeau to forge a new national identity capable of dampening the hold of regionalisms by drafting a new federal constitution and by enacting ambitious national policies swiftly led to a strengthening of the national identities that were supposed to be supplanted, and, poignantly, to the perception on the part of both English and French-Canadians that the reshaping of federal institutions had constituted a capitulation to the unjustified demands of the other side.

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 “contagion principle”, it can identify areas and interests around which there is significant trust remaining between the members of the two groups, and hope that, with time, the solidification of these areas will loosen the grip which the cognitive patterns underlying distrust have. If the mechanism upon which the principle of contagion is based 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 reemphasizing 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 streetcorner 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.

It is difficult to see how any third-party can directly affect A's attitudinal trust toward B, that is, without affecting B's behaviour. (We will be examining trusted-directed strategies below). But 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. In the present context, the relevant option to trusting is secession, and so a state can tilt the balance in favour of trusting by placing significant obstacles on the road to A's secession, or, at the limit, by making it illegal. 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.³¹ 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.

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.³² 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 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

³¹ One of the tasks I have set for myself for this year is however to discuss the recent Canadian Supreme Court decision on the constitutionality of the unilateral secession of Quebec within the context of contemporary discussions of the problem of secession.

³² Melissa Williams, *Voice, Trust and Memory...*

achieve balance between sheer numerical representation and equitable representation of regions, it does not seem facially 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 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. While I do not agree with Russell Hardin that trust can entirely be thought of as "encapsulated interest",³³ as this would break down the distinction between trust and reliance, I 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 behaviour 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 behaviour 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.

Third, incentives and sanctions aimed at restoring trust must be designed in such a way as to *support* rather than to *supplant* trust-sustaining motives. In order for trust not to become mere reliance, that is the simple expectation on the part of an agent that others will behave in ways which support or at least do not go against her interests, regardless of the *motive* behind this support or forbearance, incentive-rigging must work with motives that are already present rather than seeking

³³The following quite captures Hardin's view well: "Most of us are somewhat like bank tellers: we are sucred 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. 203-204.

to replace them. This is one of the reasons behind the principle of realism: one cannot hope to repair trust if it is not already to some degree present. Trust must build on trust, at least in situations where it is accompanied by areas of significant *distrust*. Incentive-rigging thus must assume that we are already possessed of mixed motives toward one another, and it must make it more attractive for us to act on our trust-enhancing motives than to act on our more trust-damaging ones, or to put the point another way, it must make it more costly for us to act on on ther latter motives than on the former. It must not supply us with motives that we do not already have. Much could be written about what this would require sa far as institutional design is concerned, but I would make simply make the two following remarks. First, there is a case to be made for the use of sanctions rather than rewards. The idea behind incentive-rigging is that we already possess trust-supporting motives, but are sometimes too weak-willed not to allow more trust-damaging motives to hold sway. The task is therefore to disable these motives, or at least to make them more costly to act on. Constraints and sanctions seem more appropriate to this task; there is always the risk that providing incentives for trustworthy behaviour will have the perverse effect of transforming trust into mere reliance.

Second, sanctions should be as fine-grained as possible. That is, they should be geared toward restoring *trust* and not simply toward eliciting certain kinds of behaviour. They should therefore ideally make it more costly for an agent *not* to act on certain motives, rather than simply making it more costly not to *act* in certain ways. The kinds of reputational sanctions described by republican theorists such as Philip Pettit seem best suited to this task. The main idea behind Pettit's argument³⁴ 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 repuation.

With these conditions in mind, I think that the incentive-rigging strategy is a plausible one. Some might argue that *any* use of incentives and disincentives is incompatible with trust. I take up this objection in the following section.

As a result of this (very brisk and probably incomplete) survey of the tools that are at the disposal of liberal-democratic states seeking to repair damaged trust between groups, we can conclude that the state is not at first glance powerless in this area. The types of strategies we have surveyed are moreover not mutually exclusive. Keeping in mind that context-sensitivity is essential when deciding whether to attempt to repair damaged trust through containment, replacement, incentive-rigging, trust-spreading, or what have you, there seem no *apriori* reason to practice one to the exclusion of the others.

³⁴The argument is developed principally in Pettit, "The Cunning of Trust".

Now, some of these strategies dovetail nicely with some of the measures which other theorists have argued to be requirements of *justice*. Thus, for example, containment, which requires that groups be immunized from the impact which the members of other groups might have on their ability to secure an important end, converges with conclusions of Will Kymlicka's according to which justice requires that group-differentiated rights, in essence more or less extensive areas of group autonomy, be attributed to minority national groups as a condition of their members' being able to access the cultural "context of choice" which they require in order to meaningfully exercise individual choice.³⁵

However, for reasons already mooted in this paper, I am less inclined than he seems to be to accept group identities as independent variables. Such identities are in my view often at least partly constructed as a result of strategic considerations emanating from interactions between the members of initially much more fluid groups.

For the same reason, I would argue that group-differences and the accommodations and compromises which the attempt to deal with group-distrust occasions ought not to be constitutionalized. The function of constitutions is to place certain issues off the political agenda, or at least, to make it very costly and burdensome to reopen them. But as I have argued, unity-imperilling breaches of trust are often unpredictable. One cannot simply guess by looking at a snapshot of a society when and where they will occur. There is a dynamism and fluidity to group-life in liberal democracies with which constitutionalism is ill-equipped to deal. The kinds of accommodations, devolutions and are perhaps best left up to the (suitably constrained³⁶) democratic political process.

Is Social Trust Compatible with the Institutionalization of Trust?

One final question remains to be examined. The foregoing section has examined various ways in which governments can design institutions and laws so as to create an incentive structure which are best suited to reestablishing 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 rigging the incentive

³⁵ Will Kymlicka, *Liberalism, Community and Culture*, and *Multicultural Citizenship*.

³⁶I have in mind here the types of constraints upon democratic processes which, for many theorists, are required in order to make democracy truly "deliberative" rather than a mere exercise in interest-aggregation. For a small but important sample of contemporary writing on the subject, see Amy Gutmann and Dennis Thompson, *Democracy and Disagreement*, (Cambridge, Mass.: Harvard University Press, 1996); and the essays collected in J. Elster (ed.) *Deliberative Democracy*, (Cambridge: Cambridge University Press, 1998).

structure in a sufficiently clever way, we have simply succeeded in making him *reliable* rather than *trustworthy*?³⁷

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”.³⁸ 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 extreme. 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” mechanisms³⁹: 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. 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

³⁷For the difference between reliability and trustworthiness, see Richard Holton, “Deciding to Trust, Coming to Believe”, in *Australasian Journal of Philosophy*, 72 (1994).

³⁸D.O. Thomas, “The Duty to Trust”, in *Proceedings of the Aristotelian Society*, 79 (1978-1979), p. 91.

³⁹Jon Elster, *Ulysses and the Sirens*, rev. ed., (Cambridge: Cambridge University Press, 1984), ch. II.

promote trust by “rigging the 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.⁴⁰

Conclusion

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.⁴¹

Daniel Weinstock
Princeton University/Université de Montréal

⁴⁰The condition that we be able to see such arrangements as stemming from our political agency rather than as imposed upon us from above provides another reason for not constitutionalizing such arrangements, and for entrusting them to suitably constrained democratic politics.

⁴¹This paper is a distant relative of a presentation made in June 1997 to a Canadian Philosophical Association panel on “Civic Diversity and Social Unity”. My thanks to the other participants and to the audience members for lively discussion. Thanks to Helen Nissenbaum and Richard Moran who helped me immeasurably with the present version. Thanks, finally, to my research assistant, Sayumi Takahashi.

