Applying Social Identity Theory to the Study of International Politics:
A Caution and an Agenda

By Jacques E. C. Hymans
Assistant Professor of Government
Smith College
(jhymans@smith.edu)

Paper originally prepared for presentation at the International Studies Association
ABSTRACT

Scholars now routinely use "identity" as an explanatory variable in the study of international politics. While there are myriad ways in which social scientists understand the term, one important strand of research is derived from "social identity theory" (SIT), a well-established approach to intergroup relations from social psychology. SIT has enormous potential as the basis for an alternative lens or paradigm for the IR field, but there are significant issues of external validity and cross-disciplinary translation that we must tackle first before presenting it as a ready-to-wear alternative to traditional IR paradigms. The purposes of this paper are, first, to urge caution in the application of SIT to IR, and second, to suggest a joint agenda for those psychologists and political scientists who see in SIT great potential for explaining international behavior.
1. Introduction

Scholars now routinely use "identity" as an explanatory variable in the study of international politics. While there are myriad ways in which social scientists understand the term, one important strand of research is derived from "social identity theory" (SIT), a well-established approach to intergroup relations from social psychology. Prominent IR scholars such as Jonathan Mercer have claimed that SIT offers the potential to reestablish realism on psychological as opposed to materialist grounds. Mercer’s work has had the great result of putting SIT in the IR scholar’s toolbox. But although SIT has enormous potential as the basis for an alternative lens or paradigm for the IR field, there are significant issues of external validity and cross-disciplinary translation that we must tackle first before presenting it as a ready-to-wear alternative to traditional IR paradigms.

The purposes of this paper are, first, to urge caution in the application of SIT to IR, and second, to suggest a joint agenda for those psychologists and political scientists who see in SIT great potential for explaining international behavior.

The paper is organized as follows. In the following section, I present a very brief summary of SIT. In the third section, I argue that to equate SIT with realism is a far too simple reading, and I list several key variables without which SIT’s predictions are indeterminate. In the fourth section, I suggest the need to marry two auxiliary theories to

---

1 Thanks to Marilyn Brewer, Michael Griesdorf, and participants at the workshop on “Identity Measurement” at the Weatherhead Center for International Affairs, Harvard University, and at the workshop on “Political Psychology and Behavior” at the Center for Basic Research in the Social Sciences, Harvard University, for their comments on an earlier version of this paper.

2 The boom in political science research using the concept of identity is charted in Rawi Abdelal, Yoshiko M. Herrera, Alastair Iain Johnston, and Terry Martin, “Treating Identity as a Variable: Measuring the Content, Intensity, and Contestation of Identity,” paper prepared for presentation at the American Political Science Association meeting, San Francisco, August 30-September 2, 2001, esp. pp. 4-6.

the application of SIT to IR—issue-framing on the front end, and emotional decision making on the back end. The fifth section summarizes the main points of the paper.

2. The Basics of SIT

SIT research is fundamentally concerned with understanding the variable nature of the “self.” In SIT and its close cousin, self-categorization theory, a person has not just one, “personal” self, but rather several selves that correspond to ever-widening circles of group membership. Different social contexts may trigger an individual to think, feel and act on the basis of (for example) his personal, family, regional, occupational, partisan, or national “level of self.” A “social identity,” according to this perspective, is an individual-based understanding of what defines the group. This is something of a counterintuitive notion for political science, which is more familiar with the sociological perspective that the group defines the group (through institutionalized discourse) and even that the group defines the individual (through an institutionalized assignment of roles). The distinction between SIT and other identity approaches, a distinction that political scientists often ignore, can be visualized in terms of the below table (see next page):

---

8 One relevant political science analogy, though, would be to pocketbook versus sociotropic voting.
Table 1: Different Takes on “Identity”

<table>
<thead>
<tr>
<th>Psychological</th>
<th>Personal identity</th>
<th>Group identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychodynamic approaches</td>
<td></td>
<td>Social identity theory</td>
</tr>
<tr>
<td>Sociological</td>
<td>Role theory</td>
<td>Post-modern discourse theory</td>
</tr>
</tbody>
</table>

With its distinctive object of explanation, SIT is not necessarily incompatible with other approaches. In particular, given that few individual citizens develop entirely unique notions of their group, there should be a role for sociological theories to explain broad commonalities in individuals’ social identities. But the integration of SIT with sociological theories is not easy.⁹

According to SIT, at each level of self we have two deep-seated motivations: to define ourselves and to enjoy positive self-esteem. The pursuit of self-definition at every level of self pushes the individual to compare and contrast his “ingroup” with “outgroups.” Meanwhile, the pursuit of self-esteem pushes the individual to make sure that the comparison is flattering. As a result, SIT tells us, inherent in any social identity is a basic tendency for “ingroup favoritism” at the expense of outgroups in terms of attribution of motives, evaluation of qualities, and distribution of goods.¹⁰

The key tests of SIT have consisted of experiments with “minimal groups”—experimental subjects assigned on a random or trivial basis to one or another camp, with

---


no face to face interaction with other ingroup members. Henri Tajfel, the originator of SIT, made the remarkable finding that the mere assignment of individuals to such minimal groups is enough to spark strong tendencies for them to discriminate in favor of their ingroup at the expense of the outgroup. Though they have nothing personally to gain in the experiments, subjects again and again perceive fellow ingroup members (as opposed to outgroup members) as having more correct motives for their actions, as doing better or more successful work, and as meriting a proportionately larger share of goods. This favoritism occurs even when the structure of the situation does not dictate a competitive relationship between ingroup and outgroup.

The implications of SIT for the study of international politics are potentially great. To summarize these in one sentence, SIT suggests that psychological factors, perhaps even more than material ones, may be at the root of the failure of nations to bring about international equity, justice, and peace. However, I would stress that SIT at present does not offer us a ready-to-wear theory of international politics. There is a large chasm to be bridged between the laboratory conditions of most SIT experiments and the real world. At the same time, laboratory experiments in the SIT tradition have actually produced a more complicated picture of reality than is acknowledged in many IR appropriations of the theory. In the next section of this paper, I point out some of the variables that IR scholars should investigate as they apply SIT to their objects of explanation.

---

11 The progress of these experiments over time is chronicled in the annual European Review of Social Psychology, 12 vol. (London: Wiley, 1990-present).
3. Towards a more nuanced application of SIT to IR

SIT clearly holds great promise for the study of IR, but only if applied with care. The rest of this paper will highlight some of the difficulties in making the leap from experimental psychology to empirical political science. Specifically, in this section I argue that IR scholars have generally overlooked the crucial distinction made in SIT between positive ingroup favoritism and negative outgroup discrimination. This oversight has led them to the incorrect conclusion that simple issues of cognitive categorization are sufficient to produce a tendency for intergroup conflict. I then suggest some key intervening variables that complicate even further the equation between basic cognitive self-categorization and intergroup relations: (1) the perceived relevance of different outgroups for ingroup identity; (2) the perceived legitimacy and fixedness of intergroup power/status hierarchies; and (3) the degree of ingroup ontological uncertainty.

Ingroup love vs. outgroup hate

Scholars who wish to apply SIT to international politics have generally overlooked or glossed over the crucial distinction its advocates make between ingroup love and outgroup hate.13 SIT experiments routinely find a preference for one's ingroup over the outgroup in terms of evaluation, liking, or allocation of resources.14 What they do not find is that this preference necessarily or even typically translates into negative

---

feelings toward the outgroup.\textsuperscript{15} It is also important to note that even strong feelings of dislike for an outgroup need not produce aggression against it.\textsuperscript{16} Yet this is what many IR scholars have understood the SIT literature to say. This is perhaps due to contamination from the Freudian-influenced approach to identity, a much more common approach in IR, in which the ingroup love-outgroup hate pair is something of an article of faith.\textsuperscript{17}

Ignoring the ingroup love-outgroup hate distinction can lead to premature “falsification” of applications of SIT to IR. For instance, Margaret Hermann and Charles Kegley argue that from an SIT perspective, democracies' tendency not to fight each other results from their categorization as members of the same ingroup. The other side of the coin for them is that democracies should have an exaggerated tendency to fight with non-democracies, i.e., members of the outgroup.\textsuperscript{18} Testing these twin hypotheses, Bruce Russett et al. find that democracies are unlikely to fight each other, and that they are no more likely than non-democracies to fight with other non-democracies. They thus


\textsuperscript{17} Two examples of such an approach are David Campbell, \textit{Writing Security: United States Foreign Policy and the Politics of Identity}, Revised Edition (Minneapolis, MN: University of Minnesota Press, 1998), and Vamik D. Volkan, \textit{The Need to Have Enemies and Allies: From Clinical Practice to International Relationships} (Northvale, NJ: Jason Aronson, 1988).

conclude that the SIT explanation for the democratic peace is disconfirmed. But in fact Hermann and Kegley’s prediction, however reasonable, does not follow from SIT, which distinguishes between ingroup love and outgroup hate. Russett et al.'s finding that democracies are unlikely to fight each other, while being no more likely than non-democracies to fight with non-democracies, is actually entirely compatible with SIT.

In this case the distinction between ingroup love and outgroup hate "saves" the SIT approach to an important issue in IR; but in general, the distinction serves greatly to reduce the utility of simple cognitive categorization as a one-stop shop for understanding international conflict. If simple cognitive categorization is not enough to produce dislike and negative outgroup discrimination, let alone aggression, then in order to understand international Hobbesian anarchy, other factors—whether ideational or material in nature—need to be brought back in to the picture. Below I list three such factors that would seem crucial to include in any attempted application of SIT to IR.

*Various others vs. the “Other”*

Scholars applying SIT to international politics generally understand it as depicting a world divided by the black-white dichotomy of "ingroup" and "outgroup": whoever is not "in" is "out." For instance, this inside-outside dichotomy is crucial to Jonathan Mercer's claim that SIT provides psychological foundations for realism. According to Mercer, because of the ingroup-outgroup distinction, nations commit the "fundamental attribution error" in interpreting other nations' acts; that is, they uniformly view them—

---


allies and adversaries alike—in the worst possible light. So we should perceive acts of allies or adversaries that are contrary to our interests as the result of their basic disposition, and acts that are in line with our interests as the result of the constraints imposed by the objective situation. In assuming that nations make a black-white dichotomy between "us" and "everybody else," however, Mercer is making a leap beyond what SIT experiments really tell us. While it is true that the vast majority of SIT experiments have been set up to explore interactions between “us-them” dyads, this is a limitation of the SIT literature rather than a clearly enunciated, experimentally verified statement about the way that groups tend to perceive the world.  

On the theoretical plane, in one of the foundational articles on SIT, Tajfel and Turner argued that “ingroups do not compare themselves with every cognitively available outgroup: the outgroup must be perceived as a relevant comparison group.” Only relevant comparison outgroups, Tajfel and Turner suggested, would bear the brunt of ingroup favoritism tendencies. So SIT clearly understands that not all others can be called the “Other.”

Even for the set of relevant comparison outgroups, there could be differential discriminatory tendencies. In particular, one could surmise that the more outgroups an ingroup uses to define itself, the less tendency for it to imagine a black-white dichotomy with any one of them. There is a crying need for more work on this hypothesis, and more

---

21 This limitation sometimes leads psychologists to force the analysis of intergroup relations into such dichotomies even when multiple groups are noted, e.g. in Samuel L Gaertner et al., “The Common Ingroup Identity Model: Recategorization and the Reduction of Intergroup Bias,” in *European Review of Social Psychology*, vol. 4 (1993), p. 18.


23 So far SIT has not elucidated how outgroups become relevant for identity comparison, but it appears that the norm is to compare the ingroup to outgroups of similar status, except in cases of perceived illegitimate status hierarchies. See Rupert Brown and Gabi Haeger, “Compared to What? Comparison Choice in an International Context,” *European Journal of Social Psychology*, Vol. 29, No. 1 (January-February 1999), pp. 31-42.
generally on the predictions of SIT in multipolar contexts. But there are plenty of examples in international history that point in this direction. For instance, Henri Pirenne, an historian whose vision of Belgian identity had a powerful impact on many of his compatriots, saw Belgium as having adopted the best and avoided the worst aspects of both French and German cultures. The conclusion for Pirenne was that while Belgium was "white," neither France nor Germany was "black"; moreover, the two neighbors were certainly not equivalent or interchangeable. The bottom line here is that IR scholars cannot afford to assume what SIT experiments too often do, that the dyad is the universal condition of intergroup politics.

**Settled vs. unsettled status/power hierarchies**

IR scholars have tended to pose power- and identity-based explanations as alternatives. This is encouraged by the nature of much SIT research, which is often silent on the issue of the relative status or power of the groups under study. But this silence should not be taken to imply that SIT-type processes stand independent of a group’s relative rank in the pecking order. Rather, in the design of most SIT experiments there is an implicit assumption of rough status and power parity. Moreover, the logic of SIT theory suggests that its findings of ingroup bias may in fact be dependent on this assumption.

---

27 This point has been made in the social psychology literature itself. See S. H. Ng, *The Social Psychology of Power* (London: Academic Press, 1980).
Donald Horowitz, in his SIT-influenced work on ethnic conflict, distinguishes between what he calls "unranked groups," or groups that have not (yet) established a clear pecking order between them, and “ranked groups,” those that have. Horowitz argues that unranked groups have a greater tendency to demonize the other and have more conflicts with each other than do ranked groups. For Horowitz, the greater intensity of relations between unranked groups is due to the felt need to establish a hierarchy between them, a need that ranked groups have already met. The basic setup of most SIT experiments implicitly produces a situation of “unranked groups”—precisely the context in which Horowitz would expect demonization and conflict to be more prevalent. The distinction Horowitz makes between ranked and unranked groups has clear relevance for international relations as well, although the international law fiction of interstate equality might make unranked relationships the norm. Indeed, the extent to which that norm produces unsettled identity hierarchies in IR could make for a fascinating research project.

A small body of experimental work on the question of relative rank (in terms of status or power) in the SIT tradition suggests an even more nuanced hypothesis. Itesh Sachdev and Richard Y. Bourhis have found that members of groups with relatively high rank display the typical tendency for ingroup favoritism vis-à-vis groups of lower or equal rank, but that members of groups with relatively low rank actually display a

---

29 There is a small SIT-influenced literature on status issues in inter-group relations, and it generally tends to reinforce this point. See Hagendoorn, "Intergroup Biases in Multiple Group Systems," esp. p. 224; John H. Duckitt, *The Social Psychology of Prejudice* (New York: Praeger, 1992). See also the next paragraph.
30 Status and power are clearly separate concepts, but in practice they are often very hard to distinguish, and at this early stage of development in the literature it is difficult to make such fine-grained distinctions. However Sachdev and Bourhis do make an attempt to distinguish them experimentally (see following footnote).
tendency for *outgroup* favoritism vis-à-vis groups of higher rank.\(^{31}\) In other words, Sachdev and Bourhis suggest that the reason for less intergroup conflict in the “ranked group” situation may be not so much the result of a lower tendency by those of higher rank to discriminate, as of the *acceptance* by at least some of those of lower rank of their “inferiority” and thus the legitimacy or at least inevitability of discrimination. So, whereas a straightforward SIT hypothesis would argue that there should be more discrimination and thus more conflict in relations between peers, Sachdev and Bourhis’ experimental evidence implies a more complex story. It may be necessary to distinguish between different *types* of discriminatory tendencies—those in “unranked group” situations, which are born of self-defense, and those in “ranked group” situations, which are born of contempt. Clearly more experimental work is needed here, but the bottom line for IR scholars is that they would be mistaken to portray identity and power explanations as simple alternatives.

*Level of ontological uncertainty*

A typical political science criticism of SIT is that as most of its work has focused on “minimal” groups—groups whose members have no shared history or significant common traits, except for the basic fact of categorization—it therefore cannot understand the intensity of conflicts between real groups with long histories, such as the Israeli-Arab

\(^{31}\) Status and power were studied in separate experiments. To measure the role of “power,” subjects were placed in groups with different degrees of control over the distribution of goods. To measure the role of “status,” subjects were assigned ostensibly different performances on a “creativity test.” This operationalization is not entirely convincing, but it is a start. Itesh Sachdev and Richard Y. Bourhis, “Social categorization and power differentials in group relations,” *European Journal of Social Psychology*, Vol. 15, (1985), pp. 415-434; Sachdev and Bourhis, “Status differentials and intergroup behavior,” *European Journal of Social Psychology*, Vol. 17, (1987), pp. 277-293.
conflict. It is certainly important to consider the external validity of experimental evidence, something that social psychologists have only begun to do. But the logic of SIT suggests that in terms of its external validity, we could entertain exactly the opposite hypothesis from the one typically made by political science criticisms of SIT. One interpretation of the motivational mechanism for SIT-type processes of evaluation and discrimination is that the individual has a desire to decrease group ontological uncertainty—in other words, the individual wants to reinforce his sense of the reality of the group. There is a parallel here with the previous section of the paper, which focused on a tendency towards intergroup discrimination stemming from uncertainty about relative rank. If ontological uncertainty tends to produce intergroup discrimination, then in fact SIT may find its strongest confirmation among minimal groups. This is because such groups are based on next to nothing, and thus they desperately need to establish firm boundaries with the outgroup.

The more minimal the group, the more prone it may be to exhibit strong ingroup favoritism. One reasonable IR hypothesis that emerges from this is that the processes of evaluation and discrimination identified by SIT may be most evident in the international behavior of recently "imagined" nations, such as Tanzania or Guyana. They may also be powerful in the case of seemingly culturally "thin" supranational identity categories such as the "civilized world" or the "club of democracies," whose apparently large effects on

---

policy are otherwise difficult to explain. By contrast, for more settled identity categories, such as “Portugal,” the path between basic cognitive categorization and external behavior would likely be much more indirect. In such cases, history—or more precisely, historical memory—could serve just as easily to muffle as to amplify the basic cognitive tendencies toward discrimination.

Since the mechanism producing discrimination here is not minimality per se, but rather ontological uncertainty, it is important to note that things other than minimality could produce such uncertainty. These points relate to the large psychological literature on the “contact hypothesis,” to which SIT’s contribution has been quite substantial. SIT researchers have found that intergroup contact can be unsettling and lead to a stronger motive for boundary definition and hence discrimination under many circumstances. To avoid this, they recommend that the context be manipulated to decrease the salience of the group self. This can be done by raising the salience of alternative selves, such as the personal self or a wider self that includes the “Other” within it. If this cannot be done, then SIT’s best advice boils down to the proverb that good fences make good neighbors.

Although the term “boundaries” in social psychology denotes a cognitive reality, not a territorial one, on the IR plane territory is central to most definitions of nation-statehood. So it would not be surprising if the penetration or contestation of physical

---


boundaries led to heightened levels of intergroup discrimination and bias.\textsuperscript{37} This suggests a possible basis in SIT for the intense passions generated by Alsace-Lorraine, Kashmir, Taiwan, the Falklands/Malvinas Islands, and many other international hot spots.

\section*{4. SIT, issue-framing, and emotional decision making}

To summarize the argument of this paper so far: self-categorization is an important factor in SIT’s explanation of intergroup discrimination, but by itself it is not enough to produce IR realist attitudes or behavior. SIT’s predictions of attitudes towards outgroups start looking much more like realism if three other conditions are met: (1) if the outgroup is not merely an outsider but the “Other,” that is, the basis of ingroup identity construction; (2) if the hierarchy between the ingroup and the outgroup is perceived as unsettled; and (3) if the behavior or mere presence of the outgroup causes ontological uncertainty in the ingroup. In this section, I argue that even if all these conditions are in place, the SIT perspective still implies a vision that is quite distinct from realism. For the application of SIT to IR cannot stand alone, but rather must be married to two auxiliary theories, one—issue-framing—that that is prior to any application of SIT and one—emotional decision making—that follows in its wake. I consider each of these in turn.

\textit{Need for an auxiliary theory of issue-framing}

What makes SIT a social-\textit{psychological} as opposed to a sociological theory of identity is its focus on the central role of the individual in defining his own understanding

\textsuperscript{37} Note the similarity of the logic here to the major theory of ethnic groups developed by the anthropologist Fredrik Barth. Fredrik Barth, ed., \textit{Ethnic Groups and Boundaries: The Social Organization of Culture}
of his group “level of self.” Indeed, in SIT experiments the term "intergroup relations" is something of a misnomer, if that term is taken to mean the interaction of groups as roughly unitary actors. Rather, the experiments typically investigate individuals' actions or perceptions made on behalf of the collectivity. This is a point too often ignored in IR, with its facile anthropomorphism of states as “actors.” It is very important to the success of SIT experiments to cue the group "level of self" on the part of the individual. Indeed, many SIT experiments focus on that very point—the different perceptions and behaviors exhibited by the same individual depending on whether the "personal" or "group" self has been activated. In experimental conditions, it is relatively easy to cue and then to confirm that the individual is acting on the basis of a particular group "level of self." But in the real world, given the complexity of most international issues, it is generally hard to assert that the objective nature of the issue at hand will necessarily spark one or another “level of self.”

Moreover, in international politics the choice is not only between the “personal” or “group” self, it is also between various group selves—e.g., is the decision maker thinking and acting on behalf of his family, his party, his nation, or on behalf of a supranational category such as the “club of democracies”? As Emanuel Adler has pointed out, the possibility of multiple “levels of self” is a crucial SIT insight that

---


38 Indeed, social psychologists define intergroup behavior as "whenever individuals belonging to one group interact, collectively or individually, with another group or its members in terms of their group identification." Muzafer Sherif cited in Marilynn Brewer, "Category-based vs. Person-based Perception in Intergroup Contexts," European Review of Social Psychology, Vol. 9 (1998), p. 77.


Mercer’s initial article ignored. Thus the application of SIT to IR must be accompanied (indeed preceded) by an application of one or another theory of issue-framing. Perhaps the hardest part of this for many IR scholars to swallow is the high level of attention to the precise decision-making context that any issue-framing theory requires. But some IR applications of SIT have paid serious attention to this issue, notably the work by Bruce Cronin. Cronin does an excellent job of identifying the conditions for how the trans-state “Italian” identity came to trump local political identities among Italian elites in the mid-19th century.

Need for an auxiliary theory of emotional decision making

Emanuel Adler cites SIT to support his argument that community is “not as a matter of feelings, emotions, and affection, but as a cognitive process through which common identities are created.” Though this ignores the centrality of the self-esteem motivation in SIT, in general Adler is correct to assert that SIT has traditionally framed identity more as a matter of the mind than of the heart. But many current SIT researchers view their cognitive approach to identity as a limitation to be rectified. Recent work in social psychology suggests that different cognitive appraisals of relationships can produce different emotional responses in the appraiser. Marilynn Brewer and Michele Alexander have spelled out five types of relationship patterns depending on the perceived intergroup complementarity of goals and the perceived status and power hierarchies.

---

42 He does a less complete job of explaining why other possibly salient group identities, such as family, were superseded. Bruce Cronin, “From Balance to Community,” op. cit.
43 Adler, “Imagined (Security) Communities,” op. cit., p. 263.
Drawing on work in IR by Richard Herrmann and Michael Fischerkeller, Brewer and Alexander suggest that different “outgroup images” may be correlated with different intergroup emotions: the “ally” with admiration and trust, the “enemy” with anger, the “dependent” with disgust and contempt, the “barbarian” with fear and intimidation, and the “imperialist” with jealousy and resentment. Efforts to link SIT to emotional responses and eventually to emotional decision making are still in their infancy, but as Rupert Brown puts it, “If SIT is to help in the understanding and resolution of these social conflicts then it must be elaborated to take account of and predict such displays of negative emotion and behavior.”

5. Conclusion

One of the primary purposes of this paper has been to caution IR scholars against testing hypotheses based on an oversimplified version of SIT. SIT is not a ready-to-wear theory of international politics. There is a great danger that going too far, too fast with SIT will lead to the premature "falsification" of SIT on the international plane. But this caution should not dissuade IR scholars from working to develop a robust paradigm out of the core SIT findings. In making the leap from lab to field, IR scholars should be attentive to complement a solid understanding of SIT’s findings with, at the very least, a focus on nations' perceived relevant comparison “Others,” their perceived relative power or status, their ontological uncertainty, and theories of issue-framing and emotional decision making.

The paper also suggests an agenda for SIT itself. Many of the issues of external validity raised in this paper could be clarified, if not resolved, by more experiments. Social psychologists, perhaps in tandem with IR scholars, could do much more work on multipolar situations; they could do much more work on relationships with a clear pre-existing hierarchy; and they could do much more work to establish a bridge from minimal to actual social identifications. The task will be long and arduous but certainly worth undertaking.