Identity as a Variable

Rawi Abdelal, Yoshiko M. Herrera, Alastair Iain Johnston, and Rose McDermott

As scholarly interest in the concept of identity continues to grow, social identities are proving to be crucially important for understanding contemporary life. Despite—or perhaps because of—the sprawl of different treatments of identity in the social sciences, the concept has remained too analytically loose to be as useful a tool as the literature's early promise had suggested. We propose to solve this longstanding problem by developing the analytical rigor and methodological imagination that will make identity a more useful variable for the social sciences. This article offers more precision by defining collective identity as a social category that varies along two dimensions—content and contestation. Content describes the meaning of a collective identity. The content of social identities may take the form of four non-mutually-exclusive types: constitutive norms; social purposes; relational comparisons with other social categories; and cognitive models. Contestation refers to the degree of agreement within a group over the content of the shared category. Our conceptualization thus enables collective identities to be compared according to the agreement and disagreement about their meanings by the members of the group. The final section of the article looks at the methodology of identity scholarship. Addressing the wide array of methodological options on identity—including discourse analysis, surveys, and content analysis, as well as promising newer methods like experiments, agent-based modeling, and cognitive mapping—we hope to provide the kind of brush clearing that will enable the field to move forward methodologically as well.

Introduction

Redeeming identity for the social sciences

For the past two decades, the attention given to the concept of identity—both in the social sciences and in the world at large—has continued to rise. Multiple disciplines and subfields are producing an expanding literature on the definition, meaning, and development of ethnic, national, linguistic, religious, gender, class, and other identities and their roles in processes of institutional development. Yet despite this flurry of activity, the social sciences have yet to witness a commensurate rise in the analytical rigor with which the word gets used.

The ubiquity of identity-based scholarship suggests an emerging consensus that identities, as Rogers Smith has observed, are “among the most normatively significant and behaviorally consequential aspects of politics,” yet the literature has been plagued by a big problem. To the chagrin of the social scientific community, it is in large part this same ubiquitous sprawl of scholarship that has undermined the conceptual clarity of identity as a variable. The wide variety of conceptualizations and definitions of identity have led some to conclude that identity is so elusive, slippery, and amorphous that it will never prove to be a useful variable for the social sciences. Rogers Brubaker and Frederick Cooper have even argued, in the most important critique of identity scholarship to date, that it is time to let go of the concept of identity altogether, and to move beyond a scholarly language that they suggest is hopelessly vague and has obscured more than it has revealed. Even we must concede that the current state of the field amounts to definitional anarchy.
Yet we, on the other hand, are not so quick to jettison a generation’s worth of scholarship, much of it provocative and valuable. Instead, we feel it is time—past time, really—to invest identity with the analytical rigor and methodological imagination needed to make it a measurable variable across the social sciences. Nor do we think it need be especially difficult to do so. To the contrary, the elements of our new analytic framework for identity, which we propose in this article, are already implicit in the existing literature.

Social identity scholarship suffers from two sets of problems: conceptual issues and coordination gaps. The main conceptual questions that the field has yet to answer satisfactorily are: how can we compare different types of identities; and how can we exploit theoretical advances in operationizating identity as a variable? Among “coordination” problems we include the lack of consistency and clarity in defining and measuring identities, the lack of coordination of identity research at both the cross-disciplinary and cross-sub-field levels, and missed opportunities to take advantage of expanded methodological options.

The analytic framework developed in this article addresses these problems and offers a way forward. Our article offers a definition of collective identity as a social category that varies along two dimensions—content and contestation. The progress of scholarly work on social identities depends on developing an analytic framework that allows for comparison and differentiation among the many kinds of identities; being able to differentiate between types of content is the key to such a framework. Content describes the meaning of a collective identity. The content of social identities may take the form of four, non-mutually-exclusive types: constitutive norms; social purposes; relational comparisons with other social categories; and cognitive models. Contestation refers to the degree of agreement within a group over the content of the shared category. Our conceptualization thus enables collective identities to be compared according to the agreement and disagreement about their meanings by the members of the group.

Finally, our article takes stock of the methodological options for identity work. So far, identity scholarship has limited itself to a narrow methodological band, taking little notice of newer, less traditional options that have much to offer the field. Consequently, many existing efforts to measure identity have been either too hard and simplistic (relying on blunt survey instruments or census data, for example) or too soft and impressionistic (such as relying on the individual scholar’s account of identity narratives).

We are proponents of methodological eclecticism, particularly with regard to identity work. Addressing the range of methodological options, we advocate six as being especially well-suited: discourse analysis, surveys, and content analysis, and newer methods like experiments, agent-based modeling, and cognitive mapping.

Our article thus offers two paths for social scientific work on identity—by developing a more rigorous, more precisely defined analytic framework, and by providing a methodological roadmap for further integrated progress in identity scholarship.

**Analytic Framework for Identity as Variable**

We believe that the problem at the heart of identity scholarship is the absence of an analytic framework that is broad enough to serve the majority of scholars working on identity, yet narrow enough not to include social phenomena that are distinct from, though related to, identity. Building upon the brush-clearing work already done by others, we took upon ourselves the task of developing an analytic framework that will 1) enable scholars to compare types of identities (e.g., ethnic, national, religious, gender, class, etc.); 2) allow for nuanced operationalization of the theoretical sophistication of identity scholarship; and 3) promote coordination across identity scholarship while providing a conceptualization that is flexible enough to allow researchers to tailor it to their own particular needs. Moreover, we believe the conceptualization of identity in our analytic framework is already implicit in almost all of the research on identity that we have surveyed; what we have to offer is not something new or out of left field but rather a conceptual apparatus that allows for the integration, not just of future scholarship, but especially of the mountain of already existing scholarship on which it is based and whose implicit assumptions it draws out into the open.

We define a collective identity as a social category that varies along two dimensions—content and contestation. Content describes the meaning of a collective identity. The content of social identities may take the form of four, non-mutually-exclusive types:

- **Constitutive norms** refer to the formal and informal rules that define group membership.
- **Social purposes** refer to the goals that are shared by members of a group.
- **Relational comparisons** refer to defining an identity group by what it is not, i.e., the way it views other identity groups, especially where those views about the other are a defining part of the identity.
- **Cognitive models** refer to the worldviews or understandings of political and material conditions and interests that are shaped by a particular identity.

Contestation refers to the degree of agreement within a group over the content of the shared identity. Far from being understood as fixed or unvarying, collective identities, in this conceptualization, vary in the agreement and
disagreement about their meanings. The relevant aspects of the content of collective identities vary a great deal, and it is impossible a priori to specify them all.

We believe that these four types of content encompass the variety of meanings in social identities, while contestation over content addresses the fluidity and contextual nature of identities. Moreover, we argue that every social identity includes all of these types of content, with greater or lesser degrees of contestation over aspects within content types.

Our analytic framework for identity has much in common with definitions and conceptualizations already used implicitly by many scholars, and the words “content” and “contestation” are standard in the scholarly literature. The literature includes many other words for the variation in identities as well, but we hold that such alternative conceptualizations and nomenclatures are either subsets of content and contestation, or simply not useful enough to remain part of the lexicon of identity. We developed this typology while surveying the existing literature, where scholars were already measuring identity either implicitly or explicitly along these lines. These four types, each illustrated in detail below, encompass the range of variation for which the field has so far tried to account.

**Constitutive norms**

The normative content of a collective identity specifies its constitutive rules—the practices that define that identity and lead other actors to recognize it. The rules that determine group membership and putative attributes of the group can also be thought of along these lines. This normative content, the set of constitutive rules, may be bundled together into one or more coherent “role” identities. The normative content of an identity derives from a broader set of social norms that emanate from multiple centers of authority. Norms can thus be unwritten or codified—in other words, social or legal—so long as they appear to fix meanings and set collective expectations for members of the group. These practices cause group-recognition and are thus, necessarily, obligations of individual members of the group. When practices that lead to recognition are also understood as obligations, they may be valorized by the group as ethical.

Constitutive norms do more than identify the “proper” or “appropriate” behavior for a particular identity, though such a regulatory effect is important, as in role theory. The effect that is even more powerful is recognition—constitutive norms are the very actions that lead others to recognize an actor as having a particular identity. They are also distinct from social purposes (i.e., shared interests or preferences). Rather than specifying the ends of action, norms help to define social meaning by establishing collective expectations and individual obligations. Thus, constitutive norms do not determine the preferences of a group; rather, they define the boundaries and distinctive practices of a group.

Illustrations of constitutive norms drawn from political science scholarship suggest that practices that lead relevant others to recognize an identity can be either conscious or taken for granted. The degree to which such practices are habituated or internalized (that is, the degree to which individual members are socialized) is an empirical question. Regardless of the degree to which constitutive practices are unconscious, unquestioned, or taken for granted, such norms are integral parts of the social meaning of an identity. As Price and Tannenwald have argued, over the course of the late twentieth century the content of the identity “civilized state” evolved a great deal. By the end of the century, “civilized states” did not employ nuclear or chemical weapons in their armed struggles. Klotz traced the emergence of the norm of racial equality in international society. In security communities, the factors that prevent defection are not institutional per se, but are based on the development of shared notions of in-group identification where interaction has literally eliminated defection (war) as a possibility—where there exists the “impossibility of imagining violence.” In this conceptualization, the reasons to act in a particular way are found in a decision to perform a role, not in a decision to choose between optimizing paths to some preferred outcome.

Much of the scholarly literature on socialization is also implicitly about the normative content of identities. What is at stake in socialization is ultimately the internalization of constitutive norms—the process by which the collective expectations of the members of an identity group come to feel taken for granted by new members. Checkel’s research on Europe’s constitutive norms for citizenship policies falls into this category. Indeed, the European Union’s own Copenhagen Criteria for determining the acceptability of potential members—in a nutshell, a market economy, a democratic polity, and respect for human rights—represent an explicit assessment of the constitutive norms that define European-ness for current and potential group members.

The process by which constitutive practices are internalized or habituated, may be manifested in three ways. First, norms may bias choice, meaning that certain behaviors are consciously ruled out or discounted as inappropriate for one’s identity. The commonly used phrase, “logic of appropriateness” might best describe this level of internalization. Second, norms may reduce the level of consciousness in choice. Semi-conscious choice would mean options are barely considered, or only fleetingly considered, and are dismissed out of hand. “Common sensible” choice might capture this form of internalization. Third, norms may be so deeply internalized that they are acted upon completely unconsciously, out of habit. As Fierke has written, drawing on Wittgenstein “rules are lived rather than consciously applied.”
as the logic of habituation. Options are simply not considered. Practices are just followed.14

Social purposes
The content of a collective identity may be purposive, in the sense that the group attaches specific goals to its identity. This purposive content is analytically similar to the common sense notion that what groups want depends on who they think they are. Thus, identities can lead actors to endow practices with group purposes and to interpret the world through lenses defined in part by those purposes.15 Whereas the normative content of an identity refers to practices that lead to individual obligation and social recognition, the purposive content of an identity helps to define group interests, goals, or preferences. Both the normative and purposive content of an identity may impose obligations on members, but in distinctive ways: constitutive norms impose an obligation to engage in practices that reconstitute the group, while social purposes create obligations to engage in practices that make the group’s achievement of a set of goals more likely.

The notion of the purposive content of identity is already implicit in the literature, although it has seldom been expressed this way before. The construction that pervades identity scholarship—who we are influences what we want—specifies a shared purpose.16 Horowitz similarly has written of the “special missions” ascribed by some members to their groups. Smith’s theorizing of economic, political, and ethically constitutive “stories of peoplehood” also can be understood in these terms, creating the basis for narratives of purpose. Kelman has produced important research on the connection between specific territorial claims and national identities. Yashar explored the purposive claims to a more equitable form of citizenship of indigenous movements in Latin America. Reus-Smit found moral purpose to be central to the history of modern statehood and concomitant claims to making citizens and creating justice.17 For Reus-Smit, societies of states, from ancient Greece, to Renaissance Italy, to absolutist Europe, and finally to our modern international system, have been based on fundamental moral purposes that have varied a great deal. These examples cover a wide variety of empirical questions but have in common an emphasis on the purposive meaning derived from an identity.

The scholarly literature on nationalist movements and national interests and identity as well.19

Relational comparisons
The content of a collective identity is also relational to the extent that it is composed of comparisons and references to other collective identities from which it is distinguished. An identity may be defined by what it is not, i.e., by some other identities. The relational content of collective identities can be thought of as the discursive formulations of the relations between groups of people that compose social reality. Barnett provides an excellent relational definition of identity in his work on the Middle East peace process. He has written that identity represents “the understanding of oneself in relationship to others. Group identities, in short, are not personal or psychological, they are fundamentally social and relational, defined by the actor’s interaction with and relationship to others; therefore, identities may be contingent, dependent on the actor’s interaction with others and place within an institutional context.”20

Scholars have already identified a number of relational characteristics of collective identities, including, among others, the extent to which one social identity excludes the holding of another (exclusivity); the relative status of an identity compared to others; and the existence or level of hostility presented by other identities. Examples of the implicit relational theorizing that forms the basis for much identity research are rife in all areas of the social sciences. Klandermans explored the foundational distinction between social movements and the “authorities.” Neumann described the importance of a constituting “other” for the creation of European identity. For Bartelson, the relational content of an identity is, following Derrida, more revealing than other putatively self-referential narratives. Bell described the process of constructing French nationalism with England’s barbarianism as a focal point against which to define France. Bailey examined four cities including San Francisco and Birmingham, Alabama in order to...
explore relational aspects of gay identity. Thomas Risse undertook case studies of the 1956 Suez Crisis and the 1962 Cuban Missile Crisis in examining the NATO alliance in his work on shared liberal values and democratic norms as the basis of collective democratic state identities.21 Our analytic framework makes more explicit the relational component already implicit throughout the literature and seeks to standardize it as one of the four types of identity.

Relational content is also crucial for social identity theory (SIT), which hypothesizes that the creation of in-group identity will tend to produce competitive behavior with out-groups, because the process of in-group identity creation by necessity requires, or leads to, the devaluation of out-groups. In social identity theory, the central causal process in behavior derives from in-group and out-group differentiation, not the roles or identity traits per se that are attributed to in-groups and out-groups.22 In this case, action is in some sense a reaction to, and conditioned by the existence of, those who are different. Some relationships (those with groups socially recognized as similar) will be more cooperative than others (those with groups recognized as different) even if the same issue is at stake (such as territory, power, or status). Under certain scope conditions, SIT-based arguments predict conflict with out-groups socially recognized as similar) will be more cooperative than others (those with groups recognized as different) even if the same issue is at stake (such as territory, power, or status). Under certain scope conditions, SIT-based arguments predict conflict with out-groups regardless of the content of the identity—we, we are peace-loving, but you are not, and because of this difference you threaten our peace-lovingness; therefore anything goes in dealing with your disposition to threaten us.

Cognitive models

There are many ways to think about the cognitive content of social identities, as we find throughout the literature. In the broadest sense, a cognitive model may be thought of as a worldview, or a framework that allows members of a group to make sense of social, political, and economic conditions.23 The cognitive content of a collective identity describes how group membership is associated with explanations of how the world works as well as descriptions of the social reality of the group—a group's ontology and epistemology. Being French, for example, may entail a particular way of interpreting the world.24 For some scholars, the "cognitive turn" in the study of identities is critically important. According to Brubaker, Loveman, and Stamatov, what cognitive perspectives suggest, in short, is that race, ethnicity, and nation are not things in the world but ways of seeing the world. They are ways of understanding and identifying oneself, making sense of one's problems and predicaments, identifying one's interests, and orienting one's action. They are ways of recognizing, identifying, and classifying other people, of construing sameness and difference, and of "coding" and making sense of their actions.25

Similarly, a cognitive model may consist of "ways of reasoning" that are specific to particular identity groups. This was Peng and Nisbett's finding in a survey of Chinese and Americans. This result sparked further debate over the ways in which Eastern versus Western identity affects ways of thinking.26 In addition, Gurung argued that different cultures (Chinese and Western) have different understandings of what constitutes "knowledge."27 In an analysis of the 1893 celebration of the Silver Jubilee of Shanghai's Municipal Council of International Settlement, Goodman argued that different communities taking part in the jubilee (English and Chinese) had different readings of the festivities and indeed "different mental universes."28

Identities can strongly affect interpretation and understanding not just of the present but of the past as well. Smith argued that "subjective perception and understanding of the communal past by each generation . . . is a defining element in the concept of cultural identity." More particularly, Azzam found that Muslim identity may shape memories of shared colonial experience. Identities may also shape conceptions of the future; psychologists have found that cultural differences are associated with different views on what constitutes "the good life" or "well-being."29

Cognitive models affect not only broad worldviews and temporalities but also understandings of self, group, and other. One of the primary ways that this happens is through language. Mar-Molinero has argued that "language is a means by which human beings grow to understand themselves and then to understand and share with those who speak the same language." Causal attribution is another cognitive activity affected by identity. Klandermans has argued that identity "not only emphasizes the commonality of grievances, it also establishes the group's opposition to the actor held responsible . . . thus causal attributions are an important element in the identity component . . . this element is related to the construction of a cognitive schema which comprises causes and solutions for the adverse situation."30

The literature has also produced many cases demonstrating how identities can affect understandings of political and economic interests. In the realm of politics, identities can affect conceptions of legitimacy, shared interests, and policy choices, as well as preferences for political leaders and parties. Kelman has argued that national identity allows members of a group to see their state as legitimate; Feng demonstrated that Hainanese groups have different perceptions of political and economic conditions on the island; Shabad and Slomczynski found that identity shaped "orientation" towards transition issues in Poland; Adler explored how being part of a transnational identity group (international arms control specialists) shaped understandings of security concerns; and Valenzuela and Scully showed that voters from different classes had different values and therefore assessed political leaders and parties differently. Connecting class and gender in the workplace, Canning argued that identity discourses...
Article | Identity as a Variable

shape the understanding and meaning of work, which differ by gender.31

Identities such as ethnicity and region can also shape interpretations of material conditions and economic interests. Risse et al. have argued that “collective identities define and shape how actors view their perceived instrumental and material interests and which preferences are regarded as legitimate and appropriate for enacting given identities.” 32 Herrera has explored how regional identities within Russia have led to different views of regional economic conditions. There is also evidence that identities affect understandings of land itself, including understandings of access and rights. Identity may indeed shape perceptions of territory, which also shape perceptions of culture. Goodman has argued that provincial identity in Shanxi, China was shaped in concert with a specific “interpretation of the area’s centrality to the development of Chinese culture.”

Cognitive content, rather than implying an alternative theory of action, implies a theory of interpretation. The attention to the cognitive shows us both how identity affects how actors understand the world, and, consequently, how their material or social incentives for particular actions will be influenced by their identities.

We believe that our analytic framework’s four non-exclusive content types, plus the element of contestation discussed below, preserve the restless dynamism that characterizes current identity work. At the same time, by providing more analytical coherence, we hope to move beyond the current crisis of definitional anarchy and towards collaboration and integration of identity scholarship across the social sciences.

Contestation

The content—the collective meaning—of identities is neither fixed nor predetermined. Rather, content is the outcome of a process of social contestation within the group. Indeed, much of identity discourse is the working out of the meaning of a particular collective identity through the contestation of its members. Individuals are continuously proposing and shaping the meanings of the groups to which they belong.

Specific interpretations of the meaning of an identity are sometimes widely shared among members of a group and sometimes less widely shared. At a minimum, then, contestation can be thought of as a matter of degree—the content of collective identities can be more or less contested. Indeed, the further apart the contending interpretations of a collective identity prove to be, the more that identity will be fragmented into conflicting and potentially inconsistent understandings of what the group’s purposes or relations should be. Such disputes, occurring within one or more of the four types of content, might be prompted by differences over how exclusive the member-
compose a group often seek the recognition of their identity by others, and that recognition—both formal and informal—clearly influences the particular goals associated with an identity.

The range of behaviors and practices that compose the process of contestation is broad, and, as with other aspects of identity formation and reformation, the study of contestation is most usefully understood as an empirical matter. The potential texts to which scholars have access in measuring and evaluating the process of contestation are many. More importantly, the process of contestation can be either explicit (and therefore intentional) or implicit (and therefore unplanned).

Explicit debates about the meaning of an identity tend to be self-referential. The controversy that followed the publication of Samuel Huntington’s book *Who Are We?: The Challenges to America’s National Identity* is rather straightforward: in his book Huntington literally seeks to define, historically, the content of American national identity, and his critics responded with alternate histories and alternative contents.55 The essence of the debate is clear, however—American intellectuals were contesting the meaning of their national identity in books; in book reviews from the pages of *Foreign Affairs, Foreign Policy,* and the *New York Times,* and in articles written for other intellectuals, policy makers, and the educated lay reader. The experience is not unusual in comparative context, as similar explicit debates have emerged in every country.

Even more common is the everyday, implicit contestation of identity, which takes place among members of a group without their consciously seeking to revise or remake the meaning of their identity. If it is a small enough group, this process may take place in conversation or bilateral written communication. In modern societies, journalists and the media play an important role in constructing meanings. Then there are the ongoing claims and counter-claims offered by those who aspire to lead a group, regardless of its size. Thus political debates, party platforms, and speeches are designed to evoke a sense of collective self and are examples of the process of contestation.

Because the content of an identity is the product of contestation, the very data that a scholar extracts from a group elucidate, in manner and degree, the members’ consensus and disagreement about the constitutive norms, social purposes, relational comparisons, and cognitive models of their collective identity. Contestation therefore includes the degree of within-group agreement about the constitutive norms of an identity; consensus and congruence of the social purposes ascribed to an identity; agreement about meanings attached to out-groups; and coherence of shared cognitive models.

By considering the level of contestation regarding each type of content within identities, one arrives at a necessarily constructivist approach to identity without having to assume that actors on the ground view their identities as constructed. Where there is little contestation, one might conclude that that part of identity content is taken for granted or considered “natural.” Thus, one can appreciate some apparently “primordial” aspects of identity without taking a primordialist theoretical stance that denies the possibility for contextual contestation, that is, contestation at different times and places.

**Some conceptual advantages of our framework**

Finally, before moving on to the methodologies of identity work, let us consider what our framework does not do. First, we are not attempting to account for personal identities or identities of individual persons. Individuals and groups remain analytically distinct objects, each requiring its own conceptualization.36

The essential difference lies in the *collective* meaning inherent in social identities. Whereas individual identities are subjective, collective identities are “intersubjective”—comprised of shared interpretations of group traits or attributes. Individuals may, and almost always do, ascribe different meanings and purposes to the same collective identity; that is, the meaning of a particular collective identity is defined through a process of contestation among individuals who essentially propose alternate collective meanings. In that sense, understanding the interaction among constituent individuals and their groups—or agents and identity structures—is a crucial part of the analysis of social identities. But it is as members of a group that the practices of individuals are most easily analyzed with social scientific methods. Attempting to impute the identity of the individual *qua* individual is, in other words, a matter more for psychology, whereas social psychology attempts to describe the connections among individual and collective identities.37

For this reason, *salience* and *intensity,* two terms common in the identity literature, are recast more precisely within our framework for social identities using the concepts of content and contestation. Salience is a critical variable used by psychologists to study the multiple and overlapping identities of individuals as individuals; specific contexts are said to increase the salience of one identity over another. Our framework focuses on the *meaning* of collective identities, a property that salience cannot address. We would further suggest that the sort of variation described by salience is less common than variation in meaning (types of content) and its contestation.

The issue of salience is also related to “groupness,” along the lines proposed by Brubaker and Cooper.38 Groupness describes the degree to which individuals identify with a group (or, in other words, how salient that particular collective identity is to its constituent individuals). Here, again, we find that focusing on the level of attachment to a group bypasses the *meaning* of the group to its members. Identities
cannot exist without meanings ascribed to them by their members, and we argue that the degree to which members of a group share those meanings is a variable captured by the concept of contestation.

Intensity, like salience, often appears in the literature with regard to individuals, in this case with the idea that some constituent members of groups attach greater importance to a specific collective identity than do others. We argue that, like salience, intensity cannot comprehend the degree to which group understandings are shared. Additionally, the notion of intensity often conflates two analytically distinct variations in identity. The international relations literature on nationalism, for example, regularly includes references to “hyper-nationalism.” Scholars generally use this term to describe, without analytical distinction between the two, both the extremeness of the views of a nationalist society, as well as the level of agreement among members of that society about such extreme views. Our analytic framework enables more precise distinctions. In our terms, the first idea concerns the content of an identity, while the second is a question of its contestation. Thus, we argue that content and contestation are better ways of getting at the multiple concepts covered by the term intensity.

Methods for Measuring the Content and Contestation of Identity

Which methodologies are best suited to measuring the content and contestation of identity? The literature has so far relied primarily on discourse analysis, surveys, and content analysis, all of which have been used to measure the normative, purposive, relational, and cognitive content of an identity, as well as contestation of content. We did not discover any systematic links between these methods and the types of content they were used to measure, although nearly all studies of identity included some sort of case study.

There are also three less common methods—experiments, agent-based modeling, and cognitive mapping—that we feel offer great promise to supplement the dominant methods, although the research agendas for their widespread incorporation into identity research are still being formulated. We advocate these six methodologies as the best suited for identity research for the reasons outlined below, and we conclude by offering our reflections on the path ahead for identity research at this exciting moment in the field. We hope that this methodological roadmap might facilitate further integration of identity work across the disciplines and subfields.

The state of the art in identity research combines these methods, with great sophistication, to create the most comprehensive analysis of identity possible. Cynthia Kaplan and Henry Brady employed surveys, discourse analysis, and content analysis in their study of seven groups: Estonians and Russians in Estonia, Tatars and Russians in the Republic of Tatarstan, the Komi and Russians in the Komi Republic, and Russians in Russia. Paul Sniderman and his colleagues combined three surveys with experiments designed to recover the meaning of the threat putatively posed by immigration into the Netherlands. They thereby gained insight into the relational (vis-à-vis immigrants) and cognitive (the influence and threat of immigrants on economic well-being, coherence of cultural identity, and community safety) content of Dutch national identity. And Donald Sylvan and his colleagues have used surveys, discourse analysis of interviews, content analysis, and experiments in their study of the relationship between Palestinian and Israeli identity. Unfortunately, there are very few examples so far of this sophisticated blending of methods, and much work remains to be done to demonstrate the variety of insights into identity’s content and contestation that can be gleaned from different methods.

Discourse analysis

For our purposes, discourse analysis is the qualitative and interpretive recovery of meaning from the language that actors use to describe and understand social phenomena. The very notion of discourse is usually understood as a collection of related texts, constituted as speech, written documents, and social practices, that produce meaning and organize social knowledge. The aim is to analyze such texts in order to discern or interpret the intersubjective context of the speech of actors. Scholars have relied on structured and semi-structured interviews as well as their own informed interpretations of a variety of texts, including policy statements and white papers, political party platforms, newspaper articles, classic texts of prominent public intellectuals, speeches of political leaders, and the minutes of government meetings at which important policy decisions were taken.

Unlike statistics, programming, or modeling, discourse analysis requires deep social knowledge, interpretive skills, and a familiarity with a body of interrelated texts in order for scholars to recover meanings from a discourse. The critical task for scholars is to reconstruct the intersubjective context of some social phenomenon—in our case, a collective identity—in order to account for an empirical outcome. Discourse analysis thus can be considered the qualitative contextualization of texts and practices in order to describe social meanings.

Discourse analysis has been particularly useful for studying the relational content of identity because the one, perhaps universal, distinguishing feature of an in-group is shared modes of communication, in other words, shared linguistic practices. These not only allow communication—hence, polarization in the self-categorization theory sense—but they are also markers themselves of group membership. Meta-discourse analysis, focusing on the uses of linguistic hedges,
emphatics, and attitude markers, has also been used to observe levels of affect towards self and other.45

**Surveys**

Surveys, long the backbone of research in public opinion and political behavior, are useful for identity research as well.46 Structured interviews of large numbers of people allow individuals and groups to offer and establish their own self-definings.47 Open-ended survey and interview instruments in particular allow respondents to present their own notions of who they are, what matters to them and why, and how they decide who is or is not included in their group, without being pigeonholed by the researcher’s prior biases and interpretations. Surveys and interviews have the obvious advantage of allowing researchers to ask specific questions about identity. These techniques allow interested researchers to directly address questions of content and contestation.

Long before identity research per se came into vogue, surveys were already implicitly producing the kinds of data that would later characterize the field. Surveys are fairly straightforward in the way they tap into the content of identities. Their questions often inquire directly into self-described attributes, attitudes, and practices that respondents believe that they should express as a member of X social group. Early work on party identification done by the Social Survey Center at the University of Michigan, for instance, simply asked respondents how important it was to them to think of themselves as a member of their political group. Early work on party identification done by the Social Survey Center at the University of Michigan, for instance, simply asked respondents how important it was to them to think of themselves as members of a political group. This type of question was adapted to the analysis of social identity.48 Other surveys ask equally straightforward questions. The World Values Survey asks about how proud respondents are to have an identity as X (nationality).49 The use of surveys to analyze the content and contestation of identities has since been well established. In measuring the degree to which Dutch respondents see Dutchness as important to their identities, Sniderman et al. asked very direct questions about the degree of pride in being Dutch and the degree to which respondents personalized criticisms of the Dutch.50

Survey data have proven particularly useful in exploring relational aspects of identity. David Laitin’s work on culture and identity in the European Union used surveys conducted in the six original EU countries, as well as subsequent and applicant EU members. Asking questions specifically focused on language and religious beliefs, he explored the relationship between these factors and prospects for European integration among countries in Eastern Europe. Michael Barnett, on the other hand, examined extant public opinion polls to investigate the relationship between Israeli identity and the Oslo peace accords.51

Semantic differential protocols, by providing information about a group’s prototypical traits and characteristics, offer another way that surveys can explore the relational content of identities.52 Through a process of induction (often requiring participant observation), the researcher develops a list of the typical adjectives used by a social group to describe self and other. These are then included in a survey instrument that asks respondents to place self and other along a 5-, 7-, or 9-point scale between an adjective and its polar opposite (strong/weak; peaceful/warlike; masculine/feminine; civilized/barbaric, etc.). This information can then be used in role-theoretical analysis to identify appropriate behaviors associated with particular phrases, such as “Ours is a country that prizes free markets,” or “Ours is a country that abhors non-democratic systems.” These descriptions can then be mapped onto normative preferences.

Two additional aspects of relational content can be discerned from this kind of survey. First, the mean and dispersion of responses between groups can be compared to indicate the degree of shared stereotypes about self and other within and between groups (e.g., do Americans agree about what traits Americans have and what traits Chinese have, and do they agree with Chinese about what traits Americans and Chinese have?). This information allows one to construct “meta-contrast ratios.” The ratio (whether calculated qualitatively or quantitatively) is one of the central indicators in social identity theory of the degree of within-group and between-group polarization. The larger the ratio—that is, the more extreme the means and the tighter the dispersions around the means for the two groups—the more likely the groups will, ceteris paribus, see each other in stereotyped and competitive terms. Tracking such movement in means and dispersions between two or more groups may be a useful early warning of growing inter-group competition as well as growing intra-group repression, exclusion, or policing.

Second, the mean and dispersion will also indicate where the boundaries of the in-group lie. That is, it will identify traits that are considered beyond the pale. Extreme means with tight dispersion, for instance, indicate clearly what kinds of traits are considered out-group-like. This information is critical in predicting how the in-group will police itself and punish those with liminal identities. Burke and Tully found in their study of girls’ and boys’ stereotypes of girlness and boyness that those who were viewed as possessing traits that lay in between stereo-typical girl or stereo-typical boy traits were more likely to be teased and harassed.53 The more dispersion there is around the mean, of course, the wider will be the range of traits that will be considered acceptable to the in-group. For all these reasons, we find semantic differential analysis particularly useful for tapping into the relational and constitutive content of social categories.

**Content analysis**

Content analysis is, according to Neuendorf, “a summarizing, quantitative analysis of messages that relies on the
Article | Identity as a Variable

scientific method. The unit of analysis is the “message component,” typically summarized quantitatively. Whereas discourse analysis relies on the interpretive skills of the scholar to mediate between the reader and the text or practice under study, content analysis is designed to limit mediation. The critical step is the creation of a coding scheme that is written out in great detail in order to ensure reliability among coders of the text. Where discourse analysis would interpret texts and practices in their inter-subjective contexts, content analysis treats the individual texts as meaningful on their own and produces quantitative content summaries thereof.

Content analysis has not yet been used as widely for the measurement of identities as discourse analysis and surveys, but its usefulness has been amply demonstrated. Herrera used content analysis of regional newspaper articles to examine local understandings of regional economic conditions in Russia. The results helped explain the relationship between such shared regional interpretations of the economy and support for regional sovereignty movements. Elders and Lüter employed quantitative content analysis in examining editorials on the Kosovo war in five German newspapers across the political spectrum. They uncovered not only fundamental purposive content that modern Germans generally ascribe to their identity—commitments to being a “loyal NATO member,” a “part of Europe,” and, because of the Nazi experience, an outspoken defender of human rights—but also the variation in the emphases placed on those purposes by the Left, Center, and Right. David Laitin used content analysis to examine the terms used to describe Russian speakers in a range of newspapers in the former Soviet states in his analysis of the formation of a Russian-speaking identity group.

The two basic raw results of content analysis are frequency counts of both key words and categories of terms. The latter allows the virtually unlimited categorization of textual material. Words can be clustered along shared dimensions, and the categories themselves can be analyzed in terms of relative frequency. The categories can be tailored to each project’s individual research needs, as the researcher is in charge of building his or her own dictionaries. Content analysis can also provide concordances, also known as key-word-in-context (KWIC) analyses. A KWIC analysis will list all instances where a particular term is mentioned in a text. Such data can be very useful for discerning the semantic, grammatical, or substantive qualities of identity language. Software for performing quantitative analysis has improved substantially in recent years, and is increasingly available in languages other than English.

Experiments
Experiments, the first of the three less traditional methodologies in our survey, offer the advantage of unparalleled control and assessment of causality in the measurement of identity. Although their use has been mostly confined to psychology and social identity theory, their use in identity research is increasing.

Experiments are meant to answer particular research questions. Each experiment begins by designing a particular protocol which randomly assigns subjects to various conditions. Each condition typically manipulates one or a very few variables of central interest. This design feature allows the experimenter to determine the cause of any observed changes in outcome among individuals or groups. Random assignment ensures that any emerging differences derive from the experimental manipulation and are not merely the consequence of pre-existing or systematic divergences between individuals or groups. In this way, experimental procedure allows true leverage in making causal arguments.

Despite the relative rarity of their use in political science, experiments have so far proven to be an excellent choice for capturing the internalization of normative elements of an identity. Though they called what they captured a “culture” rather than an “identity,” Nisbett and Cohen used an creative combination of lab and field experiments in their work on the role of honor in Southern white male identity. Their experiments—such as asking newspaper writers from different cultural backgrounds to tell a story using the same facts or to editorialize about the same story—could be easily adopted by other identity researchers. There is surely much to be learned by discovering how different social groups “editorialize” about appropriate behavior given descriptions of their own in-group and different outgroups. We can also imagine an experiment that might ask what would happen if the norms of an experimental group were changed exogenously, by the experimenter, either by changing the rules for payoff or by introducing powerful new members to the group who espouse different role conceptions.

Social psychology, unlike political science, possesses a long tradition of experimental investigation into various aspects of personal and collective identity. Indeed, most work on SIT within social psychology rests on experimental evidence. Henri Tajfel and Michael Billig conducted the original experiments which resulted in the development of SIT. These experiments included one that divided British adolescents into two groups, each of which was asked to evaluate some modern art. Students were then told that they and others had preferred Klee’s art to Kandinsky’s. Without ever meeting the other members of their “group,” subjects proceeded to divide points worth money between their group and the other. Consistently, subjects allocated more points to their own group than the other at a ratio of about 2:1. Interestingly, this bias appears particularly pronounced in individualistic cultures.
Agent-based modeling

Agent-based modeling is a computational methodology with a relatively short history but a great deal of promise. Whereas experiments offer scholars the opportunity to manipulate the social institutions of, at times, artificial societies of real individuals, agent-based modeling deals with artificial societies of simulated individuals. According to Lars-Erik Cederman’s recent review, agent-based modeling “allows the analyst to create, analyze, and experiment with, artificial worlds populated by agents that interact in non-trivial ways and that constitute their own environment.” Computation, in this method, is used to “simulate agents’ cognitive processes and behavior.”

Although Cederman and several other scholars have demonstrated the usefulness of agent-based modeling in the study of identity, it is not yet widely used. The most significant advantage of agent-based modeling is the opportunity it affords scholars to explore the microfoundations of constructivist theories of identity formation, reproduction, evolution, and even transformation. It is not yet clear, however, that agent-based modeling can be understood as a methodology for measuring the content and contestation of a specific identity.

Cognitive mapping

Cognitive mapping breaks down selected texts from a decision-making process into all of their component cause-effect relationships. The researcher then determines whether those causal relationships are negative or positive—that is, whether a change in direction in the causal concept leads to a similar or dissimilar change in the effect concept (e.g., does an increase in arms lead to an increase or decrease in security?). The technique can help uncover the deep structure of an argument—the presence or absence of certain cause-effect assumptions, and the consistencies and inconsistencies across cause-effect arguments. These maps can be compared across actors within an identity group, or aggregated within the group and compared with the maps of out-groups, to determine what cause-effect relationships are shared or not shared by actors and the degree to which they are shared.

Adapted for identity research, cognitive mapping could be one way of observing whether a group regards certain roles or behaviors as appropriate, and whether others are ruled out of bounds or simply not even considered. Internalization might be shown as a cognitive map in which identity-consistent practices are believed to achieve positive effects for group utility, while plausible or possible alternative actions are simply not found in the map. To determine whether, say, a security community exists, one needs to show that there is an “impossibility of violence” imagined in a relationship with another actor. One could thus look at cognitive maps of a decision-maker or a group of decision-makers to see whether violence was or was not imagined, or whether it did or did not appear as a cause concept negatively linked to some effect (as it would have to be if it had been discarded as a too-costly option). Then one would have to look at a comparative case of dispute with similar content and intensity but with an out-group actor. If “violence” was imagined in the second cognitive map, if it was a conscious cause concept leading to a particular effect, only then could one conclude that in the first dispute violence was indeed “unimagined,” thereby proving the security community’s existence.

Conclusions

Despite the proliferation of identity research in recent years, the social science community has yet to provide the analytical rigor that would render the concept of identity usable as a measurable variable across the disciplines and subfields. We have offered in this article a new analytic framework that we believe can move identity research beyond its current impasse, while preserving the dynamism that has characterized the research so far.

We defined collective identity as a social category that varies along two dimensions—content and contestation. We delineated identity content into four, non-mutually-exclusive types: constitutive norms; social purposes; relational comparisons with other social categories; and cognitive models. And we argued that there is always some level of in-group contestation over this content, implying that social identities vary in agreement and disagreement about their norms, boundaries, worldviews, analytics, and meanings.

We believe that this framework can account for the variation in identities that scholars have already been implicitly and explicitly measuring. We hope that our framework will provide greater theoretical commensurability among conceptions of identity in political science and the other social sciences, while still recognizing and valorizing a diversity of approaches. Our goal is not to “discipline” identity, or to impose a new, narrow semantic strait-jacket on scholars who seek to treat identity as a variable. Rather, by categorizing identity scholarship and its methods in a synthetic framework and highlighting complementarities among conceptualizations and methods, we aim to encourage more coordination and explicit comparison among scholars working on identity.

In outlining a definition of identity and describing methodologies for measurement, we also aim to encourage creativity in thinking about identity. We think any definition
of identity must address the issues of content and contestation, and we believe that there are a variety of research methods amenable to identity research. Rather than closing off any approaches, we hope that, by outlining specific options, we have instead conveyed a sense of the wide range of methodological tools open to scholars interested in identity research. We think a commitment to empirical research requires both attention to rigorous methods, as well as the open-mindedness that allows researchers to adapt to the specific demands of a particular research question. Attention to measurement helps complete the story of how a social identity came to be what it was at a particular historical moment, as well as how it might be changing, or in a process of re-formation.

Our framework also helps to clarify differences among types and casual consequences of identities. We also hope that our conceptualization and the discussion of measurement methods have made a useful contribution to understanding how identities affect the behavior of actors—in other words, the relationship between identity and action. We argue that the definition of identity, i.e., content and contestation, can affect predictions about action. Our definition, and in particular the four types of content that we have outlined, specifically addresses theories of action, such as SIT or role theory, as well as models of cognition and the purposive goals of a social group.

A longer-lasting contribution of this work may be our drawing explicit connections between alternative conceptualizations of the variation in identities and the methods available to measure them. The result of greater attention to common conceptualizations and research methods will support more rigorous and replicable studies of identity, help scholars to understand the wide range of tools available for analyzing identity, and overcome some of the reluctance of mainstream political science to incorporate identity variables into explanatory models.

Notes
1 Smith 2004, 302.
2 Brubaker and Cooper 2000.
3 Fearon 1999 divides personal from social identities, and then further divides social identities into type and role identities. Brewer and Gardner 1996 distinguish three types of identity: personal, collective, and relational. Our framework considers the relational identity to be one of four types of content of collective identities. Chandra and Laitin 2002 present a general classification scheme for identities. They refer to categories, attributes, and dimensions as the three main components of an identity. Categories are the immediate term used to describe an individual’s identity (e.g. working class). Attributes are the qualities individuals are expected to express to meet membership criteria for a social category. Dimensions are the range of categories that make up a typology (working class, middle class, aristocracy together constitute the typology “class”).
4 Monroe, Hankin, and van Vechten 2000.
5 Katzenstein 1996; Ruggie 1998a.
6 Hopf 2002.
7 Price and Tannenwald 1996.
9 On security communities, see Adler and Barnett 1999. On state identities and state practices in general, including national security policies, see Jepperson, Wendt, and Katzenstein 1996.
12 Weldes et al. 1999.
14 Hopf 2002.
15 Analogously, see Goldstein and Keohane 1993 on “principled beliefs.”
16 See Gutmann 2003.
21 Klandermans 1997; Neumann 1999; Bartelson 1998; Bell 2001; Bailey 2002; Risse-Kappen 1996.
23 Some use other terms besides “worldviews.” Denzau and North 1994, for example, use “shared mental models.”
28 Goodman 2000, 921.
31 Kelman 1969; Feng 1999; Shabad and Slomczynski 1999; Adler 1992; Valenzuela and Scully 1997; Canning 1996.
33 Cox 1986.
36 Brubaker and Cooper 2000 move back and forth between personal and social identities in their discussion of alternatives to identity. However, it seems that they are dissatisfied primarily with the use of
individuals' identities to explain individuals' behaviors and practices. Had they separated their analysis between individual and social identities, they might not have been as pessimistic regarding prospects for further research on social identities.

38 Ibid.
39 Mearsheimer 1990.
40 Brady and Kaplan 2000.
42 Sylvan 2004.
43 For guidance on conducting discourse analysis, see Phillips and Hardy 2002; Wodak and Meyer 2001; chapters 9–12 in Bauer and Gaskell 2000; and Titscher et al. 2000.
45 Abdi 2002.
46 See, for example, Citrin et al. 1990; and Citrin et al. 1994.
47 For guidance on surveys and interviews related to identity research, see chapters 3–5 in Bauer and Gaskell 2000. On interviews, see Briggs 1986; McCracken 1988; Rubin and Rubin 1995; and Holstein and Gubrium 2003. On surveys more generally, see Fink 2003.
48 Hooper 1976.
49 For an application of these questions in developing measures of nationalism and patriotism, see Furia 2002. See also the measures for national identity and pride in Smith and Jarkko 2001; their paper draws on a 1995 cross-national survey on national identity conducted by the International Social Survey Programme (http://www.issp.org/natpride.doc).
50 Sniderman et al. 2004, 48.
51 Laitin 2002; Barnett 1999.
52 Burke and Tully 1977; Osgood, Tannenbaum, and Suci 1990.
53 Burke and Tully 1977.
54 Neuendorf 2004. For the discussion of the difference between content analysis and discourse analysis, see Herrera and Braumoeller 2004 and Fierke 2004.
55 Recently there have been advances in computer-aided content analysis, but most research is still based on human coding.
58 For guidance on conducting quantitative content analysis, see Neuendorf 2002; chapters 8, 16–17 in Bauer and Gaskell 2000; Titscher et al. 2000; Popping 2000. For further discussion of the advantages and disadvantages of different content analysis software for the analysis of identity, see Lowe 2002, 2004.
59 For example, the Yoshikoder, a cross-platform multilingual content analysis program developed by Will Lowe for the Harvard Identity Project allows users to attribute quantitative values to different descriptors in the dictionary. The user can then choose, for instance, a particular identity category (say, an ethnic group) and determine the relative valence attributed to that category by the text. The Yoshikoder works with text documents, whether in plain ASCII, Unicode (e.g., UTF-8), or a national encodings (e.g., Big5 Chinese). It is open-source software, released under the Gnu Public License, meaning free for academic use and available at http://www.yoshikoder.org.
60 Nisbett and Cohen 1996.
63 Axelrod 1997a and 1997b.
64 Cederman 2001, 16.
66 Axelrod 1976.

References
Article | Identity as a Variable


