Foundations of a Theory of Social Forms

Lászlo Pólos, Michael T. Hannan, Glenn R. Carroll

ERIM REPORT SERIES RESEARCH IN MANAGEMENT			
ERIM Report Series reference number	ERS-2000-29-ORG		
Publication status / version	draft / version March 2000		
Number of pages	27		
Email address first author	Lpolos@fbk.eur.nl		
Address	Erasmus F	Erasmus Research Institute of Management (ERIM)	
	Rotterdam	School of Management / Faculteit Bedrijfskunde	
	Erasmus Universiteit Rotterdam PoBox 1738		
	3000 DR Rotterdam, The Netherlands		
	Phone:	# 31-(0) 10-408 1182	
	Fax:	# 31-(0) 10-408 9020	
	Email:	info@erim.eur.nl	
	Internet:	www.erim.eur.nl	

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website: www.erim.eur.nl

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS			
Abstract	Sociologists frequently invoke the concept of form when analyzing organizations, collective action, art, music, culture and other phenomena. Nonetheless, the form concept has not received careful theoretical analysis, either generally or in specific context. Using the tools of formal logic, set theory, and algebra, we propose a language for defining social forms that is suÆciently general to incorporate feature-based, position-based, and boundary-based approaches to defining forms. We focus on organizational forms although we intend for our conceptualization to be general. We define forms as a type of socially coded identity. We de_ne identity in terms of social codes that specify the properties that an entity can legitimately possess. These codes can be enforced by insiders or outsiders. We claim that one knows that a social code exists when one observes that departures from the codes after periods of conformity cause a devaluation of the entity by relevant insiders and/or outsiders. This construction allows us to de_ne a population as the set of entities with a common minimal external identity in a bounded system in period. The minimal property ensures that we localize to the most speci_c socially enforced identities. The reliance on identities instead of forms allows us to de_ne populations that never achieve form status and to extend population de_nitions back to the period of early legitimation. Research design implications follow.		
Library of Congress Classification (LCC)	5001-6182	Business	
	5546-5548.6	Office Organization and Management	
	5548.7-5548.85	Industrial Psychology	
	6951	Industrial sociology	
Journal of Economic Literature (JEL)	М	Business Administration and Business Economics	
	M 10	Business Administration: general	
	L 2	Firm Objectives, Organization and Behavior	
	L 29	Firm Objectives, Organization and Behavior : Others	
European Business Schools Library Group (EBSLG)	85 A	Business General	
	100B	Organization Theory (general)	
	240 B	Information Systems Management	
	80 J	Industrial sociology	
Gemeenschappelijke Onderwerpsontsluiting (GOO)			
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen	
	85.05	Management organisatie: algemeen	
	85.08	Organisatiesociologie, organisatiepsychologie	
	85.08	Organisatiesociologie, organisatiepsychologie	
Keywords GOO	Bedrijfskunde / Bedrijfseconomie		
	Organisatieleer, informatietechnologie, prestatiebeoordeling		
	Organisatiesociologie, Concepten		
Free keywords	Foundation of organization theory, Forms, Populations, Identity		
Other information	ĺ		

Foundations of a Theory of Social Forms¹

László Pólos Eötvös Loránd University, Budapest and Erasmus University, Rotterdam

> Michael T. Hannan Stanford University

Glenn R. Carroll University of California, Berkeley

March 30, 2000

¹This research was supported by ERIM of Erasmus University, the Center for Formal Studies in the Social Sciences at Eötvös Loránd University, the Stanford Graduate School of Business Faculty Trust, and the Institute of Industrial Relations at U.C. Berkeley. We benefitted greatly from the advice of Jim Baron, Jaap Kamps, Bart Nooteboom, Gábor Péli, and Ezra Zuckerman.

Abstract

Sociologists frequently invoke the concept of form when analyzing organizations, collective action, art, music, culture and other phenomena. Nonetheless, the form concept has not received careful theoretical analysis, either generally or in specific context. Using the tools of formal logic, set theory, and algebra, we propose a language for defining social forms that is sufficiently general to incorporate feature-based, position-based, and boundary-based approaches to defining forms. We focus on organizational forms although we intend for our conceptualization to be general. We define forms as a type of socially coded identity. We define identity in terms of social codes that specify the properties that an entity can legitimately possess. These codes can be enforced by insiders or outsiders. We claim that one knows that a social code exists when one observes that departures from the codes after periods of conformity cause a devaluation of the entity by relevant insiders and/or outsiders. This construction allows us to define a population as the set of entities with a common minimal external identity in a bounded system in period. The minimal property ensures that we localize to the most specific socially enforced identities. The reliance on identities instead of forms allows us to define populations that never achieve form status and to extend population definitions back to the period of early legitimation. Research design implications follow.

Foundations of a Theory of Social Forms

Introduction

In many areas of central interest, contemporary sociologists invoke the concept of form. For example, in analyzing French collective action across four centuries, Tilly (1986) describes the "forms of contention" that arose and spread through the provinces. These include seizures of grain; collective invasions of forbidden fields, forest and streams; attacks on machines; serenades; tendentious holiday parades; forced illuminations; turnouts; strikes; demonstrations; petition marches; planned insurrections, and electoral campaigns. Likewise, in developing a sociology of art, Becker (1982) characterizes various activities involving works of music, literature, film, television, sculpture, painting, and photography as "art forms." In organizational sociology, all major theorists occupy themselves in one way or another with "organizational forms," often used to describe established ways of producing most specific goods and services.

What exactly is a form? What roles do forms play in sociological theories? Despite the centrality of the concept in the research literatures on organizations, collective action, art, and culture, surprisingly little attention has been paid to clarifying these questions. Although ambiguity about the meaning of such a key concept might have been useful at early, exploratory stages of inquiry, it now hampers efforts to sharpen theoretically driven research programs; it also likely thwarts attempts at theoretical unification across disparate areas of the discipline.

This paper addresses this deficiency. We develop the foundations for a new general approach to social forms. We think that such work can clarify and organize many of the ideas invoked by the form concept. It might also stimulate theories in other areas of sociological study where the form concept might be useful.¹

Although more abstract than entities such as artistic movements or bureaucracies, social forms are routinely identified by persons and corporate actors. Indeed, most treatments of forms emphasize their socially constructed nature. So, in defining forms of organizations or other social entities, the conceptual apparatus should build toward theories that would provide empirically testable answers to the question: What do social agents recognize when they "see" a form; or, more precisely, how do they identify form boundaries? That is, the conceptual apparatus should be capable of supporting theories that reflect the ways in which social agents distinguish forms.

This complicated task becomes all the more difficult if the framework finds expression in "natural" language. Foundations would be far more secure if they rested on a formal language that minimizes ambiguity. Use of such a language

 $^{^{1}}$ We can imagine theoretical constructions such as the following being useful: forms of ethnicity, medicine, religion, etc.

for this purpose has other advantages as well. It permits checks on consistency, allowing us to learn whether the various substructures of the foundation interrelate felicitously. The abstraction gained in the translation to a formal language also eases the task of using the same ideas in social contexts outside the realm of organizational analysis, where we concentrate our current efforts. Even when labeled otherwise, form ideas figure prominently in analyses of such diverse phenomena as social protest, governance, contracting, and kinship. We attempt to construct the language so that it is useful in these, and other, contexts.

In recent sociological literature, the techniques of rational reconstruction and logical formalization have been used to reconstruct and improve the argumentative structure of theories. This paper, though also based on formal logic, differs in two important ways. First, instead of explicating existing texts of theoretical ideas, we draw on the research experiences of others and ourselves in examining social form distinctions made in the real world. We try to identify the processes that underlie these distinctions and construct a formal representation of the processes at work and the resulting distinctions in social forms.

The second difference concerns the formal machinery. Systematic analysis in sociology invariably employs the tools of first-order logic (including classical mathematics). We think that this choice will not do here. As we explain in some detail, we build notions of identity and form on the applicability of certain social codes. Expressing the codes requires a first-order language. Defining identities and forms in terms of predicates relating to the applicability of codes therefore requires a *second-order* language, one that allows quantification over sentences (formulæ) of the language, not merely over the objects in the universe of discourse (as in a first-order language). Moreover, we depart from other recent logical formalizations of sociological theories by paying explicit attention both to syntax and semantics, as we also explain below. For these two reasons, the formal structure we build departs extensively from convention.

The real test of the value of foundational work is whether it proves useful in building specific theories-a task primarily for future efforts. Nonetheless, some insight into potential value can be gained now by learning whether the revised foundations suggest changes in current research practice. We try to show that the scheme that we develop leads to rethinking of the basic unit in much sociological research: populations. We argue that the new scheme demands changes in the definition of population and in research practices in identifying populations empirically.

Forms in Organizational Theory and Research

Although we intend our approach to be general, it is helpful to focus on an area of application to give a sense of the motivations for our approach. We consider organizational sociology, where the notion of form has long held center stage.

Definitions of Organizational Forms

If any approach to defining organizational forms can be regarded as the standard, it is one that regards forms as particular clusters of features. The par excellence example is Weber's (1924) specification of rational-legal bureaucracy in terms of the nature of authority (professional expertise in evaluating abstract rationalized codes), procedures (impersonal exercise of authority and reliance on written rules and files), and the employment relation of the official (bureaucratic employment as a career of full-time work, office separated from the private sphere, and compensation by salary). Weber introduced his specification of rational-legal bureaucracy as an ideal type, meaning among other things that he had abstracted from the historical details in identifying the properties that define the form. Nonetheless, it was crucial to Weber's analysis—and it is to our approach—that the rational-legal form was instantiated in real organizations. For instance, Weber sought to capture the essential properties of the bureaucratic organizations of the late 19th-century Prussian state. Yet, he noted that the form was not limited to its Prussian instance; for instance, he argued that the civil-service bureaus created by the Progressive reform movements in the United States in the first decades of the twentieth century also fit the form. In the language of contemporary theory, the Prussian and American instances are two populations with the common form of rational-legal bureaucracy.

Subsequent developments of the feature-based conception recognize some features as more important than others in distinguishing forms. In this vein, analysts utilize the distinction between core and periphery and have identified forms with a set of core features. That is, organizations displaying the same core features are treated as belonging to the same form. Various analysts' notions of form differ mainly in envisioning different sets of features in the core.

The various feature-based approaches treat form distinctions as reflecting only structural arrangements, meaning that forms can be assessed in purely technical terms. Such a possibility can now be realized in the study of biotic evolution. Organizational sociology seems unlikely to reach such a position, because distinctions among forms appear to reflect social processes of boundary creation. This possibility has been explored in a second, less developed, line of work that defines forms in terms of the clarity and strength of *social boundaries*. Hannan and Freeman (1986) argued that the processes that create and reproduce the boundaries—social network ties, closed flows of personnel among a set of organizations, technological discontinuities, social movements articulating the interests of a set of organizations, and so forth—are the key to understanding forms. When these processes operate strongly and dominate the processes that blur forms, then the world of organizations is organized by forms.

DiMaggio (1986) proposed a similar approach, based on finding structural equivalence in flow networks [see also Burt (1993)]. Structural equivalence partitions a set of entities into equivalence classes that might be argued to represent forms. This proposal would actually define a population in our framework, because it pertains to a localized set of interacting entities. Forms, as we construe them, are more abstract entities; they potentially extend over space and time. Nonetheless, bringing network ideas into the picture is a very valuable step. So, we follow this lead. That is, our proposed definition of forms can be built in terms of network ties or other kinds of relational properties, such a position in a distribution of size or status.

Roles of Forms

In our view, the form concept plays three major roles in organizational theory and research. First, researchers use notions of form to define populations for study. In the typical framework, localized sets of actors having the same form constitute a population. Much recent research conducts analysis on populations of organizations so defined, including research on populations as art museums (Blau 1995), audit firms (Boone, Bröcheler, and Carroll 2000), banks (Barnett and Hansen 1996; Lomi 2000), baseball teams (Land, Davis, and Blau 1994), brewers (Carroll and Swaminathan 1992; 1998), credit unions (Barron 1999), ethnic and women's social-movement organizations (Minkoff 1999), investment banks (Park and Podolny 1999), newspapers (Olzak and West 1991; Dobrev 1999), semiconductor manufacturers (Podolny, Stuart, and Hannan 1995), symphony orchestras (Allmendinger and Hackman 1996), television stations (Sørensen 2000), wineries (Swaminathan 1995), and worker cooperatives (Russell 1995). The value of such research depends heavily on the populations' studied representing instances of meaningfully specified forms.

In its second common role, form refers to a particular configuration of properties. Examples include the multi-divisional or M-form (Williamson 1975) and the network form of organization (Podolny and Page 1998). In many cases, analysts have drawn attention to these particular configurations on the assumption that the form in question has some special adaptive advantage over others.

The examples of the M-form and network form illustrate yet another meaning of form: structural architecture. Architecture and form generally refer to different kinds of properties. Forms relate to identity, and architectures are rarely decisive for identity. In particular, architectures often vary considerably within forms. For instance, M-forms can be found in some, but not all, firms in many populations. Second, organizations also routinely change architectures without changing their identities. So, without denigrating the importance of architecture, we insist that it be distinguished from form.

The third common role of the form concept serves to differentiate between core and peripheral features. According to inertia theory (Hannan and Freeman 1984; Barnett and Carroll 1995), changing a core feature exposes an organization to great risk of mortality, but change in peripheral features does not. According to this theory, the core features are those that regulate form membership. Clear and detailed specifications of forms are needed to test such an argument.

In perhaps the most serious limitation, current definitions fail to link forms and identities. If forms are not architectures, then they presumably involve social and cultural typifications, agreed-upon classifications of entities into types. Our efforts to understand such typifications in empirical research on diverse kinds of organizations leads us to suggest that such processes build upon identities. So we seek to define identity and form in closely related terms.

An Approach to Forms

We think that systems of classification of forms will have most sociological value if they build on a theory of action. We suggest that forms be considered as recognizable patterns that take on rule-like standing and get enforced by social agents. Our choice of language for expressing this idea deserves some discussion. We want to use a term whose denotation and connotation include both cognitive recognition and imperative standing. When we formulated these issues in terms of social and cultural rules, we found that he connotation of articulated regulation is so strong that the cognitive dimension gets slighted. So we use the notion of code, which can be understood as both (1) a set of signals, as in the genetic code, and (2) a set of rules of conduct, as in the penal code. Our use of "code" reflects both meanings.

We begin by stating a formal language, using set theory and logic. We deploy these tools along with some standard sociological notions about code violation and identity. We surely do not claim originality for the basic ideas. We do think that the reformulation lends a new crispness to the old ideas.

The language must contain the elemental building blocks for identities and forms. As in the feature-based conception of form, features are central to our definition of form. As in the network-based conception, relations are also central. These relations might involve actual social ties; or they might represent resource dependencies or relative structural positions. For simplicity, we will refer to both features and relations as properties.

We introduce a language for expressing properties as well as constraints on properties. We represent properties as functions. Let \mathbf{v} denote a set of possible values of a property, say the alternative forms of authority, \mathbf{o} the set of objects, here social entities such as organizations, collective actions, contracts, and so forth, and \mathbf{p} the set of time periods. A property, $f(o, \pi)$, is a function $f: \mathbf{o} \times \mathbf{p} \to \mathbf{v}$. Such functions take an entity and a period and give the values of the properties.

Identities and Codes

The idea that forms should be grounded in identities raises a vexing question: What provides the basis of identity in the kinds of worlds we are considering? One standard answer says that identity inheres in the constancy of a set of properties. For instance, Leibniz's rule states that two objects are identical if and only if they possess exactly the same properties. This rule makes sense as a way of specifying identities for so-called extensional objects, those for whom all properties are fixed. Social entities, however, are intensional, their properties can change over time; and, some changes do not disrupt identity. So there is no point in looking for social identity in constancy.²

We reach the unavoidable conclusion that neither the actual values of the properties nor the properties themselves can be the carriers of identity. What, then, do social entities preserve when they maintain their identities? In short, they satisfy constraints. An identity constrains what an entity would/could be and what is expected and not expected of it. On the formal side, this idea can be expressed as positing that entities are described in terms of social codes relating to properties and constraints over properties.

Because we want to bring network position and other relational properties into the framework, we make what might seem to be an odd strategic choice. We define the properties that support identity broadly as including even properties involving the entity itself. We do so because sentences about network properties refer to the presence and absence of ties of the focal entity with others. For example, forms of contention require relational ties among contenders, opponents and others: "each of these forms of action links some concrete group of people to some other individual, group or groups. Each originates and changes as a function of continuing interaction—struggle, collaboration, competition, or some combination of them—among groups" (Tilly 1986: 4). If formal sentences for such images are to be well formed, both ego and alter must be included in the set of objects to which the sentences refer.

If properties are defined as functions, then each property has a range, a set of possible values. For instance, organizational status, which can range over the positive numbers, has a broad range; but the form of ownership, which includes only a modest number of values, e.g., proprietorship, partnership, collective ownership, state ownership, and shareholder ownership, has a narrow range. Many relational properties will obviously have very wide ranges.

The feasible values of a property should be distinguished from the values allowed by social codes. The feasible sets might usefully be taken as given in analyzing identity, because they are set by considerations that depend neither upon the actual entities under consideration nor on the codes defining their identities. For instance, some properties face technical constraints. Widespread availability of inexpensive information technology has increase the range of production possibilities, e.g., the option of marketing and selling products over the Internet was not in the feasible set until recently. Other properties face constraint from laws and general cultural rules, e.g., slave labor is now excluded in most societies from the set of options for the form of the relationship between the corporate entity and its members. Some identity-related restrictions go beyond the technological and legal constraints. At the extreme, an identity allows only one value for a particular property (an indispensable property), in which case an entity cannot change the value of this property without risking loss of

²It might be tempting to try to finesse this issue by conceptualizing identity in terms of some temporal functions defined on the properties. Unfortunately, such identification would require complete knowledge of the property values at every time point. Even full knowledge of an entity's past would not be enough. One also has to know its future to provide a satisfying definition of identity in terms of the time path of properties. Pólos (1999) discusses the technical details.

its identity.

Some restrictions that support identity have a conditional nature. Think of an identity that allows multiple values for each of two properties but requires that certain combinations not be allowed. For instance, the property of brokerage requires that an entity have ties to two or more other entities that are not directly tied. Each of these elements of identity can be expressed in terms of constraints.³

Aspects of a code might appear in recorded format. Modern legal codes tend to develop this explicit nature. Similarly, some of the cultural code that regulates forms of art might be recorded. But, most of the code regulating social behavior appears to be implicit and unrecorded, except perhaps by social critics. In a broad class of situations, researchers can only learn about the codes from the actions social actors take when they observe code violation.

Explicit codes are sometimes well preserved over centuries, even in the absence of any enforcement mechanisms. For instance, one can still read the legal codes of Hamurabbi. Implicit codes, on the other hand, are more fragile, and, in the absence of enforcement mechanism, they tend to vanish.

Since most social code is a composition of implicit and explicit codes, the reconstruction of social codes from the explicit part might be problematic. This might be responsible for the common difficulty people have with recalling obsolete cultural forms.

A Language for Coded Identities

Now we consider the common characteristics of these constraints to provide a formal definition of social identity. Each type of constraint mentioned above can be constructed from sentences (or formulæ) of a formal language.⁴ A companion paper (Pólos, Hannan, and Carroll 1999) provides a complete formal definition of the proposed language. Here we sketch the main ideas informally.

As we see it, a useful formal language for identities and forms must connect two⁵ semantic domains that refer to different ways of assessing the relation between entities (social actors) and coded constraints. The first holds that the entity is perceived as satisfying the constraint, and the second holds that the constraint operates as a *default* for the entity.

³Constraints circumscribe acceptable behavior in terms of ranges within which properties have to lie. This is often explained in conditional form. Our use of the notion of constraints was inspired by Barwise and Perry (1999) even though we did not adopt their formalism to represent constraints.

 $^{^{4}}$ In formal terms, the language of constraints is a first-order language, which means that it allows quantification only over the *objects* in the domain of discourse. Our formulation makes statements that quantify over the sentences/formulæ of this language. Therefore, it is a second-order formulation. Whenever we clarify certain ideas by formal means, we use a second-order language.

⁵It might prove useful in some applications and extensions of these ideas to add a third semantic domain that concerns whether an entity actually satisfies a constraint, that is, obeys the code. As this idea is not needed here, we do not introduce the third semantic in this paper.

The social-perception domain refers to the world of perceived objects and to the properties they are perceived to possess. The first semantics for our language provides meanings for such sentences or formulæ. To indicate this, we use the symbol p as a subscript. The truth value of a sentence expressing a code refers to perceived facts, whether a particular object is perceived as satisfying the code: $[\kappa]_p$, where the subscript p indicates that the truth value is evaluated in terms of the perceived facts of the case: whether a code is perceived to be satisfied. It is easy to see that this semantic domain requires a *partial* truthvalue assignment. Certain statements are perceived to be true, some others are perceived to be false; but some statements might not be perceived as either true or false.

In the domain of defaults, truth values follow the following rule: If a code is a default for an entity, then the entity is assumed to satisfy the code unless there is evidence to the contrary that over-rules this default. We denote the default truth value of any sentence expressing a code, κ , by $[\![\kappa]\!]_d$, where the subscript indicates that the truth value of the sentence is evaluated according to the default semantics. If the sentence κ is true for a particular entity, then the default truth value of the sentence equals one when the entity variable is "substituted" by the name of this object.

Two relations between these semantic domains deserve note. First, defaults "fill gaps" in perception. Social perception is by nature partial: it might be that the violation of a particular code is not perceived. If perception matters, then it is important to find ways how the absence of perception can be eliminated. We assume that actors treat the default codes as being satisfied between events of perception (since there is no evidence on satisfaction between such events,). Second, we assume that a code is a default for an entity if observed code violations have observable consequences for the entity. We will model these consequences in terms of a sudden drop in the valuation of the entity. In these terms, the relation between the perceptual semantics and the default semantics can be expressed as follows. A code is a default for an entity if the perceived violation of this code results in a sudden drop in valuation.

We extend these notions by defining a set of predicates to be used in defining social identities and forms. These predicates tie the semantics and the syntactical elements of our conceptual scheme. First, we define a predicated for perceived satisfaction using the semantics of social perception:

Definition 1 (Perception of code satisfaction)

Object ω is perceived as satisfying code κ at the time point τ :

$$\mathcal{P}_S(\omega, \kappa, \tau) \longleftrightarrow \llbracket (\kappa) \rrbracket_n = 1;$$

object ω is perceived as not satisfying code κ at τ :

$$\mathcal{P}_{\neg S}(\omega, \kappa, \tau) \longleftrightarrow \llbracket (\kappa) \rrbracket_n = 0;$$

object ω is not perceived as satisfying code κ at τ :

$$\neg \mathcal{P}_S(\omega, \kappa, \tau) \longleftrightarrow (\llbracket (\kappa) \rrbracket_n = 0 \lor \llbracket (\kappa) \rrbracket_n = 2);$$

object ω is not perceived as not satisfying code κ at τ :

$$\neg \mathcal{P}_{\neg S}(\omega, \kappa, \tau) \longleftrightarrow (\llbracket (\kappa) \rrbracket_p = 1 \lor \llbracket (\kappa) \rrbracket_p = 2).$$

Second, we define default code status, $D(\omega, \kappa, \pi)$, which reads as "code κ is a default for object ω in period π ." We define this predicate with reference to the default truth value.

Definition 2 (Default code) Code κ is a default for object ω in period π :

$$D(\omega, \kappa, \pi) \longleftrightarrow \llbracket \kappa \rrbracket_d = 1.$$

These definitions yield the following lemma:

Lemma 1 The semantics of perceived satisfaction gives the undefined value, the truth value gap, if neither satisfaction nor non-satisfaction is perceived. 6

$$\neg \mathcal{P}_{\neg S}(\omega, \kappa, \pi) \land \neg \mathcal{P}_{S}(\omega, \kappa, \pi) \longleftrightarrow \llbracket (\kappa) \rrbracket_{n} = 2.$$

Now we can spell out the formal relation among these semantic domains:

Postulate 1 If a code is a default and possible non-satisfaction is not perceived, it is just as good as if the satisfaction had been perceived.

$$D(\omega, \kappa, \pi) \wedge \neg \mathcal{P}_{\neg S}(\omega, \kappa, \pi) \longrightarrow \mathcal{P}_{S}(\omega, \kappa, \pi).$$

(Note that this is exactly how defaults are used in computation.)

Code Violations and Valuations

We need some principle for separating actual, socially meaningful identities from potential identities. We build on the idea that real identities come to have a rulelike status. This means that the identities consist of codes with the property that violations of defaults have observable consequences (Meyer and Rowan 1977; Becker 1982; Jepperson 1991). The κ -valuation function introduced below tells the direction and strength of the social approval of an entity in terms of its perceived satisfaction of applicable codes.

Understanding how normative judgments by agents combine to create, sustain, and apply codes is a fundamental problem in the social sciences. We do not pretend to offer a general solution to this problem. The framework we propose conforms with most relevant theory: it requires simply that such codes come into existence and that controllers of valued resources use these codes

⁶The partial nature of the perceptual semantics all us to give a formal semantics for default in a classical manner (Pólos, Hannan, and Carroll 1999). An alternative approach, using a nonmonotonic logic can be found in Pólos, Hannan, and Kamps (1999).

in granting and withholding access to the resources. We imagine that the details vary enormously for different identities (and forms) and among systems.⁷ The processes that translate individual judgments into consequences for actors presumably differ across cases (DiMaggio 1997). Indeed, sometimes different groups of evaluators impose different and possibly inconsistent constraints on an identity such that a given action is viewed positively by one set and negatively by the other. Although such identities are fragile and are unlikely to be sustained, the case of conflicting demands needs attention. But, we do not attend to these complications here; we simply assert that certain social codes exist and get enforced, without specifying who does the enforcing and how.

We want to express formally the idea that violation of a social code after a period of conformity lowers an entity's valuation precipitously. However, violation has negative consequences only when the entities controlling consequences treat the code in question as a default for the entity. The issue of applicability complicates the job of uncovering codes. We do not want to assume that a researcher already knows all relevant social codes. Suppose that an analyst comes across an entity that does not conform to some expected code-like pattern. Should she conclude that this unusual entity violates a social code? We think that the answer to this question is No, because the entity might satisfy another—yet unknown—code. To accommodate this complication, we suggest relying on a more limited test. This test focuses on a set of entities that conforms to what might be a social code and examines the consequences of an entity's violating the possible code after a period of perceived conformity. If an observed violation of the code after a period of conformity causes evaluations to drop sharply, then we conclude that we have identified a social code.⁸

This reasoning holds even when an entity changes in a way that leads to perception violation of one social code but perceived satisfaction of another code that it did not satisfy previously, e.g., an organization departs from its longstanding tradition and switches to a currently fashionable design. In this case, the entity does not satisfy a social code for only a vanishingly small interval. Yet, the violation of the former code should be costly. The matter turns on what codes are defaults. If the entity has an observed history of conforming to a code, then the relevant evaluators ordinarily come to treat this code as a default. Hence, the shift to conformity with some other code does, in the example, constitute a violation.

So the next step is to define a predicate for code violation, $V(\omega, \kappa, \sigma)$, which reads as "entity ω violated code κ during the infinitesimally short period σ ."

⁷For instance, Becker (1982) identifies consensus among art-world participants as paramount, especially those recognized as experts. He also claims that art works must be justified in terms of a logically coherent and defensible set of aesthetic principles. With respect to forms of contention, Tilly (1986) argues that two main sources have to do with the peculiarities of capitalist development and of concentration of political authority. Hannan and Freeman (1986) contend that organizational forms arise from various segregating processes.

⁸If codes and their application are known, then violation can be defined irrespective of prior observed conformity. We introduce the special restriction here only because we are considering common situations in which the researcher might not know codes.

Definition 3 (Code violation) Entity ω violates a code, κ , in the time interval σ if the code is a default for the entity and there are two time points τ, τ' such that the entity is perceived as satisfying the code or is not perceived as not satisfying it in any time interval within σ prior to τ , but is perceived as failing to satisfy it in $[\tau, \tau') \subseteq \sigma$.

$$V(\omega, \kappa, \sigma) \longleftrightarrow \exists \tau, \tau', \pi \left[(\sigma \subseteq \pi) \land ([\tau, \tau') \subseteq \sigma) \land D(\omega, \kappa, \pi) \land \forall \tau'' [([\tau'', \tau) \subseteq \sigma) \land (\mathcal{P}_S(\omega, \kappa, [\tau'', \tau)) \lor \neg \mathcal{P}_{\neg S}(\omega, \kappa, [\tau'', \tau))) \land \mathcal{P}_{\neg S}(\omega, \kappa, [\tau, \tau'))] \right]$$

(Here, as elsewhere in the paper, σ is a variable that ranges over infinitesimally short intervals, as contrasted with π which ranges over longer intervals. We assume that the expression $[\tau, \tau')$ is well formed only if $\tau < \tau'$.)

According to this definition, all that matters is whether the code is a default for an entity and whether the entity is perceived as failing to satisfy it. It follows from this approach that violation of a code usually prompts an evaluator to shift from automatic cognition to deliberative cognition (DiMaggio 1997).

Many social entities, such as organizations, protest actions, and artistic schools, are composed of members who are themselves actors. So, in general, we want to consider cases in which entities face evaluation by both insiders and outsiders. The various kinds of evaluators might have different perspectives and might impose different, perhaps conflicting, demands. At this stage of development of the conceptual framework, we elide most differences and assume that evaluations can be meaningfully aggregated within each of two sets of actors: insiders and outsiders.

Valuation can involve all sorts of social reactions based upon identity. For instance, it can involve judgments by members about whether a corporate actor remains faithful to its traditions and culture. Valuation can also mean the literal calculation of pecuniary value in the sense that investors in stock markets explicitly or implicitly make calculations of the expected value of firms. When these expectations involve conformity of forms with certain coded identities, then the process we are discussing is relevant to the definition of identities. Zuckerman (1999) documents such a process in the world of securities analysts, employees of stock brokerages who follow industries and firms and make forwardlooking assessments of the stock market prospects of firms. These valuations depend upon a variety of objective and subjective measures of performance. Zuckerman also shows that they depend upon the conformity of the firms with the category schemes used by the analysts. Firms that operate in markets and industries that do not fall in one of the analysts' categories are less likely to be followed by analysts. This reduces their attractiveness to investors, and their stock market returns fall accordingly.

Similarly, Becker (1982: 135) shows that the classifications of aestheticians and critics affect artists and their activities: "whether free-form jazz is really jazz and whether fashion photographs are really photography and therefore art, are discussions, among other things, about whether people who play freeform jazz can perform in jazz clubs for existing audiences and whether fashion photographs can be exhibited and sold in important galleries and museums." These artistic valuations affect material rewards essential to the production of art, intangible benefits that shape careers, and even the life and death of whole genres.

One important class of valuation pertains to judgments about the social worth of an identity, such as the distinction between high-brow and low-brow art forms. We focus on a different kind of valuation, one that depends upon judgments about whether a social entity follows the relevant default codes. (Even critics who find little artistic value in a genre can make judgments about the degree to which an artist or an art work conforms to the rule of the genre.) We refer to such judgments as κ -valuations, for short. The agents need not be able to articulate fully the bases of their valuations. Becker (1982: 199) observes that "artists find it difficult to verbalize the general principles on which they make their choices. They often resort to such noncommunicative statements as 'it sounds better that way,' 'it looked good to me,' or 'it works'."

Of course, such valuations might reflect legal or regulatory codes and their enforcement. State authorities, professional bodies, and other collective actors frequently enact formal rules that govern identities, especially in the case of organizations. For instance, the world of American financial-service organizations was sharply constrained by regulation for roughly fifty years in the middle of this century, with legal codes stipulating a set of allowed combinations and forbidding other combinations (such as combining commercial banking and investment banking in the same organization).

We introduce κ -valuation functions pertaining the judgments of insiders and outsiders. Although it would be extremely important in empirical research to measure these functions directly, we simplify here by concentrating on the changes in valuations over a period. Because we want to consider punishment of deviations, we single out situations in which valuations drop substantially and instantaneously. We represent this idea as follows. Let $\Delta_i(\omega, \kappa, \sigma)$ mean that the insider κ -valuation function for object ω , with respect to code κ , dropped substantially in instantaneous period σ ; and let $\Delta_e(\omega, \kappa, \sigma)$ do the same for external valuation.⁹

We use the κ -valuation functions to define predicates that tell that a code is socially enforced: $\mathcal{SCI}(\omega, \pi)$ reads as "the sentence κ is an internal social code in period π ," and $\mathcal{SCE}(\omega, \pi)$ is defined similarly for external identity.

Definition 4 (Social code enforced internally/externally) An internally (externally) enforced social code is a default code whose perceived violation causes a

⁹If big changes in the κ -valuation of an identity can take place instantaneously, then κ -valuations of identities do not serve as a basis for potentially stable social identities. So, we make the seemingly reasonable assumption that κ -valuation functions change smoothly over brief intervals for entities that are perceived to satisfy their default codes over the subinterval.

discontinuous drop in an entity's κ -valuation by insiders (outsiders)

$$SCI(\kappa,\pi) \longleftrightarrow \forall \omega, \sigma \left[D(\omega,\kappa,\pi) \land V(\omega,\kappa,\sigma) \land (\sigma \subseteq \pi) \to \Delta_i(\omega,\kappa,\sigma) \right].$$
$$SCE(\kappa,\pi) \longleftrightarrow \forall \omega, \sigma \left[D(\omega,\kappa,\pi) \land V(\omega,\kappa,\sigma) \land (\sigma \subseteq \pi) \to \Delta_e(\omega,\kappa,\sigma) \right].$$

The two types of codes might disagree. Perhaps nowhere do the differences between insiders and outsiders become so visible as with art forms, where works that belie the external public's common-sense view of art at the same time gain high esteem within insider artistic circles. Becker (1982: 146) describes such divergence in discussing Marcel Duchamp's exhibitions of a signed shovel and a signed urinal:

The common-sense critique of these works is that anyone could have done, that they require no skill or insight, that they do not imitate anything in nature because they are nature, that they do not express anything interesting because they are no more than commonplace objects. The critique of those with finer sensibilities is much the same. Nevertheless, those works gained great renown in the world of contemporary visual art, inspiring many more like them.

A Formal Definition of Social Identity

We move from social codes to social identities in two steps to make the analytic strategy clear. First, we collect and name all of default social codes for an actor. We want to reduce the identity to its core, stripping away incidental features that might be attached. We achieve this by defining the *building blocks of social identity*. It will simplify things to separate the ingredients of internal identity (**be**).

Social identities, which can potentially be separated from particular carriers, should be defined abstractly. But, we want to ensure that we are working with tangible codes, those that actually serve as defaults for some entities for at least some of the relevant period. Therefore, we start by defining the tangible building blocks.

Definition 5 (Tangible building blocks for identity)

$$\mathbf{bi}_{\omega,\pi} = \{\kappa | SCI(\kappa,\pi) \land D(\omega,\kappa,\pi)\}; \\ \mathbf{be}_{\omega,\pi} = \{\kappa | SCE(\kappa,\pi) \land D(\omega,\kappa,\pi)\}.$$

Note that the D predicates tie the codes to entities in these formulæ.

Next, we free the building blocks from the carriers. Certain identities might exist even when no actual entity carries them. It is sufficient that there was an entity that carried this identity. But, enforcement mechanisms for an identity do not last forever. If no one is subject to a code, then its enforcement mechanism weakens and eventually vanishes. We can characterize the pharaoh's identity even though there are no pharaohs; but the Pharaoh code lacks means of enforcement. We represent these considerations by introducing a parameter, θ , that tells how long it takes for an empty identity to lose its rule-like status and thus its enforcement. Let $\Lambda(\pi)$ denote the length of period π . We implement the idea that enforcement of an identity code dissipates if the identity is continuously vacant for $\Lambda(\pi) > \theta$.

Definition 6 (Abstract building blocks for identity)

$$\begin{aligned} \mathbf{b}\mathbf{i}_{[\tau_{1},\tau_{2})} &= \left\{ x \mid \exists \omega, \tau \left[(\tau_{1} < \tau \leq \tau_{2}) \land x \in \mathbf{b}\mathbf{i}_{\omega,[t_{1},t)} \right] \land \\ \neg \exists \pi \left[\pi \subseteq [\tau_{1},\tau_{2}) \land (\Lambda(\pi) > \theta) \land \forall \pi', \omega \left[\pi' \subseteq \pi \to x \notin \mathbf{b}\mathbf{i}_{\omega,\pi'} \right] \right] \right\}; \\ \mathbf{b}\mathbf{e}_{[\tau_{1},\tau_{2})} &= \left\{ x \mid \exists \omega, \tau \left[(\tau_{1} < \tau \leq \tau_{2}) \land x \in \mathbf{b}\mathbf{e}_{\omega,[\tau_{1},\tau)} \right] \land \\ \neg \exists \pi \left[\pi \subseteq [\tau_{1},\tau_{2}) \land (\Lambda(\pi) > \theta) \land \forall \pi', \omega \left[\pi' \subseteq \pi \to x \notin \mathbf{b}\mathbf{e}_{\omega,\pi'} \right] \right] \right\}. \end{aligned}$$

These complicated constructions say the following: For some period, the abstract building blocks for internal/external identity is the collection of those elements that are part of the internal/external identity for some entity for at least some time point within the interval as long as there is no subinterval longer than θ in which these elements are not part of the internal/external identity for some entity.)

The second step in moving from codes to identities introduces *composition rules* that combine building blocks in particular ways. Just as a recipe specifies a sequence of combinations of ingredients along with operations on those ingredients, an identity code imposes a pattern on the social codes that underlie the identity. For instance, in forms of organization, organizational identities often entail conditional constraints across the component codes, as we discuss below for microbreweries and brewpubs. In forms of contention, given sequences of activities come to be expected, as can be seen in Tilly's (1986: 394) description of the French turnout:

Workers in a given craft who had a grievance against the employers of their locality went from shop to shop within the locality, calling out the workers to join them in a march around the town, ended the circuit with a meeting at the edge of town, voted to make certain set of demands, sent a delegation to the employers, declared a work stoppage, and enforced it as best they could throughout the town until they reached an agreement with the employers.

In forms of music, a variety of different factors apparently come into play. Consider Peterson's (1997: 218) description of what it takes to produce the American country-music form: Music and performance are vital to the audience, but signifiers are also vital. The boots, the hat, the outfit, a soft rural Southern accent, as well as the sound and subjects of the songs, all help. Finally, being able to show a family heritage in country music is perhaps the strongest asset among authenticity claims. Many artists recall learning first from their mothers or playing in a family band.

A plausible interpretation of the widespread use of prototypes or templates in business, art, and other domains is that they embody not only the feature values inherent in a code, but also its composition rules. Prototypes are useful in attempts to recreate or copy a code, especially when the composition rules are tacit or, at least, not explicit.

We denote the composition rules for inside and outside identities as the function ι and ι' .

Definition 7 (Tangible internal/external identity codes) A tangible internal (external) identity is a set with the property that each of its members is an internally enforced composition of building blocks of identity for some entity.

$$\mathbf{iic}_{\omega,\pi} = \{ \alpha | \exists A \subseteq \mathbf{bi}_{\omega,\pi} \land \alpha = \iota(A) \land \iota(A) \in \mathbf{bi}_{\omega,\pi} \};\\ \mathbf{eic}_{\omega,\pi} = \{ \alpha | \exists A \subseteq \mathbf{be}_{\omega,\pi} \land \alpha = \iota'(A) \land \iota'(A) \in \mathbf{be}_{\omega,\pi} \}$$

In each case, the ingredients used to build an identity are a subset of the full set of ingredients. The formulæ state that a particular composition of the subsets of ingredients is coded as an identity. So the social coding process can pay attention to all or only to some of the ingredients in constructing identities. Note that this definition stipulates that the composition—rather than the elements of identity—that gets enforced. This property gives identity its unitary character.

Again, we free the codes from their carriers:

Definition 8 (Abstract internal/external identity codes) An internal (external) identity is a set with the property that each of its members is an internally (externally) enforced composition of building blocks of internal (external) identity.

$$\mathbf{iic}_{\pi} = \{ \alpha | \exists A \subseteq \mathbf{bi}_{\pi} \land \alpha = \iota(A) \land \iota(A) \in \mathbf{bi}_{\pi} \};\\ \mathbf{eic}_{\pi} = \{ \alpha | \exists A \subset \mathbf{be}_{\pi} \land \alpha = \iota'(A) \land \iota'(A) \in \mathbf{be}_{\pi} \}.$$

An obvious step in developing the notion of code-based social identity considers the consequence of violations:

Theorem 1 An observable violation of an internal (external) social identity causes a discontinuous drop in an entity's κ -valuation by insiders (outsiders).

$$\forall \omega, \pi, \alpha, \sigma [((\sigma \subset \pi) \land \alpha \in \mathbf{iic}_{\omega, \pi} \land V(\omega, \alpha, \sigma)) \longrightarrow \Delta_i(\omega, \alpha, \sigma)]; \forall \omega, \pi, \alpha, \sigma [((\sigma \subset \pi) \land \alpha \in \mathbf{eic}_{\omega, \pi} \land V(\omega, \alpha, \sigma)) \longrightarrow \Delta_e(\omega, \alpha, \sigma)]$$

This theorem provides the analytical leverage needed to provide a definition of social identity that can guide empirical research. It defines identities as those that have rule-like status, viz., defecting from compliance with the code harms the entity's valuation.

Grading Identities

Social identities often form nestings. Of course, in any given context, it is a matter for empirical research to document actual nesting patterns and to identify the processes that generate them. For instance, we know that if an organization has the identity "craft labor union," then it usually also has the identity "labor union", but not vice versa). Thus, we would conclude that craft union is a refinement of the labor-union identity, that the former is a more restrictive identity.

To define the specificity ordering of codes, we need the following:

Definition 9 (Sharper relation) Code κ is sharper than code κ' in period π , $\kappa \prec_{\pi} \kappa'$, if κ is a default then κ' is a default and perceived satisfaction of κ implies perceived satisfaction of κ' , but default status (perceived satisfaction) of κ' does not imply default status (perceived satisfaction) of κ .

$$(\kappa \prec_{\pi} \kappa') \longleftrightarrow \forall \omega [D(\omega, \kappa, \pi) \to D(\omega, \kappa', \pi) \land \neg (D(\omega, \kappa', \pi) \to D(\omega, \kappa, \pi)) \land \mathcal{P}_{S}(\omega, \kappa, \pi) \to \mathcal{P}_{S}(\omega, \kappa', \pi) \land \neg (\mathcal{P}_{S}(\omega, \kappa', \pi) \to \mathcal{P}_{S}(\omega, \kappa, \pi))].$$

Definition 10 (Ordering of tangible (internal/external) identity codes)

$$(\alpha \prec_{\pi}^{ic,\omega} \alpha') \longleftrightarrow (\alpha, \alpha' \in \mathbf{iic}_{\omega,\pi} \lor \alpha, \alpha' \in \mathbf{eic}_{\omega,\pi}) \land (\alpha \prec_{\pi} \alpha'))$$

We can define, *mutatis mutandis*, orderings of abstract (internal/external) identity codes:

Definition 11 (Ordering of abstract (internal/external) identity codes)

$$(\alpha \prec_{\pi}^{ic} \alpha') \longleftrightarrow (\alpha, \alpha' \in \mathbf{iic}_{\pi} \lor \alpha, \alpha' \in \mathbf{eic}_{\pi}) \land (\alpha \prec_{\pi} \alpha')).$$

Of course, not all pairs of identities can be ordered by specificity, e.g., jazz and country music.

It will be important in applying our framework to have an explicit definition of the sharpest external identities that apply to an entity. We refer to identities in the plural here because the lack of a hierarchical structure of identities allows entities to have several identities each of which satisfies the property that no other identity is sharper.

Definition 12 (Minimal external identities) $\operatorname{eic}_{\omega,\pi}^{\min}$ is a minimal external identity for an entity in a period if it is an external identity for the entity and none of the entity's other external identities is sharper.

$$\mathbf{eic}_{\omega,\pi}^{\min} = \left\{ (\alpha | \alpha \in \mathbf{eic}_{\omega,\pi}) \land \neg \exists \alpha' \left[\alpha' \in \mathbf{eic}_{\omega,\pi} \land (\alpha \prec_{\pi}^{ic,\omega} \alpha') \right] \right\}.$$

Note that an actor need not have a unique minimal identity.

Forms

After reviewing the role of forms in diverse kinds of sociological arguments, we concluded that forms could best play their assigned roles if forms are defined as social identities that possess three properties. First, forms are built from external identity codes. Identities might be established primarily—or even exclusively—by evaluations of insiders: think of a secret society whose structure has avoided notice by outsiders. In contrast, forms are cultural objects with broader significance. As Becker (1982: 163) explains for art,

... in principle any object or action can be legitimated as art, but ... in practice every art world has procedures and rules governing legitimation which, while not clear-cut or foolproof, nevertheless make the success of some candidates for the status of art very unlikely. Those procedures and rules are contained in the conventions and patterns of cooperation by which art worlds carry on their routine activities.

So we define the default social codes to entities as the external identity codes. (Of course, they might be enforced internally as well; but that is not required.)

Second, unlike identities, forms pertain to multiple entities. Sociologists want forms to be patterns with broad reach. So we also add this requirement. But, how broad must they be?

Here we take advantage of knowledge about density-dependent legitimation processes (Hannan and Carroll 1992; Carroll and Hannan 2000). According to this theory and research, external identities gain taken-for-granted status as a consequence of growth in the number of actors with the identity. As this number grows from zero, the constitutive legitimation of the identity increases to some ceiling. In empirical research (that also specifies the other part of the theory of density dependence concerning density and competition) we find evidence that the implied growth trajectories of legitimation with growing density vary among identities. So we specify a *form-specific application number*, denoted as $\nu(\phi)$ which gives the number of entities to which a social identity must apply for the identity to gain the standing of form.

Third, forms last longer than ordinary identities. Recall that we defined abstract identities as having the capacity to persist for a period of length $\Lambda(\pi) = \theta$ even when vacant. To represent the idea that forms have greater persistence, we assume that they can be enforced after much longer periods of vacancy, $\Lambda(\pi) = \Theta$, where Θ is much larger than θ .

These three considerations motivate the following definition:

Definition 13 (Form) A form is a default external identity code for at least $\nu(\phi)$ -many different entities at the beginning of a period and the length of the maximal

sub-period within which ϕ is not an external identity of any entity is below Θ .

$$\Phi(\phi, [\tau_1, \tau_2)) \longleftrightarrow \exists \tau \left[(\tau_1 < \tau \le \tau_2) \land (\nu(\phi) \le \left| \left\{ \omega \mid \phi \in \mathbf{eic}_{\omega, [\tau_1, \tau)} \right\} \right| \right) \right] \land$$
$$\neg \exists \pi \left[\pi \subseteq [\tau_1, \tau_2) \land (\Lambda(\pi) > \Theta) \land \forall \pi', \omega \left[\pi' \subseteq \pi \to \phi \notin \mathbf{eic}_{\omega, \pi'} \right] \right].$$

This definition relies on the construction of the set $\{\omega | \phi \in \mathbf{eic}_{\omega,[\tau_1,\tau)}\}$. This is the set of entities for whom the code ϕ is part of their external set of identities $\mathbf{eic}_{\omega,\pi}$. We use the standard notation for the cardinality of a set $|\cdot|$; the cardinality is the number of distinct objects in a set. The definition says that ϕ can be a form only when the number of such entities is at least as great as $\nu(\phi)$.

This definition allows for the possibility that forms can become empty for a time without the relevant codes changing. Under such circumstances, the external enforcement is obviously counterfactual.¹⁰ We do this by requiring that the $\nu(\phi)$ condition be satisfied for some (half-open) interval, $[\tau_1, \tau)$, where τ_1 is the starting point of the time interval under consideration. Even if the number of entities that satisfy the external identity condition drops below $\nu(\phi)$ (perhaps even reaching zero) after τ , the form still exists for the remainder of the interval. This construction leaves it to the researcher to determine the appropriate width of the interval, depending on the context. We imagine that the width of an appropriate interval for artistic forms might be considerably wider than for organizational forms.

An empirical illustration helps to show what it means for a form to persist during a period of inactivity. Carroll and Swaminathan (1998) plotted the number of firms (density) in the population of American brewers over the history of the population. The density falls to zero in 1920 with the onset of the national Prohibition and stays at zero until 1934, the year following the end of Prohibition. Interestingly, the level and trend in density after Prohibition is what one would have expected if the trend to 1920 had been continued and there was no Prohibition. The form seems to have been unaffected by the many years of inactivity. For another example, see Dobrev's (1999) analysis of the evolution of the population of newspapers in Bulgaria over regimes of free and highly regulated press.

To complete the framework, we tie forms to entities using the predicate $\mathcal{F}(\omega, \phi, \pi)$:

Definition 14 (Default classification) An object is classified as subject to a form if the codes specifying the form apply to it by default:

$$\mathcal{F}(\omega,\phi,\pi) \longleftrightarrow \Phi(\phi,\pi) \wedge D(\omega,\phi,\pi).$$

 $^{^{10}}$ This situation can be paraphrased as follows: Had there been a violation of the formdefining constraints, the violator would have been expelled from the form and would have to face all of the consequences.

Form membership, as we define it, does not imply that a default member actually conforms to the codes specifying the form. Rather, it means that these codes are defaults. It should be noted that classification can be unintended or undesired. As with identities generally, form classification gets externally conferred.

Our conceptual scheme yields the following result as a theorem:

Theorem 2 If a social entity violates the code that specifies a form to which it is classified, then its valuation by outsiders drops sharply.

$$\forall \omega, \pi, \sigma \left[\left(\mathcal{F}(\omega, \phi, \pi) \land (\sigma \subseteq \pi) \land V(\omega, \phi, \sigma) \right) \longrightarrow \Delta_e(\omega, \phi, \sigma) \right].$$

Subforms

Form distinctions are ordinarily partially nested, because identities are partially nested and forms are special kinds of identities. Social worlds sometimes contain sets of forms that have subforms that, in turn, have subforms, and so forth. Such a hierarchical arrangement allows the set of social codes to take an especially simple form. Hierarchical systems can be easily comprehended (DiMaggio 1997). Code violations can be seen clearly and, therefore, can be punished easily in such a world.

The key property for conceptualizing hierarchical relations among forms is the subform relation, which we denote by \prec_{π}^{ϕ} . (The subscript π indicates that the subform relation might be period-specific.) This subform relation is the sharper relation (\prec_{π}) applied to forms. This means that subforms are more constrained than the forms that they specialize.

Definition 15 (Subform) Suppose that ϕ_1 and ϕ_2 are forms in period π : $\Phi(\phi_1, \pi)$, and $\Phi(\phi_2, \pi)$. Then ϕ_1 is a subform of ϕ_2 iff ϕ_1 is sharper than ϕ_2 :

$$(\phi_1 \prec^{\phi}_{\pi} \phi_2) \longleftrightarrow ((\phi_1 \prec_{\pi} \phi_2) \land \Phi(\phi_1, \pi) \land \Phi(\phi_2, \pi))$$

We close this section with a theorem concerning the nesting of forms, as we define them.

Theorem 3 Default form classification propagates upward: default classification in a subform of a form implies default classification in that form.

$$\forall \phi_1, \phi_2, \omega, \pi \left| \left((\phi_1 \prec_{\pi}^{\phi} \phi_2) \land \mathcal{F}(\omega, \phi_1, \pi) \right) \longrightarrow \mathcal{F}(\omega, \phi_2, \pi) \right|.$$

Clarifying the Status of Populations

The notion of population plays a central role in social research. Nearly all researchers, influenced by statistical formulations, claim to make inferences to some population. But, populations are often defined opportunistically, with availability of data determining the bounds of the entity; and the concept of population as a sociological idea is poorly articulated. White (1992: 17) refers to the mainstream sociological conception of population as the "bag of beans" view. As long as the population concept remains underdeveloped, the link between abstract theory and systematic empirical research remains tenuous. We think that the formal language we are proposing can help to remedy this situation. This section sketches a theoretical conception of population that builds on notions of socially enforced identities.

The most sustained attention to population in sociology comes from the ecological tradition. There populations are defined as bounded sets of entities with a common form (Hawley 1968; Hannan and Freeman 1977). Populations are limited by some social-system boundary, which reflects barriers to the operations of relevant social processes such as the flow of information, competition, and regulation. Getting this part of the specification right is crucial for defining a population in our scheme. In this respect, nothing has changed from the usual standard in ecological research. However, our proposed conception of forms leads us to depart from the conventions of this approach. These departures reflect the insight gained from the formalization that points to a way to eliminate two ambiguities in the conventional ecological definition of population, not noted previously.

The first ambiguity arises because the codes that specify identities and forms are generally partially nested. An entity for which the social code specifying a particular external identity is also subject to the defaults for any higherlevel identity in the nesting. Which of the many possible sets are meaningful populations? A useful definition of population must provide clear guidance on this issue.

There seems to be widespread agreement that population definitions have most value for theory and research when they pertain to the most specific external (sub)identity applicable to an entity. The whole point of defining and identifying populations is to allow analysis of *local* social structures and the interactions they entail. A useful specification of a population should single out a set of entities that are expected to interact strongly both because they fall within the same system boundary and they share a common, highly-specific external identity. This reasoning motivates the first important departure from the standard approach: we define populations in terms of minimal external identities—not forms.

The second ambiguity concerns the theoretical standing of population at inception when populations are defined in terms of forms. It is helpful in thinking about this issue to consider working backward through the history of a population. As we move backward in the history, we eventually reach the point at which the population-defining form takes hold in the culture. Then imagine pushing the history back further, continuing to the point at which we see the first known entity that satisfies the rule that eventually becomes a form. This is not a purely imaginary exercise. Most contemporary research on the evolution of populations of organizations tries to push observations back to this kind of origin (Carroll and Hannan 2000), partly because we have learned that observations on the very early history are crucial for examining processes of legitimation (Hannan and Carroll 1992). But, what is the theoretical status of such a collection of entities in the period before a form has been established? If we define populations in terms of forms, we lack a warrant for extending the definition of the population back to the origin. Obviously, we would prefer a definition of population that allows meaningful specification of a single unit that spans the entire history.

Our foundational material suggests that defining populations in terms of external identities—instead of forms—provides a potential solution. We can imagine that each new socially enforced external identity initiates a population. Most such populations will amount to little. However, sometimes other entities come to share this identity and a process of legitimation begins. As the number of entities with this default identity grows, the taken-for-grantedness of the identity strengthens until it reaches a ceiling (Hannan and Carroll 1992). Our definition of form reflects this reasoning in stating the requirement that an external identity gains the status of form only when some specific number of instances, $\nu(\phi)$, exist. Now we can interpret $\nu(\phi)$ as the ceiling, the level of density at which the identity takes on the character of a form. The period in a population's history between its inception and the time at which density surpasses $\nu(\phi)$ is the crucial period of legitimation in the sense of taken-for-grantedness.

Defining populations in terms of minimal identities resolves the two ambiguities. The minimal property ensures that populations are localized to the most specific socially enforced identities. The reliance on identities instead of forms allows us to define populations that never achieve form status and to extend meaningfully the definitions of population back to the period of early legitimation.

We think that this approach offers another advantage. Organizational analysts are often struck by what they see as a strong social-movement flavor in the early history of a population with activists seeking to promulgate definitions and codes of conduct for members of the movement (Carroll and Hannan 2000). In our conception, such activity plays an extremely important role. To see this, consider the question: how are external identities enforced before an identity has become sufficiently taken for granted to be regarded as a form? Activists and movement entrepreneurs play precisely this role. Absent the webs of cooperative action and standard setting that characterize early histories of many successful population, social identities are unlikely to find external enforcement. The absence of strong enforcement makes it unlikely that an identity will achieve the standing of form.

We also need to formalize the notion of boundedness. We introduce a predicate $L(\omega, \lambda, \pi)$ that reads as "actor ω is located in a particular bounded system λ in period π ." As we explain in the next section, just because we can write a predicate for boundedness does not mean that it is a simple matter to establish the relevant boundaries. Because the relevant boundaries depend upon the processes under study, getting the boundaries right will often be as difficult as getting the forms right.

Now we can provide an improved formal definition of population:

Definition 16 (Population) A population is a set of entities with a common minimal external identity in a bounded system within a period.

$$\Pi(\phi,\lambda,\pi) = \left\{ \omega \mid (\phi \in \mathbf{eic}_{\omega,\pi}^{\min}) \land L(\omega,\lambda,\pi) \right\}.$$

Implications for Research Design

We think that this new formulation has far-reaching implications for practice. First, meaningful specification of a population depends upon (1) precise understanding of the social codes that specify externally enforced identities and (2) proper delineation of the system boundaries, i.e., choosing L appropriately. These are challenging requirements. Correct delineation of the social codes and system boundaries requires intensive study of the institutional context within which a set of organizations operates.

We think that these requirements also have strong implications for designs of empirical studies. Current research pursues two main design philosophies: one tries to get broad representativeness of the world of organizations over relatively short time periods (e.g., the National Organizations Survey [Kalleberg et al. 1996]), and the other, the population design, focuses narrowly on a segment of this world but uses a long temporal perspective (Carroll and Hannan 2000). The representativeness design maximizes variability in forms, while the population design minimizes such variability. The representativeness design would seem to provide more information about forms. However, it seems implausible, in the present state of knowledge, to identify the forms meaningfully. If many forms are represented, then the analyst must specify correctly an enormous number of social codes.

Our experience, based on application of the simpler population design to a relatively small number of arenas (audit companies, automobile manufacturers, banks, credit unions, life insurance companies, manufacturing firms, brewers, labor unions, newspapers, and semiconductor manufacturers) indicates that several years of close study is required to grasp the important institutional details that provide the key information about identities and forms. Whatever one decides about the merits of the two designs, an appreciation of the centrality of social codes in defining forms leads to a deep re-examination of the issues involved.

Second, because social codes and relevant system boundaries vary over time, population definitions are historically specific and change over time. Therefore, it is usually a mistake to impose an invariant population specification over long historical periods. The American brewing industry provides a useful illustration of the challenges (Carroll and Swaminathan 1998). Over several centuries, all brewers were what today are called microbrewers, because they used hand-crafted methods and traditional ingredients. They also operated in local markets, with the city or town defining the boundaries of competition. Over this period, then, the industry consisted of a great many localized populations, with each population exemplifying the same external identity. Of course, such enterprises were called brewers, not microbrewers. The microbrewing identity has an oppositional flavor; such an identity requires the existence of another identity that is not micro.

During this century, the brewing industry bifurcated in form, as some firms began to produce beer using industrial methods and new recipes, which included many additives not culturally sanctioned previously, and others retained traditional methods. By the end of World War II, the mass-production brewers had attained dominance in the industry, which became exceedingly concentrated. At the same time, the boundaries shifted to the national—and even the international—level. At that time, there was a national population of massproduction brewers and a lingering populations of traditional brewers.

Within the past two decades, population boundaries have shifted once again. Movements of local associations of hobby brewers succeeded in gaining legislation in various states that permitted traditional brewers to operate in a new mode. Specifically, the new form integrates production, sale, and consumption at a single site. This arrangement was—and still is—forbidden to the major brewers. It has come to be known as the brewpub form. The brewpub's emergence occurred hand in hand with development of another form: the micro-brewery, which is a small-scale brewery producing for distribution. The numbers of both types of breweries have grown into the hundreds in little over a decade (Carroll and Swaminathan 1998). A crucial part of the story of the rise of these two types is that they became external identities and then forms in the sense in which we use the terms. The activists in the movement succeeded in creating two distinctive identities that differ sharply from the identity of mass brewer. The key was posing the issues in terms of authenticity, involving slavish conformity to a set of specific traditional methods and the invocation of nostalgic images of craft production.

The rise of the two forms, of course, amounts to a recreation of the historic forms in the industry. This development has complicated the industry's topography of forms in two ways. First, the diversity of forms has increased. Second, the boundaries of the populations defined by these forms differ. The industrial, mass-brewer population is still national. So there is a single population of mass brewers. The brewpub population is, by virtue of its dependence on consumption at the site of sale, exceedingly local with respect to market competition, but not flows of information and legitimacy. The microbrewer population stands in an intermediate position, with some brewers emphasizing their local roots and others striving to operate on a national level. According to this account, the brewing industry has persisted for hundreds of years, but the populations of brewers have changed markedly over time.

A third implication of the definition bears on the enumeration of populations in the context in which form distinctions are nested. Consider the example of American national labor unions. Hannan and Freeman (1987; 1988) analyzed the ecological dynamics of the entire population of American labor unions and conducted analyzes of the two important subidentities and subforms: craft unions (in which workers are organized according to craft or occupational distinctions) and industrial unions (in which workers are organized by site of employment irrespective of craft or occupation). According to their understanding, the craft and industrial forms were minimal. The proposed definition would suggest that the Hannan-Freeman study amassed data on three populations: craft unions, industrial unions, and unions that did not fit either minimal form but still fit the labor-union form. In this view, the set of all unions does not comprise a meaningful population and should not be the main focus of analysis. Instead, analyses of the subforms should be given priority. Similarly, our sketch of the brewing industry shows that the same kind of implication holds for analysis of the contemporary period: the set of brewers does not constitute a single population.

Conclusion

¿From a substantive perspective, the most important feature of our proposed approach is its grounding of forms in identities. Our conception of identity, while consistent with much sociological thinking on the subject, expresses identity in terms of constraint. Potential identities are constraints that inform defaults, which tell what can be expected of an entity that possesses the identity. Social valuations convert potential identities into real identities, real in the sense that violations of the default assumptions has the price of lowered valuation.

Making these ideas sufficiently precise to serve as a basis for theory and research on the evolution of forms requires careful specification of defaults. Our approach develops a formal language for constraints and defaults using formal semantics. This strategy imposes obvious costs: gaining intuition about the approach requires close study of the somewhat dense formalism. We think that the approach has potential benefits that more than outweigh the costs. It provides a unified treatment of issues that heretofore have been analyzed in isolation: organizational identity and organizational form. It also improves the clarity of the formulation and allows checks for consistency among its various facets.

References

- Allmendinger, Jutta and J. Richard Hackman. 1996. "Organizations in Changing Environments: The Case of East German Symphony Orchestras." Administrative Science Quarterly 41:337-69.
- Barnett William P. and Glenn R. Carroll. 1995. "Modeling Internal Organizational Change." Annual Review of Sociology 21:217-36.
- Barnett William P. and Morten Hansen. 1996. "The Red Queen in Organizational Evolution." Strategic Management Journal 17:139-57.
- Barron, David N. 1999. "The Structuring of Organizational Populations." American Sociological Review 64:421-45.
- Barron, David N., Elizabeth West, and Michael T. Hannan. 1994. "A Time to Grow and a Time to Die: Growth and Mortality of Credit Unions in New York, 1914-1990." American Journal of Sociology 100:381-421.
- Barwise, Jon and John Perry. 1999. Situations and Attitudes. Stanford: CSLI Publications.
- Becker, Howard S. 1982. Art Worlds. Berkeley: University of California Press.
- Blau, Judith R. 1995. "Art Museums." Pp. 87-120 in Organizations in Industry: Strategy, Structure, and Selection, edited by G. R. Carroll and M. T. Hannan. New York: Oxford University Press.
- Boone, Christophe, Vera Bröcheler, and Glenn R. Carroll. 2000. "Custom Service: Applications and Tests of Resource Partitioning Theory among Dutch Auditing Firms from 1896 to 1992." Organization Studies, forthcoming.
- Burt, Ronald S. 1993. Structural Holes: The Social Structure of Competition. Cambridge: Harvard University Press.
- Carroll, Glenn R. and Michael T. Hannan. 2000. The Demography of Corporations and Industries. Princeton: Princeton University Press.
- Carroll, Glenn R. and Anand Swaminathan. 1992. "The Organizational Ecology of Strategic Groups in the American Brewing Industry From 1975 to 1990." Industrial and Corporate Change 1:65-97.
- DiMaggio, Paul J. 1986. "Structural Analysis of Organizational Fields: A Blockmodel Approach." Pp. 355-70 in *Research in Organizational Behavior* Vol. 8, edited by B. Staw and L. Cummings. Greenwich Conn.: JAI Press.
- ----- 1997. "Culture and Cognition." Annual Review of Sociology 23: 263-287.
- Dobrev, Stanislav D. 1999. "Revisiting Organizational Legitimation: Cognitive Diffusion and Sociopolitical Factors in the Evolution of Bulgarian Newspaper Enterprises, 1846-1992." Organization Studies, in press.

- Hannan, Michael T. 1998. "Rethinking Age Dependence in Organizational Mortality: Logical Formalizations." American Journal of Sociology 104:85-123.
- Hannan, Michael T. and Glenn R. Carroll. 1992. The Dynamics of Organizational Populations: Density, Legitimation, and Competition. New York: Oxford University Press.
- Hannan, Michael T. and John Freeman. 1977. "The Population Ecology of Organizations." American Journal of Sociology 82:929-64.
- ------ 1984. "Structural Inertia and Organizational Change." American Sociological Review 49:149-64.
- ------ 1986. "Where Do Organizational Forms Come From?" Sociological Forum 1:50-72.
- ------ 1988. "The Ecology of Organizational Mortality: American Labor Unions, 1836-1985." American Journal of Sociology 94:25-52.
- Hawley, Amos H. 1968. "Human Ecology." In David Sills (ed.) International Encyclopedia of Social Sciences. New York: Macmillan.
- Jepperson, Ronald L. 1991. "Institutions, Institutional Effects, and Institutionalism." Pp. 143-63 in *The New Institutionalism in Organizational Analysis*, edited by W.W. Powell and P.J. DiMaggio. Chicago: University of Chicago Press.
- Kalleberg, Arne, David Knoke, Peter V. Marsden, and Joe L. Spaeth. 1996. Organizations in America: Analyzing Their Structures and Human Resource Practices. Thousand Oaks, Calif.: Sage.
- Land, Kenneth C., Walter R. Davis, and Judith R. Blau. 1994. "Organizing the Boys of Summer: The Evolution of U.S. Minor League Baseball, 1883-1990." American Journal of Sociology 100:781-813.
- Meyer, John W. and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." American Journal of Sociology 83:340-63.
- Minkoff, Debra C. 1999. "Bending With the Wind: Change and Adaptation for Women's and Racial Minority Organizations." American Journal of Sociology 104:1666-1703.
- Olzak, Susan and Elizabeth West. 1991. "Ethnic Conflicts and the Rise and Fall of Ethnic Newspapers." *American Sociological Review* 56:458-74.
- Park, Douglas and Joel M. Podolny 1999. "The Competitive Dynamics of Status and Niche Width: U.S. Investment Banking, 1920–1950." Unpublished paper. Stanford Graduate School of Business.
- Péli, Gábor, László Pólos, and Michael T. Hannan. 2000. "Back to Inertia: Theoretical Implications of Alternative Styles of Logical Formalization." Sociological Theory, forthcoming.

- Peterson, Richard A. 1997. Creating Country Music. Chicago: University of Chicago Press.
- Podolny, Joel M. and Karen L. Page. 1998. "Network Forms of Organization." Annual Review of Sociology 24:57-76.
- Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan. 1996. "Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984-1991." American Journal of Sociology 102:659-89.
- Pólos, László. 1999. "On Identity." Applied Logic Laboratory, University of Amsterdam.
- Pólos, László, Michael T. Hannan, and Jaap Kamps. 1999. "Aging By Default: Building Theories From Fragmentary Knowledge." Pp. 207-19 in Proceedings of the Fourth Dutch-German Workshop on Non-monotonic Reasoning Techniques and Their Applications, edited by H. Rott, C. Albert, G. Brewka, and C. Witteveen. Amsterdam: ILCC Scientific Publications.
- Pólos, László, Michael T. Hannan, and Glenn R. Carroll. 1999. "The Language of Social Codes." Technical Report, Rotterdam School of Management.
- Russell, Raymond. 1995. Utopia in Zion. Albany: State University of New York Press.
- Sørensen, Jesper B. 2000. "The Ecology of Organizational Demography: Tenure Distributions and Organizational Competition." Industrial and Corporate Change 8:713-44.
- Swaminathan, Anand. 1995. "The Proliferation of Specialist Organizations in the American Wine Industry, 1941-1990." Administrative Science Quarterly 40:653-80.
- Tilly, Charles. 1986. The Contentious French. Cambridge: Harvard University Press.
- Weber, Max. [1924] 1947. The Theory of Social and Economic Organization, edited by A. H. Henderson and T. Parsons. Glencoe, Ill.: Free Press.
- White, Harrison C. 1992. Identity and Control: A Structural Theory of Social Action. Princeton: Princeton University Press.
- Williamson, Oliver E. 1975. Markets and Hierarchies. New York: Free Press.
- Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Legitimacy Discount" American Journal of Sociology 104:1398-1438.

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES RESEARCH IN MANAGEMENT

Publications in the Report Series Research^{*} in Management

Impact of the Employee Communication and Perceived External Prestige on Organizational Identification Ale Smidts, Cees B.M. van Riel & Ad Th.H. Pruyn ERS-2000-01-MKT

Critical Complexities, from marginal paradigms to learning networks Slawomir Magala ERS-2000-02-ORG

Forecasting Market Shares from Models for Sales Dennis Fok & Philip Hans Franses ERS-2000-03-MKT

A Greedy Heuristic for a Three-Level Multi-Period Single-Sourcing Problem H. Edwin Romeijn & Dolores Romero Morales ERS-2000-04-LIS

Integer Constraints for Train Series Connections Rob A. Zuidwijk & Leo G. Kroon ERS-2000-05-LIS

Competitive Exception Learning Using Fuzzy Frequency Distribution W-M. van den Bergh & J. van den Berg ERS-2000-06-LIS

Start-Up Capital: Differences Between Male and Female Entrepreneurs, 'Does Gender Matter?' Ingrid Verheul & Roy Thurik ERS-2000-07-STR

The Effect of Relational Constructs on Relationship Performance: Does Duration Matter? Peter C. Verhoef, Philip Hans Franses & Janny C. Hoekstra ERS-2000-08-MKT

Marketing Cooperatives and Financial Structure: a Transaction Costs Economics Analysis George W.J. Hendrikse & Cees P. Veerman ERS-2000-09-ORG

- ^{*} ERIM Research Programs:
 - LIS Business Processes, Logistics and Information Systems
 - ORG Organizing for Performance
 - MKT Decision Making in Marketing Management
 - F&A Financial Decision Making and Accounting
 - STR Strategic Renewal and the Dynamics of Firms, Networks and Industries

A Marketing Co-operative as a System of Attributes: A case study of VTN/The Greenery International BV, Jos Bijman, George Hendrikse & Cees Veerman ERS-2000-10-ORG

Evaluating Style Analysis Frans A. De Roon, Theo E. Nijman & Jenke R. Ter Horst ERS-2000-11-F&A

From Skews to a Skewed-t: Modelling option-implied returns by a skewed Student-t Cyriel de Jong & Ronald Huisman ERS-2000-12-F&A

Marketing Co-operatives: An Incomplete Contracting Perspective George W.J. Hendrikse & Cees P. Veerman ERS-2000-13– ORG

Models and Algorithms for Integration of Vehicle and Crew Scheduling Richard Freling, Dennis Huisman & Albert P.M. Wagelmans ERS-2000-14-LIS

Ownership Structure in Agrifood Chains: The Marketing Cooperative George W.J. Hendrikse & W.J.J. (Jos) Bijman ERS-2000-15-ORG

Managing Knowledge in a Distributed Decision Making Context: The Way Forward for Decision Support Systems Sajda Qureshi & Vlatka Hlupic ERS-2000-16-LIS

Organizational Change and Vested Interests George W.J. Hendrikse ERS-2000-17-ORG

Strategies, Uncertainty and Performance of Small Business Startups Marco van Gelderen, Michael Frese & Roy Thurik ERS-2000-18-STR

Creation of Managerial Capabilities through Managerial Knowledge Integration: a Competence-Based Perspective Frans A.J. van den Bosch & Raymond van Wijk ERS-2000-19-STR

Adaptiveness in Virtual Teams: Organisational Challenges and Research Direction Sajda Qureshi & Doug Vogel ERS-2000-20-LIS

Currency Hedging for International Stock Portfolios: A General Approach Frans A. de Roon, Theo E. Nijman & Bas J.M. Werker ERS-2000-21-F&A

Transition Processes towards Internal Networks: Differential Paces of Change and Effects on Knowledge Flows at Rabobank Group Raymond A. van Wijk & Frans A.J. van den Bosch ERS-2000-22-STR

Assessment of Sustainable Development: a Novel Approach using Fuzzy Set Theory A.M.G. Cornelissen, J. van den Berg, W.J. Koops, M. Grossman & H.M.J. Udo ERS-2000-23-LIS Creating the N-Form Corporation as a Managerial Competence Raymond vanWijk & Frans A.J. van den Bosch ERS-2000-24-STR

Competition and Market Dynamics on the Russian Deposits Market Piet-Hein Admiraal & Martin A. Carree ERS-2000-25-STR

Interest and Hazard Rates of Russian Saving Banks Martin A. Carree ERS-2000-26-STR

The Evolution of the Russian Saving Bank Sector during the Transition Era Martin A. Carree ERS-2000-27-STR

Is Polder-Type Governance Good for You? Laissez-Faire Intervention, Wage Restraint, And Dutch Steel Hans Schenk ERS-2000-28-ORG