CHAPTER 12

ETHNICITY AND ETHNIC CONFLICT

ASHUTOSH VARSHNEY

"Until recently," wrote Donald Horowitz in 1985, "the field of ethnic conflict has been a backwater of the social sciences." This statement is to be taken seriously. Horowitz's Ethnic Groups in Conflict was a seminal text. For the first time in scholarly history, a book on ethnic conflict covered a whole variety of topics, ranging from concepts and definitions to those spheres of institutional politics (party politics, military politics, affirmative action) in which the power of ethnicity had become obvious and could no longer be ignored. Some important social science arguments had emerged earlier, especially on the relationship between ethnicity and nation building, ethnicity and modernity, ethnicity and consociational democracy (Lijphart 1969; 1977), and migration and ethnic conflict (Weiner 1978). But each of these works covered a specific problem at hand. Ethnic Groups in Conflict covered a wide array of topics under the umbrella of ethnicity, becoming thereby the founding text of the field.4

Over two decades have passed since then. There has been such an explosion of research on ethnicity and ethnic conflict that the field can no longer be called a "backwater of the social sciences." Especially since the end of the Cold War, the rise of ethnicity has coincided with the weakening of the customary left–right ideological

---

1 As stated in the introductory chapter of Horowitz (1985, 13).
2 For example, Brass (1974); Connor (1972); Geertz (1963); Shils (1957); Smith (1979).
3 For example, Rudolph and Rudolph (1968); and Deutsch (1966).
4 For whatever it is worth, it may be noted that according to "Google scholar", as of February 1, 2007, Horowitz's Ethnic Groups in Conflict had been cited 807 times. Some works on nationalism have been cited more, but none more on ethnicity or ethnic conflict.
axis in politics the world over, both in the developed and developing world. As a research field, too, ethnicity has become a growth industry, straddling a variety of disciplines, topics, and methods, and attracting a large number of scholars.

But have we made progress? And if so, in what ways? In a widely read evaluation of the field, published in 1998, Brubaker and Laitin were negative about the progress made:

Notwithstanding the increasing scholarly interest in ethnic and nationalist violence, there is no clearly demarcated field of social scientific inquiry addressing the subject, no well-defined body of literature, no agreed-upon set of key questions or problems. It is not simply that there is no agreement on how things are to be explained; more fundamentally, there is no agreement on what is to be explained, or whether there is a unitary phenomenon (or a coherently related body of phenomena) to be explained. Rather than confronting competing theories or explanations, we confront alternative ways of posing questions, alternative approaches to or “takes” on ethnic and nationalist violence, alternative ways of conceptualizing the phenomenon, and situating it in the context of wider theoretical debates.5

Two things should be noted about this evaluation. First, it relates only to ethnic and nationalist violence, not to the whole field of ethnicity. The latter term now covers topics as varying as ethnic identity formation, ethnic movements and protests, ethnic voting and ethnic parties, ethnic heterogeneity and allocation of public goods, ethnic diversity and economic growth rates, and ethnic riots, pogroms, and civil wars. No essay can cover all of these topics adequately. I will confine myself to only two topics: ethnic identity and ethnic conflict. I will distinguish them especially from national identity and nationalism on the one hand, and civil wars on the other. These latter topics are covered elsewhere in this volume. I will use arguments about nationalism and civil wars only to the extent that they clarify my analytic overview of the literature on ethnic identity and conflict.

Second, what Brubaker and Laitin find troubling may, in part, be viewed as a reflection of the field’s age. As King, Keohane, and Verba (1994) argued, the younger fields are like a double-edged sword. Typically, they do not have a body of theory that most scholars agree with, but the returns to entry may be great. In the established fields, strong theory exists and progress is typically marginal. In younger fields, big theoretical strides can be made.

Has the field made great strides since Brubaker and Laitin wrote their evaluation? I argue below that progress has been substantial. I also argue that wide acceptance of two concepts—mechanisms and variations—has driven the evolution of research, especially in the last ten years.

In earlier times, scholars often used to leave theory building to a link, or affinity, between structural conditions and the rise of ethnic conflict or nationalism. Gellner (1983) is the most illustrative, and well-known, example of this tendency. Gellner essentially theorized that the rise of the industrial age required nationalism, as linguistic standardization became necessary for communication between citizens and the rural masses left their village particularities behind, moving to unknown

---

5 This essay is reproduced as Brubaker and Laitin (2004), where the section I have cited is on p. 92.
cities in search of industrial employment. Given the social science norms of the 1990s, a critique was easy to launch. The fact that industrialization required nationalism did not mean that it would happen. Why should need create its own fulfillment? At the very least, we need an account of the organizations, movements, or leaders that would undertake the task of converting objective needs into actual outcomes.

The idea of variance, similarly, has made advances possible. Theorizing about ethnic violence used to be based on establishing commonalities across the many cases of violence (or sometimes based on an in-depth case study or two). By the mid-1990s, following the popularity of King, Keohane, and Verba (1994), this came to be called "selection bias," and deemed inadmissible for theory construction.

Selection on the dependent variable, it was later recognized, was not entirely without its uses. It could, for example, knock down an existing theory, if the generalizations based on similar cases led to an argument opposed to the existing theoretical orthodoxy. But in and of its own, it was not enough to generate a new valid theory. Outcome variation was a better principle to follow for theory construction. Most research in the field has followed this principle over the last decade.

Despite these advances, Brubaker and Laitin are right in one sense. Cumulation has been quite slow. Very few theories have been fully knocked over. A more rapid "creative destruction" is likely to take place in the future, especially because testing has become a norm in the field.

Existing arguments about ethnic identity and/or ethnic conflict can be divided up into five traditions of enquiry: essentialism, instrumentalism, constructivism, institutionalism, and realism. There are theories within each tradition. I will concentrate on the first four traditions in this chapter, concentrating on the core idea of each tradition and how it has evolved over time.

I will leave out realism. Brought in from the field of international relations, realism is driven by the concept of security dilemma. Realists argue that when an existing state collapses, relations between ethnic groups begin to resemble those between states in the international system, the difference between defensive and offensive ethnic mobilization disappears, and neighbors kill neighbors to ensure that they are not possibly killed in the future. Such situations are more applicable to civil wars, excluded from the purview of this essay, and discussed elsewhere in this volume.

Section 1 is conceptual. Given the number of terms moving imprecisely about in the field, clarity about what we mean by the various terms is necessary for constructing a clear analytic domain. Section 2 surveys explanations provided in the four traditions of enquiry, analyzes the inadequacies or merits of arguments within each tradition, and reviews the evolution of arguments. Section 3 presents conclusions.

---

6 For example, the arguments about conflict in Horowitz (1985) were based on the commonalities principle. In his more recent work, Horowitz has taken note of variance and dealt with it. See Horowitz (2001, ch. 12).

7 See Varshney (2006) and Laitin (2006). It should also be noted that the search for commonalities is quite valid if one is identifying the characteristics of the phenomenon or problem at hand.
1 Concepts

1.1 What is Ethnicity?

Following Horowitz (1985), ethnicity as a term designates a sense of collective belonging, which could be based on common descent, language, history, culture, race, or religion (or some combination of these). Some would like to separate religion from this list, letting ethnicity incorporate the other attributes. From the viewpoint of political identities and group solidarity, this separation is a semantic quibble. It becomes critical, however, when ethnicity and religion clash (East and West Pakistan before 1971, Kashmiri Hindus and Muslims, Irish Protestants and Catholics, black and white American Christians).\(^8\)

How is a nation different from an ethnic group? An ethnic group may do without a state of its own; a nation implies bringing ethnicity and statehood together. Nationalism therefore becomes a principle that “the political and the national unit should be congruent” (Gellner 1983, 1). This congruence may be satisfied in a federal arrangement, or may head for nothing short of sovereignty.

In official as opposed to academic terminology, another term “nationality” is also used, particularly in the former Soviet bloc. In this three-tiered classification, a nation is a group with a political and territorial home; a nationality is a large ethnic group without such a home (but with cultural rights pertaining to language and sometimes religion); and an ethnic group is a smaller collectivity, different from a nationality but not large enough to be called a nationality. In the post-1945 Yugoslavia, Croats, Macedonians, Serbs, Slovenes, and Montenegrins were called nations; Albanians, Hungarians, Bulgarians were nationalities; and Austrians, Greeks, Jews, Germans, and Poles were “other nationalities and ethnic groups.” In the 1971 constitution, Muslims of Yugoslavia were promoted from a nationality to a nation.

For a transition from an ethnic group to nationhood, territorial concentration remains central. Dispersed ethnic groups typically demand affirmative action (preference in jobs, education, political representation) and protection of language, religion, and culture. National demands for sovereignty or federalism normally come from territorially concentrated ethnic groups (Québécois, Basques, Sikhs, Kashmiris, Bengali Muslims, Eritreans, Filipino Muslims, Sri Lanka Tamils, Acchnese).

This does not, however, have to be so. The Basques in Spain have had a separatist movement; the Catalans, though territorially concentrated, have not. Tamil Nadu in India saw signs of separatism till 1962; its neighbors Kerala, Karnataka, and Andhra Pradesh never did. All of these are linguistically cohesive, territorially concentrated, and culturally distinctive states. In other words, a conjunction of territorial concentration and ethnicity may be a necessary condition for nationalism, though it is manifestly not sufficient.

\(^8\) See, however, a new proposal in Chandra (2006).
When the national demand goes beyond a federal arrangement of power, the pre-existing larger territorial nationalism is challenged: ethnicity begins to seek territoriality and therefore nationhood. Given that territoriality in the current state system also generally tends to define citizenship, a challenge to the existing notion of citizenship is also posed. Three sacrosanct principles of the nation-state system, thus, become vulnerable: territoriality, citizenship, and sovereignty. Since the number of territorially based ethnic groups is currently larger than the number of nation-states, the existing nation-state system must be considered vulnerable. Some ethnic conflicts may not remain simply ethnic; they may eventually take steps towards separatist nationalism.

1.2 Conflict and Violence

A distinction between violence and conflict is also necessary. In an ethnically plural society, where freedom of expression is not curtailed, some conflict on identity-based cleavages is typically to be expected. Indeed, such conflict may mark all multiethnic polities, authoritarian or democratic. As compared to an authoritarian polity, a democratic political system may simply have a more open expression of such conflicts. In pursuit of political order and stability, authoritarian polities may push ethnic discontent under the surface and induce long phases of ethnic silence, but a coercive outlawing, or forcible containment, often increases the odds of an accumulated outburst, when an authoritarian system starts liberalizing, or when its legitimacy begins to unravel.

Indonesia is an excellent example. During the Suharto era (1966–98), on ethno-communal issues the government had a so-called SARA policy. SARA was an acronym for ethnic (suku), religious (agama), racial (ras), and inter-group (antar-golongan) differences. These differences were neither to be mobilized, nor discussed in the public realm. In the 1980s, Suharto’s Indonesia came to be widely viewed as a stable and well-ordered society. However, by 1998, as the system began to lose its legitimacy, horrendous group violence took place on ethno-communal lines (Bertrand 2004). The former Yugoslavia is another example, although it remains unclear whether ethnic rivalries there were contained more by laws or by an ideological system which, much like the former Soviet Union, sought to create a new communist identity overriding the ethnic and national identities that had so hobbled the Balkans in the first half of the twentieth century.

In contrast, conflicts are a much more regular feature of pluralistic democracies, for if different ethnic groups exist and the freedom to organize is available, there are likely to be struggles over: which language should be used in schools and employment; whether migrant ethnic groups should be allowed entry into the country and/or given restricted rights; whether different groups should be under one civil law for marriages, divorce, and property inheritance, or multiple family laws should be derived from the diverse religious or customary codes; whether religious dress can be allowed in public spaces; whether some groups should be given the benefits of
affirmative action, how, and to what extent; whether the allocation of public resources favors some ethnic groups more than others. India and the United States are good illustrations of how democracies frequently witness such conflicts. Democracy is no guarantee that ethnic conflicts will not flare up. Indeed, some argue that democracies might give politicians incentives to play the ethnic card (Snyder 2000; Wilkinson 2004).

The conceptual issue is whether conflict is violent, or it is pursued within the institutionalized channels of the polity. When ethnic protest is channeled through parliaments, assemblies, and bureaucracies, or when it takes the form of strikes and non-violent demonstrations on the streets, it is an expression of conflict to be sure, but it is not a form of ethnic violence. Such institutionalized conflict, which can be quite healthy for a polity in many ways, must not be equated with riots, pogroms, and civil wars. The explanations for violent and non-violent conflict may also be different.

1.3 Types of Violent Conflict

One more conceptual clarification concerns the various forms of violent conflict. Collective violence, not individual violence or homicides, is at issue here. Collective violence can be defined as violence perpetrated by a group on another group (as in riots and pogroms), by a group on an individual (as in lynchings), by an individual on a group (as in terrorist acts), by the state on a group, or by a group on agencies of the state (as in civil wars).

The most widespread collective violence is typically divisible into three forms—riots, pogroms, and civil wars. Riots refer to a violent clash between two groups of civilians, often characterized as mobs. While, in riots, the neutrality of the state may be in doubt, the state does not give up the principle of neutrality. In pogroms, typically a majority community attacks an unarmed minority, and the principle of neutrality is for all practical purposes dropped by the state. The state administration either looks away, or sides with the attacking group. In civil wars, the state not only abandons the principle of neutrality, but it either becomes a combatant fighting an armed rebel group, or is physically unable to arbitrate between two armed groups fighting each other (Kalyvas 2006).

The key difference between pogroms and civil wars is that in the former, the target group—typically a minority—is hapless and unarmed, whereas in civil wars both combating sides are armed. Riots or pogroms typically precede civil wars, as in Sri Lanka in the 1980s, but all riots and pogroms do not lead to civil wars. Unlike Sri Lanka, the massive 1969 Malay–Chinese riots in Malaysia did not culminate in a civil war, nor for that matter have the Muslim–Christian riots of northern Nigeria in the 1990s led to a civil war there.9

9 The last civil war in Nigeria took place in the late 1960s. It had nothing to do with Muslim–Christian divisions. It was ethnically driven.
2 Traditions of Explanatory Enquiry

2.1 Essentialism

Essentialism is the oldest tradition of enquiry in the subfield of ethnicity and has been seriously under attack of late. It emerged at a time when the early enthusiasm, witnessed at the birth of the newly decolonized nations after the Second World War, had begun to ebb. In country after country, the story seemed to be similar. Nation building encountered serious ethnic resistance from within. Why was that so? Why could smaller ethnic identities not be subsumed under larger country- or state-level identities that governments were ostensibly seeking to create?

The first scholarly response was simply that the decolonized states were new, but ethnic, or communal animosities—sometimes also called national animosities—were old and, therefore, deeply historically rooted. The primordialism of ethnic groups was a stronger bond and a more powerful motivator of human conduct than the pull of civic ties being forged by the new states (Geertz 1963; Shils 1957). This view found its most systematic exponent in Connor (1972, 1994). As late as the early 1990s when, as discussed later, the constructivist attack on essentialism was at its full cry and only journalists were willing to use the term “ancient hatreds” (Kaplan 2003), Connor was willing to argue that “man is a national,” not rational, “animal,” and at the core of nationalism lay the notion of “shared blood” or “shared ancestry” (Connor 1994).

Essentialism in this form had three primary weaknesses. The first had to do with variations. If ethnic antagonisms were so deep-rooted, why did ethnic violence rise and fall at different times? Yugoslavia may have come apart with a nearly all-consuming violent thud in the 1980s and 1990s, but there was a long stretch of peace during the socialist period. Do institutional designs not change human motivations? Did violence at the time of Yugoslavia’s break-up show that in times of state collapse, ethnic antagonisms flare up, or that ethnic hatreds caused the collapse of Yugoslavia? Another type of variation is interspatial. Why did the same groups live peacefully in some places, but not in others? Hindu-Muslim violence often flared up in certain parts of India, not all over India (Varshney 2002; Wilkinson 2004).

Second, a lot of ethnic conflict in the world had nothing to do with old hostilities at all. Rather, older inhabitants of a land clashed with a relatively new migrant group, with little or no long history of contact. Can one establish the “primordial” or “ancient” roots of Chinese–Malay violence in Malaysia? The Chinese, after all, arrived in Malaysia mostly in the nineteenth and twentieth centuries. Similarly, it was primarily in the nineteenth and twentieth centuries that the Chinese came to Indonesia, and the Ibos flowed to northern Nigeria. Yet the anti-Chinese violence in Indonesia and the anti-Ibo violence in the Hausa-dominated northern Nigeria in the twentieth century was as ferocious as that between Hindus and Muslims, both older groups, in India.

The third attack on essentialism emanated from what came to be called the constructivist school. To talk about nations having primordial animosities, the constructivists argued, was wrong. In arguments that over time shook the
foundations of essentialism and became mainstream wisdom, constructivists argued that nations were constructed only in modern times (Anderson 1983). Before the rise of modernity, most human interactions were on a small scale. Only ecclesiastical and dynastic communities spread beyond the local and the regional. The implication was that religious or dynastic animosities could be said to be pre-modern, even primordial, but ethnic animosities had local or regional protocols. By bringing far-flung people into the frame of human consciousness, it is modernity that changed the meaning of ethnicity and also led to nationhood. To speak of primordial ethnic or national antagonisms was historically false.

Essentialism, however, did not fully disappear, as was predicted and expected. These attacks—variations, modern provenance of conflicts, and constructivism—led to a fresh honing of arguments. Accepting the inadequacies of a Connor-style argumentation, Petersen (2002) recast essentialism with psychological theories about emotions.

On "ancient hatreds," he argued:

Most academics dismiss the "ancient hatreds" argument. They show how violent interethnic "histories" are often fabrications, inventions that serve the interests of rabble-rousing elites. If "ancient hatreds" means a hatred that has produced uninterrupted ethnic warfare, or an obsessive hatred consuming the daily thoughts of great masses of people, then the "ancient hatreds" arguments deserves to be readily dismissed. However, if hatred is conceived as a historically formed "schema" that guides action in some situations, then the conception should be taken more seriously. (Petersen 2002, 62–3. emphasis mine)

In short, the existence of hatred did not require a proof about its ancient origins. Even if hatred had non-ancient origins, it could profoundly shape human behavior. Human nature was quite capable of expressing hatred. In what might be called a neoessentialist twist, Petersen turned an argument about primordial hatreds into an argument about human nature:

the motivation to participate in or support ethnic violence and discrimination [is]... inherent in human nature. Until we realize that the capacity to commit ethnic violence lies within all of us we are in danger of constantly being surprised at the emergence of forces from the "dark ages." (Petersen 2002, 1)

Petersen built four models, based on four different kinds of emotions: fear, hatred, resentment, and rage. Fear as an emotion guides individuals in situations of security threats; hatred in conditions of historical grievance; resentment in settings of status discrepancies; and rage simply expresses a desire to "lash out" due to accumulated emotions, but without a specific target. A prediction was made with respect to how each emotion would work, and a test devised in Eastern Europe. Petersen's general argument, finally, was that resentment born out of status reversals explained most of the ethnic violence in twentieth-century Eastern Europe during periods of state collapse (during and after the First and Second World Wars, and at the end-game of communism). Hatred, fear, and rage explained fewer cases of violence, but they were also present.
Emotions have been, on the whole, neglected in social science theories about ethnic conflict. It is now clear that the field will have to engage Petersen’s neo-essentialist arguments. One potential line of engagement is obvious. What is the role of institutions in reining in, or redefining, emotions? Why do these emotions explode in times of state collapse, not when state authority is firmly anchored? Does that variation indicate something about our deep-seated human nature, or about the causal role of institutions, in the outbreak of ethnic violence? A second question is about whether state collapse, even in Eastern Europe, necessarily leads to horrendous violence. Laitin (1998) argues that new identity formation after the end of communism was peaceful in the Baltic republics and Kazakhstan. What accounts for such dramatic variation?

2.2 Instrumentalism

The core idea of instrumentalism is that ethnicity is neither inherent in human nature nor intrinsically valuable. Ethnicity masks a deeper core of interests, which are either economic or political. Ethnicity is useful for gaining political power or for drawing resources from the state. That is why it is deployed so often in multiethnic societies. Conflicts take place because leaders strategically manipulate ethnicity for the sake of political power, or for extracting resources from the state (Bates 1974, 1983; Chandra 2004; Hechter 1986; Rabushka and Shepsle 1972).

This line of reasoning runs into several difficulties. Even if we accept that leaders gain by mobilizing ethnicity and that is why they deploy ethnic symbols and idioms in politics, why should the masses come along? Why do leaders in multiethnic societies so often think that ethnicity is the means to power or for extracting resources from the state, not mobilization based on economic or ideological programs? Second, if the masses were also instrumental, would ethnic collective action not be crippled by free rider problems? One can perhaps understand why it would be instrumentally rational for someone to join an ethnic movement when it is close to capturing power, but why would ethnic mobilization begin at all? A fuller account or inclusion of “selective incentives” (Olson 1965) or “commitment” (Sen 1973) is required. Third, if ex-ante odds are quite high that ethnic mobilization or protest would lead to violence by another group, or to punitive action by the state, why should anyone participate in ethnic mobilization at all? Why would instrumentally rational people take such high risks? One could propose that people are coerced into participating in ethnic mobilization, but that would have to be demonstrated, not assumed.

In different ways, some of the more widely noted instrumentalist scholarship of the last decade and a half seeks to address these problems (Hardin 1995; Fearon and Laitin 1996; Collier and Hoeffler 1998, 2004; Collier, Hoeffler, and Sambanis 2005). Extending the idea of “focal points” originally proposed by Schelling (1963), Hardin argues that the central strategic problem in ethnic mobilization is one of

---

10 These criticisms are based on Horowitz (1985, 2001), and Varshney (2003).
coordination, not one of collective action. In the latter, it is rational to "free ride," but in coordination games, it is rational to cooperate so long as others are cooperating. A "charismatic leader," a "focus," is what one needs to reinforce expectation about the behavior of others.

To understand this point better, it is worth recalling the famous Schelling example:

When a man loses his wife in a department store without any prior understanding on where to meet if they get separated, the chances are good that they will find each other. It is likely that each will think of some obvious place to meet, so obvious that each will be sure that the other is sure that it is obvious to both of them.11

The "lost and found" section of the department store, argued Schelling, could serve as one such obvious place, but not if there were many "lost and found" sections in the store. "Prominence" or "uniqueness" distinguished a focal point. That is why it could be used to facilitate the development of mutually consistent expectations. Seen this way, ethnicity could serve as a "focal point," and ethnic mobilization would simply require coordination of expectations. Ethnicity did not have to be intrinsically valued for it to be politically useful.

Though ingenious, this resolution has its own difficulties. Two come to mind immediately (Varshney 2003). First, why is ethnicity-based mobilization akin to a coordination game, but class-based mobilization a form of collective action saddled with free rider problems? Hardin's answer is that ethnicity provides "epistemological comforts of home," but that restates the problem. Why should ethnicity provide these comforts, not class or party? After all, the Marxists-Leninists had believed for much of the twentieth century that the Communist Party would be home to the new socialist man, replacing ethnicity or nationhood. Second, why should it be easier to mobilize ethnicity, despite the risks of injury, incarceration, or death? Saying ethnic mobilization is a mere coordination problem does not square with the well-known risks of ethnic conflict. In short, can one really explain ethnic preferences in an entirely instrumental way, or is recourse to the psychological or cultural foundations of ethnicity necessary?

Fearon and Laitin (1996) respond to these difficulties by restricting the domain of instrumental rationality, even while using instrumentalist assumptions for developing their core argument. Instead of asking why there is so much ethnic conflict and violence in the world, they first note that the incidence of ethnic violence is lower than is normally believed. Instead of engaging in killings, many ethnic groups, in fact, live in peace. There is a gap between actual violence and what is theoretically possible.

What would explain inter-ethnic peace and cooperation? Relying on the notion of ethnic groups as information networks, they game-theoretically generate a powerful and unexplored idea as an equilibrium solution: "in-group policing."12 Faced with provocation or attacks, a group could restrain its members from hitting back, and

11 Schelling (1963, 54). We are, of course, talking about the pre-cellphone days.
12 It should be noted that "in-group policing" remains a deductive idea, still to be systematically and empirically tested. The fear of "spiralizing" produces the other equilibrium solution, meaning individuals of one group could be expected to attack the other group indiscriminately in response to an attack, which could lead to escalating violence, which in turn would induce cooperation.
rely on similar restraining exercised by the other group. This is possible because each ethnic group has better information about its own members than about those of the other group, which in turn can allow each group to check who the in-group “opportunists” are, meaning those who would use the provocation to retaliate.

Does this mean that explanation of ethnic conflict requires no recourse to psychological theories of grievance? Careful not to make universalistic claims, Fearon and Laitin explicitly lay out the limits of their theory:

We should emphasize...that we are not offering a full causal theory of either ethnic peace or ethnic violence. We specify what we believe are important causal mechanisms that appear to have been systematically neglected... But we do not pretend that our formulation or...mechanisms we identify tell the whole causal story. A richer story would surely include...narratives of interethnic injury. It might also include the motivations stemming from indignities suffered by peoples who are considered of lower rank and who seek to overturn a rigid social ordering. (Fearon and Laitin 1996, 715)

In short, an instrumental use of ethnicity—in this case, ethnicity as a communication and information device rather than an intense form of group attachment—may explain part of the phenomenon of violence, but historical indignities and injuries may well be relevant. Exploring a variety of conflicts in different settings should begin to show which motivations are present where. This argument leads to the possibility that ethnic conflict could have pluralistic microfoundations.

Let me now turn to another new argument. Collier and Hoefller (1998, 2004) are associated with the famous “greed versus grievance” framing of ethnic violence. Though they concentrate only on civil wars, an extreme form of ethnic conflict, their argument is worth considering here. The strength of their belief in instrumentality has evolved in an educative manner.

Based on a large-N statistical model, Collier and Hoefller (1998) first argued that social scientists had been wrong to believe that civil wars were the consequence of accumulated grievances of a victimized or targeted ethnic group in society. A grievance-based argument was simply equal to accepting the discourse of rebels. Instead, a greed-based model had an infinitely better fit with data.

They model rebellion as an industry in which looting generated profits. Leaders of rebellions are driven by a desire to amass fortunes, and the masses join them, for in poor societies with very few economic opportunities, the opportunity costs of participation in a rebellion are low and the benefits—in the form of a share of the loot—quite substantial. Given their geographical concentration, natural resources are an especially “lootable commodity.” Civil wars predominantly erupt in economies highly dependent on natural resource extraction.

This was, arguably, the sharpest framing of the instrumentalist view ever witnessed in the field. And the notion that both models—greed and grievance—were tested with a large-N dataset added a new punch. Instrumentalist arguments used to be about ethnic mobilization, something decidedly less violent than civil wars, and large-N datasets were rarely, if ever, used for testing.

However, as the Collier–Hoefller argument evolved, its sharpness and universality steadily diminished. When the dataset was enlarged, coding further finessed, and
model specifications changed, they concluded that “we cannot reject one model in favor of the other” and “while the (greed) model is superior, some elements of the grievance model are likely to add to its explanatory power” (Collier and Hoeffler 2004, 577). Still later, the findings of the statistical model were subjected to carefully chosen case studies because even if the statistical model “predicted all cases of civil war onset perfectly, it would still not be able to tell us much about the process through which these outcomes (war or peace) are generated. By contrast, analyzing the process—the sequence of events and the interaction of variables in the (statistical) model over time—is the comparative advantage of case study designs.... Qualitative analysis can help us sort out the endogenous from the exogenous variables in the model” (Collier Hoeffler, and Sambanis 2005, 2).

What, then, was the final conclusion? “The distinction between greed and grievance,” they argued, “should be abandoned for a more complex model that considers greed and grievance as inextricably fused motives for civil war” (Collier, Hoeffler, and Sambanis 2005, 2).

Just as pure essentialism could not survive empirical scrutiny, pure instrumentalism also could not. Future work in this tradition is likely to be highly domain specific. Instrumental uses of ethnicity do exist and will continue to. But one will have to be clear about the kinds of questions for which either instrumentalist assumptions can be made, or instrumentalist claims can be sustained. Not all forms of ethnic behavior, or ethnic conflict, can be linked to instrumental rationality.

### 2.3 Constructivism

Constructivism is the new conventional wisdom in the field of ethnicity and nationalism. Its central idea is that our ethnic and national identities are constructs of the modern epoch. This claim is relatively straightforward for national identities, for work across a whole range of traditions shows that nations were born with the rise of the industrial age.13 Political units took the form of city-states or empires before that.14

But the argument is also made with respect to ethnic identities (Hobsbawm and Ranger 1983; Mamdani 1996; Vail 1989; Suny 2001). The claim is not that there were no Turks, Han Chinese, Tibetan, Zulus, or Scots in pre-modern times. Rather, in pre-modern times, mass identities were locally or regionally based. Only some kinds of identities—for example, the aristocratic or ecclesiastical—were extra-local or extra-regional. Modernity transformed the meaning of ethnic identities by bringing the masses into a vastly expanded framework of consciousness and meanings.

Three kinds of mechanisms have generally been identified to show how this came about—technological, ideational, and, in the former colonies, colonial policies, institutions, and practices. Anderson’s *Imagined Communities* (1983), arguably the

---

13 However, see Kedourie (1993).
14 For the link between nations and modernity, see, among others, Greenfeld (1992).
most influential text in the field of ethnicity and nationalism, calls attention to the rise of “print capitalism”—the arrival of the printing press and capitalism—as the basic mechanism through which local identities were transformed into larger national identities. The boundaries of the political community typically depended on the spread of the vernacular and the decline of “truth languages” such as Latin or Sanskrit.

The second mechanism that constructivists now routinely embrace relies on the arguments made by Taylor (1994) about how modernity brought about ideational changes in human life. In pre-modern times, one’s identity—who am I?—was given by one’s place in the traditional social structure. People accepted ascriptive social hierarchies, or their “stations” in life. In modern times, hierarchies may exist, but ascriptive hierarchies are not easily accepted. In pre-modern times, the notion of honor, reserved only for a few, marked relations between people of different ranks. Modernity has introduced us to the notion of dignity; to which all regardless of rank are entitled. Finally, the pursuit of dignity is dialogical, not monological—that is, it takes place in interaction with others. Our identity is partly shaped by recognition or its absence, often my misrecognition of others, and so a person or group of people can suffer real damage, real distortion, if the people or society around them mirror back to them a confining, demeaning, or contemptible picture of themselves. Nonrecognition or misrecognition can inflict harm, can be a form of oppression, imprisoning someone in a false, distorted and reduced mode of being. (Taylor 1994, 25)

Though acceptable and internalized in the past, much of ethnic or national assertion in the modern world is about resisting such “confining, demeaning or contemptible” pictures that the dominant groups—through colonial rulers and state bureaucracies—have often relayed to the subordinate groups. The keyword here is dignity, not material self-interest.

Given the heavy reliance on historical detail, constructivism first flourished in the discipline of history (Weber 1976; Hobsbawm and Ranger 1983; Vail 1989). In comparative politics, Anderson (1983) was the first to make the constructivist argument. Some others also joined in, especially concentrating on the structure of colonial rule and colonial policies. Laitin (1986) explained why in Yoruba politics in Nigeria, religious cleavage was missing, even though both Islam and Christianity dominated the religious landscape. He argued that for reasons of their own, the British did not allow religion to be the basis of politics in Yorubaland, electing instead to emphasize tribal cleavages. By the time they left, the tribal cleavages were so deeply institutionalized that they became the political common sense of Yorubaland. Chatterjee (1986) argued that the images of Indians British rulers created and propagated,

---

15 According to “Google scholar”, as of February 1, 2007, Anderson’s Imagined Communities had been cited over 6,300 times, followed by Gellner’s Nations and Nationalism (1,149 times), and Taylor’s Multiculturalism and the Politics of Recognition (1,205 times). All three texts are covered in this chapter.

16 Technical change—in the form of print capitalism—is the centerpiece of Anderson (1983), but Taylor’s ideational change is often implicit in his arguments about the birth of nationalism in the former colonies. “Creole pioneers” felt humiliated in Spanish America, rebelling against the Spanish rulers (Anderson 1983, ch. 4).
once they conquered India, led to the development of nationalism in India. These three examples notwithstanding, constructivism remained on the margins of how ethnicity was studied in comparative politics in the 1980s and for much of the 1990s. Constructivism’s rise in other disciplines, especially history, far preceded its incorporation in political science.

Although it is by now customary to state that constructivism has become the dominant mode of argumentation about ethnicity within political science as well, one of its key weaknesses ought to be noted. Constructivism accounts for identity formation well, but it does not do a good job of explaining ethnic conflict. Often, a distinction between identities and conflict is not drawn. The key constructivist idea on conflict is that each society has a historically constructed “master cleavage”—Protestant versus Catholic in Northern Ireland, Hindu versus Muslim in India, black versus white in the USA—and political entrepreneurs can easily insert local, often trivial, incidents, events and rumors into the “master narrative,” creating inflammable situations and instigating violence (Brass 1997, 2003). In social scientific terms, a causal role is thus assigned to master narratives and political entrepreneurs.

The problem is that the master cleavage is typically at the national level and political entrepreneurs are also available throughout the length and breadth of a country, but ethnic violence tends to be highly locally, or regionally, concentrated. In the 1960s, racial violence in the USA was heavily concentrated in northern cities; southern cities, though intensely politically engaged, did not have riots (Horowitz 1983). A mere eight cities in India, holding less than 6 percent of the country’s population, accounted for just a little less than half of all deaths in Hindu-Muslim riots during 1950–95 (Varshney 2002). Between 1990 and 2003, fifteen districts of Indonesia, in which less than 7 percent of the nation’s population lived, had close to 85 percent of deaths in all forms of group violence short of secessionary wars (Varshney, Panggabean, and Tadjoeddin 2006).

How can one explain local variations with a nation-level constant (“master cleavage,” “master narrative”) and the countrywide ubiquity of political entrepreneurs? The answer perhaps lies in (a) how local structures of some kind discourage political entrepreneurs from inserting local incidents into the master narrative; or (b) how political entrepreneurs are unable to instigate violence even when they insert local events into the larger narrative; or (c) how the presence of local or regional narratives counters the power of a master narrative (Varshney 1997). Constructivist arguments about violence are thus far built on case studies of violence, not on a comparison of peaceful and violent cases. Selection bias has led to significant weaknesses; studying variations has explanatory promise.

A final question about constructivism remains. Are constructivism and instrumentalism merely two sides of the same coin? Chandra (2001) has argued that divisions in the field of ethnicity and nationalism should simply be viewed as those between essentialists and constructivists. According to her, Geertz (1963) is an example of essentialism, and constructivist arguments include not only those made by Anderson (1983) and Laitin (1986), reviewed above, but also those made by Bates (1974), included here as an example of instrumentalist reasoning. To recall, the
latter argument is that ethnicity is a conduit for extracting resources from the state; nothing more need be said or assumed. According to Chandra, what distinguishes constructivism is the notion that "ethnic groups are fluid and endogenous to a set of social, economic and political processes" (Chandra 2001, 7).

The instrumentalist and constructivist approaches are undoubtedly opposed to the primordialist view of ethnicity, but that is where the similarity ends. According to instrumentalist reasoning, ethnic identity is not valuable in and of itself; it is basically a mask for a core of "real" interests, political or economic. As interests change, masks also do, making ethnic groups "fluid." One should, therefore, expect the same people to pick different sides of their multiple identities at different times and at different places.

This view should not be equated with constructivism. Constructivism is not about the radical short-run fluidity of identities. It is about the long-run formation, and the consequent stickiness, of identities. In Anderson's case, the argument is epochal: he discusses how the birth of print capitalism in modern times created national identities. Weber (1976) shows how peasants were turned into Frenchmen over more than a century after the French Revolution—through a conscription army and public schools. Colley's argument is about how "Britishness" emerged out of "Englishness," "Scottishness," and "Welshness" over more than a century (1707–1837), and how the presence of France as a "Catholic enemy" and a colonial empire especially blunted the historically rooted intensity of English–Scottish rivalries (Colley 1993).

Each of these scholars demonstrates how new identities came about, but it does not follow that they view identities as radically fluid. That identities are constructed does not mean that they do not become internalized and institutionalized, and acquire meaning.

Constructivism is basically about the long-run stickiness, instrumentalism about the short-run fluidity. While equally opposed to primordialism, they are fundamentally different in their assumptions, explanatory ambition, and methodological impulse.

17 One should also note that Laitin (1986) was profoundly opposed to an instrumental view of ethnic identity formation, though his positions changed later (Fearon and Laitin 1996; Laitin 1998). Consider the following arguments in the earlier book:

Rational choice theorists...cannot tell us if ultimately butter is better than guns; it can tell us that at a certain point the production of a small number of guns will cost us a whole lot of butter, and at that point it is probably irrational to produce more guns. Within a political structure, individuals constantly make marginal decisions. (Rational choice) theories can give us a grasp on how individual political actors are likely to make choices within that structure.

(Rational choice) theory cannot, however, handle long-term and non-marginal decisions. When market structures are themselves threatened, and people must decide whether to work within the new structure or hold on to the old—without an opportunity for a marginal decision—microeconomic theory is not applicable.... Structural transformations—changing the basic cleavage structure of a society—are not amenable to the tools of microeconomic theory. (Laitin 1986, 148–9)

Identity choice was not a marginal, but a structural decision. Instrumental rationality, therefore, was inapplicable.

18 The relationship between the long-run stickiness of some identities and short-run fluidity of others may have to be sorted out, but that is another matter altogether. The two should not be conflated.
2.4 Institutionalism

If constructivism has come to shape the literature of the formation of ethnic identities, institutionalism has long dominated the arguments about ethnic conflict in comparative politics. The core idea here is that the designs of political institutions—consociational or majoritarian polities, proportional representation or first-past-the-post electoral systems, federal or unitary governments—explain why some multiethnic societies have violence, and others, peace.

Ethnic pluralism, it is argued, requires political institutions distinct from those that are suitable for ethnically undivided societies. A mechanical transfer of institutional forms regardless of whether a society is marked by deep ethnic divisions can cause ethnic violence. The foundations of such arguments go all the way back to John Stuart Mill in the nineteenth century. Mill had claimed that common loyalty to a political center was a precondition for a democracy to function. A multiethnic society was likely to have many loyalties, not one. Only under the tutelage of a more politically advanced ethnic group can order be maintained and ethnic violence avoided. Tutelage was necessary until a civic consciousness towards a political center, not to an ethnic group, emerged.

Nobody can suppose that it is not beneficial to a Breton or a Basque of the French Navarre to be brought into the current of ideas and feelings of a highly civilized and cultivated people—to be a member of the French nationality... than to sulk on his own rocks, the half-savage relic of past times, revolving in his own little mental orbit, without participation or interest in the general movement of the world. The same remark applies to the Welshman or the Scottish Highlander, as members of the British nation. (Mill 1990, 385–6)

Colonial tutelage is no longer popular, but the arguments about whether multiethnic societies should have majoritarian democracies continue to be debated. Lijphart (1977) and Horowitz (1985, 1991) have defined the field. Lijphart continues to argue in favor of consociationalism, in which each ethnic group’s political and cultural affairs are left to its elite, and inter-ethnic compromises are made only at the elite level. Horowitz argues against it, suggesting that the electoral system should make it impossible for political parties to win power unless they appeal across ethnic groups, not lock them in a permanent intra-ethnic embrace. The former is more likely to lead to peace, the latter to violence.

This debate has greatly advanced our understanding of ethnic conflict. However, it has left one big problem unresolved. The Lijphart–Horowitz arguments have basically been about national-level institutions. Using national-level concepts, we certainly explain why country A, rather than country B, tends to have more ethnic violence, but we cannot understand the regional or local variations within the same country. For institutional explanations to be relevant to local or regional variance, the electoral designs or institutions must themselves vary locally or regionally.

The neo-institutionalist work of recent vintage goes in the direction of uncovering local institutional variations. Varshney (2002) argues that local variation in conflict is

19 For a review, see Reilly (2001).
best explained by whether local civic organizations, including political parties, exist and whether they integrate ethnic communities or segregate them. Wilkinson (2004) argues that in a first-past-the-post electoral system, it is the effective number of parties and the need for minority support—both of which can vary regionally and/or locally—that determine whether ethnic violence will occur or peace will obtain.

A second new development in the literature is the focus on the relationship between institutions and identity choice. Lijphart (2001) accepts that when consociational theory was developed in the 1960s and 1970s, an essentialist view of identities prevailed. In line with those times, he also assumed that ethnic identities were fixed, and appropriate political institutions were to be constructed in light of the fixity of ethnic identities.

The new literature shows how institutions can transform the salience of identities. Posner (2005) argues that since colonial times, Zambians have had two axes of identification: language and tribe. Zambia has four language groups and over six dozen tribes. Since independence, Zambia has also had two kinds of overarching institutions: multiparty rule and one-party rule. Under the former, Zambians embraced language as the basic political identity, and under the latter, they chose tribe. Under a multiparty system, they had to elect a constituency representative as well as the president. This meant that the political arena was national, and the larger identification (language), therefore, made sense. Under one-party system, only the constituency representative was to be elected, not the president. The political arena was, thus, reduced to the constituency level, and the smaller identification (tribe) became more relevant.²⁰

Such reasoning, it should be noted, was implicit in Horowitz (1985). His critique of consociationalism was, in part, based on the fact that identities could change and the elite of an ethnic group, therefore, could not be expected to keep the loyalty of that group for ever. He also argued that the changing political arena would reshape the cleavages. But in the new literature, this idea is explicit. Identity choice is squarely posed as a dependent variable to be explained. As a result, we have a more self-conscious and focused explication of the institutional determinants of identity choice (Chandra 2004; Laitin 1998; Posner 2005; Waters 1990).

A marriage of constructivism and institutionalism is the third new development in the literature. To recall, a general stickiness of master cleavages is the core idea of constructivism. Institutionalism, in comparison, has begun to accept fluidity of identities, depending on the institutional context. Can historical stickiness and conjunctural fluidity be combined?

Posner (2005) begins to show how.²¹ The contemporary choices between language and tribe in Zambia may be determined by whether the country has a one-party or

---

²⁰ It should, however, be noted that Posner's argument, though presented as one about identity choice, could as easily be constructed as an argument about electoral choice. It is not the Zambian identities that changed with the alteration in the party system, but only how Zambians voted.

²¹ Posner argues that he is, in effect, combining constructivist, institutional, and instrumental-rational arguments. The last does not appear to be true. Technically speaking, an instrumental-rational view requires that (a) the microfoundations be defined in terms of self-interest, which is not affected by "framing," and (b) given those microfoundations, collective action problems be resolved, for group
multiparty rule, as noted above. But colonial history, argues Posner, had already deeply institutionalized only two identities: linguistic and tribal. This was because of the administrative and recruitment policies and census practices of British rulers and companies. Several other axes of identification were conceivable.

The innovative marriage of constructivism and institutionalism on identity formation/choice faces some tougher challenges ahead. Van Evera (2001) has posed the important question of whether identities can be fluid, if formed or deepened by violent conflict. In other words, is Zambia an easy case? In Zambian history, is there anything like India’s Hindu-Muslim violence at the time of partition, Malaysia’s Malay–Chinese violence 1945–69, Sri Lanka’s Sinhala–Tamil violence since 1977, and the several descents into ethnic warfare in the Balkans? Van Evera claims that if violent conflict constructs, or deepens, identities, they cannot be easily reconstructed. Future research under the marriage of constructivism and institutionalism may have to respond to this challenge.

3 Conclusion

Three conclusions can be drawn from the arguments above. First, if one thinks of cumulation in Popperian terms—as progress through a systematic disconfirmation of theories—then only two theoretical ideas have been knocked over in the last ten to fifteen years. No one seriously argues any more that ethnic identity is primordial, nor that it is devoid of any intrinsic value and used only as a strategic tool. Pure essentialists or pure instrumentalists do not exist any longer. Nor is it likely that they will re-emerge, given the force of empirical evidence. Second, the traditions which produced these theories, however, continue. Innovation within has taken place, or a new set of unresolved problems promises innovation. Essentialism has moved towards an argument about human nature, especially in conditions of state collapse. Instrumentalism has sought to restrict its domain, or begun to think of models in which “greed and grievance” will be “inextricably fused.” Constructivism has to sort out whether subnational and local variations in conflict can be explained within its own guiding assumptions and principles. Institutionalist has to ascertain whether identities are fluid only under some circumstances, and how fluid they are. Third, the field has become methodologically highly self-conscious and sophisticated and that is only to be welcomed. However, methodological disputes or
methodological advances alone will not generate significant progress. Some of the most creative work in the future is likely to be problem and puzzle driven and may well emerge from border crossings and mixed approaches. Of course, not all borders can be crossed. It is, for example, not clear whether essentialism and instrumentalism can ever be brought together without grotesque internal inconsistencies. But border crossings between constructivism and institutionalism have been initiated, and should certainly be more easily possible. The results could be highly instructive.

References


22 For a recent methodological debate in the field, see the symposium on David Laitin's work Qualitative Methods (2006).


—— 2006. Recognizing the tradeoffs we make. *Qualitative Methods*, 4 (1).


