Typecasting and Legitimation:
A Formal Theory*

Greta Hsu
University of California, Davis

Michael T. Hannan
Stanford University

László Pólos
Durham University

January 21, 2009

Abstract
We develop a unifying framework to integrate two of economic sociology’s theory fragments on categorization: typecasting and form emergence. Typecasting is a producer-level theory that considers the consequences producers face for specializing versus spanning across category boundaries. Form emergence considers the evolution of categories and how the attributes of producers entering a category shape its likelihood of gaining legitimacy among relevant audiences. Both theory fragments emerge from the processes audiences use to assign category memberships to producers. In this paper, we develop this common foundation and clearly outline the arguments that lead to central implications of each theory. We formalize these arguments using modal expressions to represent key categorization processes and the theory-building framework developed by Hannan, Pólos, and Carroll (2007).

*We thank Glenn Carroll for comments and the Stanford Graduate School of Business and Durham Business School for financial support.
Categorization in market contexts has attracted considerable interest in recent years, spurred in large part by Zuckerman's (1999) seminal work in capital markets. Empirical work on this subject covers a range of topics. One branch treats category emergence, proliferation, and erosion (Carroll and Swaminathan 2000; Ruef 2000; Rao, Monin, and Durand 2005; Bogaert, Boone, and Carroll 2006; Pontikes 2008). A second addresses the consequences of categorical positions and category structures for individual producers (Zuckerman and Kim 2003; Hsu, 2006; Negro, Hannan, and Rao 2008). A third examines the role of audience members in structuring understanding of categories (Boone, Declerck, Rao, and Van Den Buys 2008; Koçak 2008; Koçak, Hannan, and Hsu 2009).

This paper focuses on two theory fragments, typecasting and form emergence that exemplify the different emphases in research approaches. Typecasting theory focuses on well-established categories and considers the implications for individual producers of specializing in one category versus generalizing across categorical boundaries (Zuckerman, Kim, Ukanwa, and von Rittman 2003; Hsu, Hannan, and Koçak 2009). Research suggests that audiences have an easier time making sense of specialists but that a clear association with a single category restricts the range of future opportunities.

Form-emergence theory considers how the attributes of the producers associated with an emerging category shape its likelihood of gaining legitimacy among relevant audiences (McKendrick and Carroll 2001; McKendrick, Jaffee, Carroll, and Khessina 2003). Work in this area finds that a category is more likely to become a well-established form when new entrants are focused on the activities with the category.

These theory fragments have progressed largely independently of one another. This is not surprising given the differences in levels of analysis and key outcomes. Yet, they are clearly conceptually connected. Both address the positioning of producers in a space of categories and the effect of such positions on an audience's understandings. In this paper, we flesh out these connections to clarify the processes that lie at heart of theories of categorization. In particular, we demonstrate that a common foundation, a theory of partiality in memberships, can unify the key claims of both theory fragments.

Behind the scenes we use the formal theory-building tools and framework developed by Hannan, Pólos, and Carroll (2007) and extended by Pólos, Hannan, and Hsu (2009). The key constructions are modal models that allow for subtle formalization of key sociological concepts such as legitimation, identity, and social form as consequences of the beliefs held by relevant audiences. As we aim to illustrate, this approach to theory building has value for producing coherent, integrative sociological theories. However, in the interest of making the argument accessible we have placed most of the technical details, including
the formulas expressing the definitions, postulates, and theorems as well as the proofs, in the Appendix.

We begin with an overview of the conceptual approach to market categorization we adopt in this paper. We then highlight key concepts from recent theoretical work by Hannan et al. (2007) on the emergence categories and forms. We extend this theory to develop a theorem that fits the typecasting imagery developed by Zuckerman and colleagues. Then, with a few additional considerations, we establish that McKendrick and Carroll’s arguments regarding form emergence also follow from the framework.

**Categorization in Markets**

Recent interest in the topic of market categorization stems from dissatisfaction with traditional approaches to conceptualizing the notion of form (as applied to organizations and other kinds of producers). In the case of organizations, scholars generally agree that form refers to “those characteristics of an organization that identify it as a distinct entity and, at the same time, classify it as a member of a group of similar organizations” (Romanelli 1991: 81-82). Yet, there has been considerable disagreement about how to approach the task of identifying the presence and boundaries of a form in a given domain (Hannan and Freeman 1986; Romanelli 1991; Pólos, Hannan, and Carroll 2002). A significant body of research identifies forms by looking for common patterns of features, suggesting that forms can be assessed in purely objective terms. Other research looks to boundary-creating processes such as social network ties and personnel flows to understand form distinctions.

Yet, there is growing consensus that such approaches lose sight of the importance of the social meanings and interpretations of contemporaneous audiences in the specification of forms for organizations and other kinds of producers in market. How can researchers ensure that the forms they study empirically actually represent instances of meaningful social units? An audience-based theory of categories (an antecedent to forms) seeks to provide a resolution to such concerns.

This approach emphasizes modeling what social agents perceive when they “see” a producer and they expect of the bearers of the identity (Pólos et al. 2002; Hannan et al. 2007). In this approach, categories are a type of collective identity; they involve a typification of commonality where audience members recognize similarities among different producers and come to regard them as members of a common set. In some cases, categories become forms—that is, they become collective identities that are highly legitimated or taken-for-granted by members.
of an audience.

This formulation explicitly addresses the messy nature of socially constructed categories. Emphasizing what audience members see relies on the use of conceptual tools for representing the degree to which a producer is viewed by an audience member as fitting into a category. Such a conceptualization allows modeling of partiality in category memberships and the impact of partiality on category dynamics (Hannan 2008). This approach also provides a natural way to address the theoretical standing of a population (a set of entities with a common form) at its inception (Pólos et al. 2002). In reframing forms in terms of identities, this approach holds that each new socially enforced identity initiates a potential population. Whether a collective identity progresses to the status of a form depends on whether other entities come to share this identity and adopt features that makes them fit closely to audience expectations.

A broad range of recent empirical work has adopted important elements of this new approach to conceptualizing categories and forms (but not necessarily cast in a model of potentially partial memberships). We provide a brief (and by no means comprehensive) overview of key lines of research here. One line uses ideas about identity to enhance understanding of form emergence. For example, Martin Ruef (2000) demonstrates for the healthcare domain that the distribution of existing organizations in identity space affects the likelihood new forms with similar identities will emerge. A key implication is that the way a novel organizational form fits within audience members’ preexisting understandings shapes the likelihood of its acceptance and legitimation. McKendrick, Carroll, and colleagues study a complementary issue: how the identities of the producers associated with a category shape the likelihood it will cohere into a highly legitimated form (McKendrick and Carroll 2001; McKendrick et al. 2003). They find a category is more likely to achieve highly legitimated status when its members are highly focused on the activities of the category (as compared to situations in which members are active in multiple categories).

Another line of research draws attention to the role of perceptions and actions of external audiences in shaping category evolution. For example, Rao, Monin, and Durand’s (2003) study of the nouvelle cuisine identity movement in French gastronomy finds that greater sociopolitical legitimacy of movement activities and theorization about nouvelle cuisine by culinary journalists and critics influenced the propensity of French chefs to “defect” from classical to nouvelle cuisine.

Research has also investigated how external audiences’ perceptions shape critical and commercial success and failure for individual producers. In his research on financial markets, Zuckerman (1999) finds that firms that fail to establish themselves clearly as members of a legitimated form in the eyes of finan-
cial analysts (by diversifying across the categories that they use in covering the market) are less likely to receive analyst coverage. This reduces their attractiveness to investors and impairs their stock market returns accordingly. And in a study of Hollywood film projects, Hsu (2006) finds that audiences have trouble making sense of films that span established genres. Genre-spanners fit poorly with audience tastes and have low appeal. This line of research suggests audience members exert strong constraints on individual producers to conform to the expectations inherent in the identities of the category with which they are associated.

The strength of category-related expectations (and thus the penalties associated with violating them) evolves over time with changes in the composition of a category. For example, Rao, Monin, and Durand (2005) find that penalties imposed on chefs who borrow elements from both classical and nouvelle cuisine weaken as the number of chefs who similarly straddle these category boundaries rises. This suggests that, as chefs increasingly engage in category-spanning activities, audience members’ categorical understandings shift as well. High levels of category spanning appear to cloud audience members’ beliefs about what it means to a member of a category. Negro, Hannan, and Rao (2008) find evidence of this kind of pattern in their study of critical reactions to elite Italian wines: as category spanning becomes widespread, the rewards for membership in the category decline and the penalties for category spanning weaken.

Researchers have also examined the changes in the repercussions of category spanning as the identity of individual producers change. For example, research on typecasting by Zuckerman and colleagues (2003) finds that conforming to the expectations of a single category increases the likelihood that a new producer will gain attention from relevant audiences. However, such a simple and clear identity restricts future opportunities outside of a producer’s initial category. So an identity that spans multiple categories can prove more beneficial for experienced producers in allowing a wider range of activities to be acceptable to relevant audiences.

These various lines of research share a common conceptualization of categories as collective identities constructed over time through the actions of interested social agents. As such, they can be viewed as fragments of a larger research program. Learning how these separate fragments fit together—or what changes must be made to make them fit—is vital to the continued development of this area of research. Such theoretical integration allows for checks on the consistency of different lines of reasoning as well as the identification of core ideas transferrable across diverse areas of theorization.

In this paper, we integrate theories of typecasting and form emergence. These theory fragments focus on causal factors and outcomes at different analytical
levels (i.e., the individual producer versus the category) as well as on categories at different stages of development. Typcasting theory focuses on highly legitimated categories (forms) and considers how the existing identity of a producer shapes its future opportunities vis-à-vis market categories. Theories of form emergence focus on earlier stages of the categorization process and consider the factors that shape the likelihood a category will attain status of a highly legitimated form. By identifying and integrating principles common to these very different but related fragments, we aim to establish a stronger foundation for future theorization and empirical work on market categorization.

To structure our theory integration, we formalize the core arguments and assumptions within each theory. A key advantage of formalization over natural-language theory development is that it pushes researchers to trace arguments through precise chains of reasoning. Theoretical arguments in natural language naturally contain both ambiguity regarding the meaning of key concepts and the relationships linking them (Hannan et al. 2007). By adopting a formal approach, we aim to minimize ambiguity and carefully reconstruct arguments central to both typcasting and form emergence in a way that highlights the specific and common underlying processes at work as well as the assumptions that might be necessary for each theory to hold. To be sure, there is a price to be paid. Formalization requires reliance on explicit simplifying assumptions, which might not fit real situations.

In their theory of categories and form, Hannan et al. 2007 establish a formal foundation on which we build. We begin by sketching the basic framework, which highlights the role of the audience in constructing categories and assigning membership to them.

**Modal Models for Legitimation**

Category emergence theory considers a domain as consisting of a dual role structure: producer (agents who make offerings in the domain) and audience member (agent who evaluate offerings and potentially reward producers of offerings that they find appealing), and a language that spells out the meanings attached to the roles.

The basic linguistic objects are labels that audience members apply to producers. Applying labels (such as “research university,” “semiconductor manufacturer,” or “dramatic actor”) facilitates cognitive processing of and communication regarding producers (Hsu and Hannan 2005). Labeling thus appears to be an important step in the social construction of categories. We begin our formal story with a labeling function, which maps audience members, producers,
and time points to (the powerset of) the set of available labels.

Labels are often paired with schemas that articulate what determines a label's applicability to a producer. Schemas provide abstract models or representations of the feature values that audience members regard as consistent with a given label. In other words, a schema establishes the meaning (or intension) of a label.

In formal terms, schemas for labels are sets of formulas that pick out a set of relevant features (or relations). They distinguish the values of those features (or relations) that are consistent with membership in a label from those that are not. An audience member's schema for a label points to the subsets of the values of the relevant features that contain the schema-conforming feature values. For example, the label "microbrewerer" might be paired with a schema including features such as a small-scale operation, traditional hand-crafted methods of production, and use of traditional ingredients.

Although agents can apply unique labels to specific producers, we are concerned with labels attached to clusters of producers that audience members believe to be members of a similar set. We refer to a pair consisting of a label and a schema for a similarity cluster as a type.¹

Following several major lines of work in cognitive psychology and cognitive science, we assume that assessments of producers' membership in a type can be partial, a matter of degree (Hannan 2008). Based on the configurations of features in a schema, an audience member regards producers with certain configurations as full-fledged members of a type, others as having a moderate or low standing as a member, and still others as completely outside the type boundary. The degree to which the producer's characteristics fit a schema is reflected in a grade of membership (GoM) function.

Types can have positive, neutral, or negative valuation. For issues related to typecasting, the interesting case concerns positive valuation. In such cases, greater fit with an audience member's schema yields greater intrinsic appeal (fit with tastes for offerings of that type). A type is positively valued when the expected intrinsic appeal of a producer's offering increases with the producer's grade of membership in the audience member's meaning of the label.

¹From the definition of a schema as a function, it follows that at most one schema can be paired with a label. Therefore, types for labels are unique.
Perception, Defaults, and Beliefs

An audience members’ experience with type members shapes the strength of her schema for the type. In some cases, an audience member might generally find type members to have a high GoM in the label—they generally display feature values that fit well the relevant schema. Such generic fit and low frequency of observed misfit causes audience members to come to take for granted that the behavior and structures of any bearers of the type label will be completely consistent with their schema for the type. Beliefs about schema conformity thereby become default assumptions of everyday life. This means the defaults are used to fill in the many gaps in perceptions that come about from incomplete information, unobservability, and ambiguity.

We use three modal operators—for perception, default, and belief—to analyze these issues (see Pólos et al. 2009). In logic, the term modality originally was used originally to refer to qualities of the truth of an expression, especially the possibility and necessity of a statement. The technical apparatus for analyzing logics with operators for possibility and necessity has been generalized to treat statements about an agent’s attitude toward an object or relation; and the term modality is now generally extended to include expressions of perceptions, beliefs, and valuations. We use this extended sense of modality.

We refer to an agent’s information state about a factual situation as a set of beliefs. Perceptions contribute to beliefs in an immediate way. What agents directly perceive updates the set of beliefs that they hold. Therefore the temporal order of perceptions matters: more recent perceptions replace older ones in case they conflict. But agents’ perceptions are generally partial, making some propositions true and others false, while leaving open the truth/falsity of others. As such partiality generates uncertainty, it is natural that mechanisms emerge that eliminate some of the gaps.

We propose that agents rely on schematic defaults to “fill in” missing facts when a relevant direct perception is lacking and an applicable default is available. That is, defaults shape beliefs only in the absence of current perception of the facts in question. And although beliefs based on taken-for-granted assumptions shape information states (and thus behavior), defaults are exposed to revision due to direct perceptions that conflict with the assumed facts.

Pólos et al. (2009) defined a model for the language containing these oper-

---

2DiMaggio (1997) discusses the sociological implications of research in cognitive science on a key related mechanism: automatic cognition—the implicit, automatic reliance on default assumptions about features embedded in schemas.
ators and provided their formal semantics. Their model was designed to satisfy the following constraints:

1. perception is partial at all time points;
2. beliefs must be grounded in either perception or taken-for-granted assumptions;
3. as seeing is believing, perception (at least temporarily) overrides earlier beliefs;
4. defaults shape beliefs (unless there is perceptual evidence to the contrary);
5. lasting beliefs develop if lasting taken-for-granted assumptions are not contradicted by perceptual evidence.

**Defaults and Induction**

Audience member’s perceptions of a producer’s fit to their schemas for a label are often partial. In some cases, an audience member sees (or treats as a default) only that a producer claims a label or that some other audience members (perhaps critics or another kind of gatekeepers) apply the label to the producer. Such situations offer the analytic leverage needed to define legitimation. The key issue is how many schema-consistent features an audience member needs to check (in terms of beliefs) before she assumes conformity with the schema for the unchecked features.

This idea can be represented in terms of a **test code**, a partial segment of a schema that an audience member uses to make inferences about fit to the rest of the schema on which she has no beliefs (perceptions or defaults). If the audience member believes that a producer “passes” the test, then she induces that the unperceived/non-default values of schema-relevant features also fit the schema. For example, an agent’s test code for the type “classical French restaurant” might consist of specific ingredients used in dishes and the naming and presentation of dishes (Rao et al. 2003). If a restaurant displays feature values consistent with the agent’s schema, then the agent will induce that other schema-relevant features, such as the role of the chef in the restaurant’s power structure and the organization of the kitchen, are consistent as well.

In technical terms, (believed) satisfaction of a test triggers the audience member to apply the default that the unperceived/non-default feature values also satisfy the schema.\(^3\)

---

\(^3\)We modify the definition of induction offered by Hannan et al. (2007, Def. 4.1), which holds
**Definition** (Induction from a test). An induction from a test is a situation in which an audience member's belief that (1) a producer bears a type label and (2) its feature values satisfy a test is enough to trigger the application of the default that the values of unchecked features (for which there is no prior belief to the contrary) also satisfy the schema.

If an audience member must check every relevant feature before assuming as a default that the rest of a producer's features match the relevant schema, then nothing is taken for granted. If only a small fraction of the relevant features must be checked (perhaps only a claim to the label), then defaults get used in a powerful way. These comparisons make the most sense when we consider the minimal test for an audience member-schema pair, the test that involves the smallest number of features.

The relative size of the minimal test for induction for fit to a schema relates directly to the degree of taken-for-grantedness of the label for the audience member. By size we mean the number of features that belong to the schema or test.

**Definition** (Taken for grantedness). The degree to which an audience member takes for granted that the untested feature values of a labeled producer conform to a schema for the label at a time point is the ratio of the size of the untested portion of the schema to size of the whole schema.

\[ g(l, x, t) = \frac{I - J}{I}, \]

where \( I \) denotes the number of features in agent y's schema for the label \( l \) at time point \( t \) and \( J \) denotes the number of features in the agent's minimal test code (given that \( l \) is a type for the audience member).

This definition sets taken for grantedness to zero if the audience member does not apply the label to the object or needs to see every (nonlabel) feature before making an induction (which is no induction at all); nothing is taken as satisfied by default. It sets taken for grantedness to one if applying the label by itself shifts the audience member to defaults about schema-conformity on all other relevant

that induction “fills in” all non-perceived feature values when a test is perceived to be satisfied. This overlooks the role of existing defaults. Recall that defaults are beliefs when there is no contrary perception. There does not appear to be any reason to think that audience members will override existing defaults based only the passing of a test on other features. So we refine the earlier conception in line with this intuition. That is, we propose that induction works on features about which the audience member has no belief (based either on perception or default).
features. In this case, the test on feature values is empty, \( f = 0 \); and the test is passed automatically whenever the label is applied.

We refer to a highly taken-for-granted type as a *concept*. That is, an audience member’s type is a concept if she treats conformity to her schemata for the type label as taken-for-granted for (nearly) all those producers/products to which she assigns the label.

**Incomplete Beliefs and Defaults in Typcasting**

Reliance on defaults about concept membership shapes how audience members regard the producers to whom they apply a label. We claim that defaults also play a key role in creating the typcasting dynamic that Zuckerman et al. (2003) highlight in their study of the careers of Hollywood film actors. As noted earlier, this research finds that actors who are strongly identified with a single type (genre) of work often find it difficult to obtain future work in other types difficult. Presumably, audience members assume that each type of work requires a distinct set of skills, so clear identification with one type of work implies that an actor lacks the skills necessary for others.

Our understanding of typcasting is that it depends on partiality of available information. Sometimes audience members have full information about the properties of some producer and can tell whether it fits one or another schema. In such situations, there is no reliance on typcasting—the agent relies on direct perception. But, when perception is incomplete, knowledge that a producer fits one type generally gets treated as evidence that it likely does not fit other types (with clashing schemas).

More generally, the typcasting dynamic suggests that the belief that a producer is a member of one type will (1) increase the producer’s appeal in exchanges of that type and (2) prevent acceptance of its membership in others. To build to that multiple-type case, we first need to consider how taken for grantedness affects the assignment of GoMs in cases in which test codes are satisfied but the audience member does not have a belief about schema satisfaction for some relevant features. In doing so, we contrast the role of concepts and mere (i.e., less taken-for-granted) types.

To simplify our formal story, we construct our arguments at the audience-member level. These results can be aggregated member by member to derive implications for an audience. In the interest of brevity, we do not develop these aggregate implications formally.

We treat simple situations in which fit to a schema can be assessed by simply counting matches and mismatches of values of relevant features to the schema.
(In more complex cases, there might be weights assigned to features such that mismatches reduce fit more when they occur on certain features or elements of schemas might be conditional, meaning that the value of one feature affects what is schema-conforming on another). We implement this restriction with the notion of a flat schema.

**Definition (Flat schema).** An audience member’s schema for a label is flat if audience members normally assign higher grades of membership to objects with more matches to the schema and fewer mismatches.

How can we represent the idea that audience members often lack complete beliefs about schema satisfaction? Because we want to make the argument general and we do not have any prior expectations about patterns, we develop a simple baseline probability model that allows us to compare situations that are alike on average. (We state the elements of the probability model as auxiliary postulates, which means that they are stated as analytical conveniences not as claims about the world.)

The first step defines a common probability over schema-relevant features that an audience member lacks a belief. Our baseline model holds that the schema-relevant features do not differ in the probability that an audience member lacks a belief about conformity to the schema. In other words, each audience member has available beliefs on a random sample of schema-relevant features.

**Auxiliary assumption 1** (Beliefs available for random samples of features). Beliefs about fit to a schema for a label are available at random for an audience member-producer pair in the sense that the probability that the audience member does not have a belief about the value of a feature is the same for all schema-relevant features.

The key intuition behind typecasting relies on a counterfactual: had the audience member had full information about two producers (who differ in their histories of prior labels and memberships), she would have no reason to prefer one to the other. According to the counterfactual, the audience member would generally regard the producers as having equal grade of membership in terms of satisfaction of her schema. To represent this counterfactual, we assume as the second element in the baseline probability model that the two producers being compared are equally likely to satisfy the audience member’s schema if the audience member had a positive belief about their values. We do so by assuming that the probability that a belief that an arbitrary feature conforms to the schema is the same for the two producers being compared. In other words, the producers are equivalent in expected-value terms.
Auxiliary assumption 2 (Common probability of schema-conforming beliefs). The probability that an audience member believes that one of a producer's feature values satisfies here schema for a label (conditional on a belief) is the same for all audience-member–producer pairs.

This probability model, when applied to flat schemas, implies a pattern that agrees with the core intuition about the constraints imposed by typecasting. We develop this implication for a simplified situation that makes the analysis tractable. The simplification considers situations in which each audience member has flat schemas for two labels, \( l \) and \( l' \), of the same length (\( I = I' \)),\(^4\) has minimal test codes for each schema, and the probabilities that the audience member has a belief about a \( l \)-schema-relevant feature and that beliefs indicate schema conformity are equal for the two triplets of producers, audience members, and time points.

Let \( \Phi[t, t'] \) indicate that the following conditions hold over the time interval beginning at \( t \) and ending at \( t' \):

1. \( l \) is a concept for the members of the audience over the interval \([t, t']\) and their schemas for it do not change over the interval:
2. \( l' \) is a type for the audience members at (at least) the end point of the interval;
3. the audience members’ schemas for \( l \) and \( l' \) are flat over the interval and have the same length \( I = I' \);
4. on average, beliefs about the \( l \)-relevant and \( l' \)-relevant feature values of all the producers in the domain are incomplete to the same degree for all audience members over the time interval and all producers fit the audience member’s schemas for \( l \) to the same degree within that interval.

Lemma 1 (Taken for grantedness and expected grades of membership). With random availability of beliefs for flat schemas with a common probability of forming a schema-conforming belief, audience members presumably assign higher grades of memberships (to producers) in the meaning of a label when the conformity with label is more taken for granted (under the conditions stated in \( \Phi \)).

\(^4\)If we allow \( I, I', J, \) and \( J' \) to vary freely subject only to the constraint that \( (I - J)/I > (I' - J')/I' \) (the relevant condition for judging the degree of taken for grantedness), the implications appear to be indeterminate.
The proof of this lemma, along with those of other lemmas and theorems, can be found in the Appendix.

According to Lemma 1, an audience member presumably assigns a higher grade of membership to a producer when her minimal test for induction is smaller. Hence conformity with the schema is more taken for granted. This result has an immediate implication about the importance of concepts in situations of partial beliefs.

**Theorem 1** (Concepts versus types and grades of membership). *When audience members have partial observations on some type-relevant producer characteristics, they presumably assign higher grades of membership in the type to objects when the type is a concept, that is, highly taken for granted (under the conditions stated in $\Phi$).*

Because we focus on positively valued types, the argument behind Theorem 1 also implies a parallel difference in the intrinsic appeal of offerings for concepts versus mere types.

**Corollary 1** (Concepts versus types and appeal). *When audience members have partial observations on some type-relevant producer characteristics, they presumably find the offerings of members of the type more appealing when the type is a concept (under the conditions stated in $\Phi$).*

**Typcasting**

The argument made to this point, together with the behavior of the modalities, yields what we regard as a somewhat surprising implication. Consider the case in which an audience member first decides that a producer passes the minimal test code for a concept and does not display any observable violations of the schema for that concept. She then later finds that the same producer also passes the minimal test code for a clashing concept and does not display any observable violations of the schemata for that concept. What happens?

To provide a formal answer to this question, we first define clashes between schemas. To simplify what follows, we consider pairs of labels whose schemas clash (for an audience member) but only outside of their minimal test codes.

---

5Hannan et al. (2007) define schema clash indirectly with a meaning postulate (MP5.1) that presumes part of what we want to derive: “Normally, the higher a producer’s grade of membership in a type whose schema clashes with that of a focal category, the lower the producer’s grade of membership in the focal type.”
**Definition** (Code clash). Schemas for a pair of labels clash outside an agent’s minimal test codes for them if (1) the pair of minimal test codes can be jointly satisfied and (2) the pair of codes cannot be jointly satisfied.

As we thought about these issues, we first reasoned that a schema clash might block an audience member from applying defaults and would yield lowered GoMs in both concepts. But we recognized that this in not how defaults work (as we have modeled them). Once defaults are set, they have the status of facts (unless and until they are overridden by new perceptions). So, in the scenarios we are considering, the audience member treats all of the schema-relevant features as satisfying the schema and also treats the defaults as facts when considering membership in the clashing concept. The result is that the audience member decides that the producer does not fit well the focal (clashing) concept, and she does not alter her judgment of the producer’s typicality in the original concept. This conclusion fits our reading of the typecasting imagery.

This argument can be built on the argument behind Theorem 1. The key step in linking this argument to typecasting is constructing meaningful simplifying assumptions that allow us to capture the key insights.

Let \( \Psi[t, t'] \) indicate that the following conditions hold over the time interval beginning at \( t \) and ending at \( t' \):

1. code clash: the audience members’ schemas for \( l \) and \( l' \) clash outside their minimal test codes and schema clashes outnumber non-clashes from the perspective of all of the audience members;
2. each audience member applies the label \( l \) to the producer \( x \) and believes that the producer \( x \) passes her minimal test code for \( l \) over the relevant time interval;
3. each audience member applies the label \( l' \) to both producers at the later time point \( t' \);
4. each audience member either does not apply the label \( l \) to the producer \( x' \) over the relevant time interval: or does not believe that the \( x' \) passes her minimal test code for \( l \) over the interval.

**Theorem 2** (Typecasting). When two concepts with schemas that clash only outside audience members’ minimal tests for them, membership in one concept at an earlier point in time (1) yields a higher fit to that schema at subsequent times but (2) reduces the fit to the other schema at a later point in time when the audience members do not generally have beliefs about a producer’s conformity to schema on all relevant features (under the conditions stated in \( \Psi \)).
In the case of positively valued types, the pattern claimed by Zuckerman and collaborators immediately follows.

**Corollary 2** (Zuckerman et al. 2003). *In the case of two concepts with schemas that clash only outside audience members’ tests for them, membership in one at an earlier point in time presumably (1) enhances the intrinsic appeal of the producer’s offering in the first concept but (2) reduces the appeal of its offering in the other at a later point in time when the audience members do not generally have beliefs about a producer’s conformity to schema on all relevant features (under the conditions stated in $\Psi$).*

Theorem 2 and Corollary 2 highlight both the benefits and the drawbacks of passing the minimal test code of a highly taken-for-granted type. On the one hand, reliance on defaults means that audience members will assign a high grade of membership to the producer and find its offerings to have high intrinsic appeal. But this restricts the producer’s ability to demonstrate fit with a clashing type in the future. Audience members will rely on prior defaults in the case of partial perception and immediately assume a poor fit with the schema and their tastes for offerings of the clashing type.

**Contrast and Taken for Grantedness**

Our rendering of the typecasting argument has potentially broad implications. Exploring them requires attention to related processes underlying taken for grantedness (legitimation).

The original theory of density-dependent legitimation held that growth in the number of producers associated with a category increases its taken for grantedness (Hannan and Freeman 1989). This basic formulation, however, did not address the idea that different producers might contribute differentially to the taken for grantedness of a category. To incorporate this possibility and generalize the theory, Hannan et al. (2007) shifted attention from density to population contrast. Contrast refers to the average grade of membership among those with positive GoM (for an audience member). In other words, a high contrast condition is one in which approximates a binary distinction (full membership versus non-membership).

Categories with high contrast stand out sharply against the background, increasing the likelihood that audience members see the cluster of producers in similar ways, which eases the gaining of consensus among audience members about the meaning of the category (intensional consensus). Hannan et al. (2007) posit that the level of legitimation of a label in the audience as a whole increases
monotonically with the level of intensional consensus about the label. Then it follows that legitimation (the average level of taken for grantedness over all pairs of producers and audience members) increases with average contrast in the audience.

Here we focus on another path, one that links legitimation to contrast at the level of the individual audience member. The contrast of a type for an audience member is the average of the nonzero grades of membership that the audience member assigns to the objects to which he assigns the type label.

For an individual audience member, high type contrast means that the producers to whom the audience member assigns the type label generally fit the concept schema well (because high contrast means that the audience member assigns either high or very low GoM in the meaning of the label to the objects in the domain). Cases of poor fit to schemas will generally be viewed as exceptions to the general rule. Such generic fit causes the audience member to come to take for granted that any producers to which she applies the type label will have schema-consistent features. The probability that beliefs about schema conformity will become defaults therefore increases with contrast.

Postulate. A type's expected taken-for-grantedness for an audience member normally increases (with some delay) monotonically with its contrast.

With these notions in hand, we return to our main substantive focus, the effects of typecasting.

**De-Novo and De-Alio Entrants**

The intuition underlying our rendition of the typecasting theorem—that assumptions regarding a producer's membership in one concept constrain beliefs about fit with others—can be usefully extended to shed light on other aspects of the dynamics of types and concepts. In this section, we demonstrate this by considering how this process of induction and typecasting relates to key findings on form emergence. We continue here to build the model at the level of the audience member.

In their seminal study of the disk-array producers, McKendrick and Carroll (2001) found that a label for an emerging category gains more legitimation from de-novo entrants (those with no prior history for the audience member) than from de-alio entrants (already existing producers who the audience member regards as diversifying from some other categories). They reasoned that audience members perceive de-novo entrants as more focused on the activities associated with the label than their diversified counterparts and thus contribute more
to audience members’ understandings of what it means to be a type member. This finding had a serious impact on the thinking of organizational theorists. However, so far it has not been integrated into the formal theoretical framework of categories.

To capture McKendrick and Carroll’s core insight, we focus on the case of de-alian entrants with clashing memberships. Just as the typecasting dynamic rested on the assumption that distinct genres correspond to distinct skill sets, the story about de-novo and de-alian entrants rests on the idea that the different producer types are associated with clashing schemas. A producer has de-novo status in a label to an audience member if she applies the label at the time point and has not previously applied any label to it. A producer has de-alian-clashing membership in a label if the audience member applies the label at the time point and also continues to apply a label assigned earlier in a clashing concept and believes that the producer passes the minimal test for the clashing concept.

When an audience member assesses the fit of a de-alian entrant from a clashing concept to a focal type, the process of induction that drives the typecasting dynamic (as stated in Theorem 2) is also at work. Membership in the clashing concept reduces the fit of the de-alian entrant in the focal type. This puts the de-alian entrant at a disadvantage as compared with a comparable de-novo entrant (when there is random availability of beliefs and a common probability of forming a schema-conforming belief for each producer).

**Theorem 3** (De-novo entrants have higher expected grade of membership). *De-novo entrants presumably have higher expected grades of membership in audience members’ types than do de-alian entrants with memberships in clashing concepts (under the conditions stated in $\Psi$).*

This grade of membership disadvantage for de-alian entrants also results in a disadvantage in terms of the appeal of their offerings to audience members.

**Corollary 3** (De-novo entrants have higher expected appeal). *The offerings of de-novo entrants presumably have higher intrinsic appeal than do those of de-alian entrants with memberships in clashing concepts (under the conditions stated in $\Psi$).*

McKendrick and Carroll (2001) suggest that the extent to which audience members perceive a set of entrants as having a type focus contributes to the taken for grantedness of the type. This idea of perceptual focus can be analyzed in terms of contrast, as defined in the previous section. Because de-novo entrants have a higher expected grade of membership in an audience member’s concept, they naturally contribute more to type contrast.
Lemma 2 (De-novo entrants contribute more to expected contrast). *De-novo entrants presumably contribute more (with some delay) to the contrast of audience member’s type than de-alio entrants from clashing concepts (under the conditions stated in \( \Psi \)).*

The core insight of the de-novo/de-alio story is an implication of the preceding argument. The implication follows from the foregoing lemma and Postulate 1. Because Lemma 2 yields a difference in the expected contrasts at the end of the period \([t, t']\) and Postulate 1 states a delayed effect of contrast on expected taken for grantedness, we express the desired theorem as holding for expected taken for grantedness at some time at or after \( t' \) (reflecting the possible delay in the effect of contrast on taken for grantedness).

Theorem 4 (McKendrick–Carroll 2001). *The expected (delayed) contribution of a set of entrants to the taken for grantedness of a type for an audience member presumably is higher for de-novo than de-alio entrants (under the conditions stated in \( \Psi \)).*

Discussion

A key aim of this paper was determine whether processes underlying typecasting and form emergence could be usefully integrated to explain dynamics central to both theory fragments. We attempted this integration by extending a model of how audience members apply labels and assess fits of producers to schema to multiple concepts.

Our theory building suggests that induction of fit to schemas might be at the core of the issues considered by these two theories. Audience members often rely on defaults regarding schema-relevant features to define the concept memberships of producers in the face of partial information. Following Hannan et al. (2007), we proposed that the concept of taken-for-grantedness is linked to induction: what defines the taken-for-grantedness of a type is the degree to which audience members automatically fill in defaults for producers who exhibit some minimal criteria for membership for that type. In the case of a single category, one implication is that audience members presumably assign a higher grade of membership to a producer for a concept versus mere type. As a result, they also presumably find the producer more appealing in the former case.

Generalizing this story to the case of multiple memberships sets the foundation for modeling the typecasting dynamic. When audience members associate a producer with one concept at an earlier point in time, the (believed) fit of that producer in that concept will be enhanced, but its fit in other types will be reduced. A similar process of induction can explain the findings by McKendrick
and Carroll (2001) on form emergence. De-alie entrants are generally believed to have worse fit with a concept schema because they already belong to a clashing concept. Hence de-alie entrants get assigned lower GoMs in a concept than their de-novo counterparts, and they contribute less to the taken for granted-ness of a type.

The intuition underlying both theories relies on a counterfactual: with complete information about the producers under comparison, audience members would assign equal grades of membership in the relevant types. Both theories suggest that, by relying on defaults to cope with incomplete information, audience members become biased in systematic ways in their perceptions. Capturing this intuition formally required the introduction of auxiliary assumptions. In particular, we stipulated random availability of beliefs and a common probability of forming a schema-conforming belief for each of the producers under comparison.

Our formalization also shows the importance of the relationship between categories. Without some clash between the codes of the relevant categories, there would be no reason to expect such systematic bias in beliefs. Zuckerman et al. (2003) argue that lay theories of skill provide the foundation for typecasting dynamics to emerge. For classification structures (such as types and concepts) to be restrictive, audience members must believe that the features or skills necessary to be a member of one differ from those required of others. They must also believe that significant type-specific investments are necessary to acquire the necessary features of each (meaning that there is a so-called principle of allocation tradeoff). These assumptions are what lead audience members to presume that a producer who has demonstrated fit with one category must lack the features necessary for others.

Clearly, a formal translation of this insight would require attention to the relationship between the membership schemas for different categories. But to capture the core intuition, we needed to formalize this in a way that audience assumptions (rather than their direct perceptions) would drive the main dynamic. More specifically, we needed audience members to assume incompatibility of membership for producers who might actually possess features consistent with the types in question. Thus, clash in codes between concepts should matter when some features are unobserved. Our treatment of induction provides a sensible way to model this by allowing us to focus on code clash outside of an audience member’s minimal test criteria.

This thought process also led us to think about the role code clashes play in McKendrick and Carroll’s theory of form emergence. Their natural language theorization did not explicitly consider the relationship between the categories de-alie producers span. Yet, our framework suggests that clashes between spanned
categories must also be present for bias against de-alio producers to emerge. We find that it is the presumption of code clashes outside the minimal test criteria that also produces a key intuition underlying this theory.

We developed our model at the level of the audience member by considering the audience member’s application of labels and assessment of fit to his/her own schema for the label. As we noted above, these results have implications at the level of the audience as a whole. When the members of an audience come to substantial agreement about the meaning of a set of labels, they will generally make similar assessments of fit of producers to schemata and engage in induction based on similar observations. Hence, the line of argument we presented in this paper applies mutatis mutandis to a comparison of categories and forms, the audience-level parallels of types and concepts.

A core tool used to develop our findings is use of modal models. Following Hannan et al. (2007), we proposed that the default operator holds when audience members encounter producers who pass their test for a type. In the case of highly taken-for-granted types, this test code is very small and processes of induction are common. However, for less taken-for-granted types, a large test means the audience member will not assume much in terms of conformity with the concept schema. In such cases, the perception operator largely applies, and only partial membership will be assigned when an audience member lacks a belief about some schema-relevant features. Together, these modalities capture in a very specific way what seems distinctive about both membership in highly legitimated types and membership in multiple market types and concepts (categories and forms).

More broadly, this integration project pushed us to identify, articulate, and check the consistency of principles that likely apply to other identity-based theories of market categorization. Theories about market categorization often concern counterfactuals—the way an agent’s position in the market’s role structure is decoupled from its actual features and exerts independent constraints on its opportunities (Zuckerman et al. 2003; Hsu, Hannan, and Koçak 2009). Our model building highlights how distinct theories conceptualize this issue in similar ways with regards to the constraints of classificatory memberships. In particular, we demonstrate the key role induction plays in producing these counterfactual dynamics. In doing so, we provide a path for articulating similar processes in theories of social structuration more generally. For example, similar inductive processes likely play a role in theories of status, reputation, and identity. Attention to this fundamental social process might provide a way for researchers in these diverse areas of research to better integrate their findings.

Our framework might also be used to develop new theoretical insights. To give one example, the basic typecasting dynamic we specify could be extended
to consider what happens when audience members gain greater familiarity with a domain and develop stronger beliefs about distinctions between type. At the audience level, this reflects increasing institutionalization of the domain. Our framework suggests that audience members should increasingly rely on defaults to induce memberships, resulting in more extreme assessments of grades of membership. As a result, the overall contrast of categories within a domain would increase.

Our study also draws attention to the need to consider relationships between codes for different market types in arguments about partiality in memberships. An important direction for future research in this area is to develop better understanding of how inter-category relationships shape market dynamics. Currently, theories of multiple memberships either do not explicitly consider relationships between categories or focus exclusively on oppositional categories. Yet, there is a richer set of inter-category relationships that one encounters in real-life markets. For example, some categories appear to overlap in expected features and seem complementary, while others do not directly clash but contain features that are largely unrelated. Our framework provides a concrete way to conceptualize relationships between category codes and provides an avenue for future research development on multiple-category memberships.
Appendix: Formal Details

The theory on which we build states (some) definitions, postulates, auxiliary assumptions, lemmas, and theorems in a nonmonotonic logic (Pólos and Hannan 2002, 2004). In formal terms, models of arguments are given in terms of sequences of intensions of open formulas. It contains a formal language to represent causal stories and defines a new kind of quantifier, denoted by $\forall$. Formulas quantified by $\forall$ state what is expected to “normally” be the case according to a causal story. The normal case is what we assume to be the case if we lack more specific information that overrules the default. The implications of a set of rules with exceptions, provisional theorems, are the logical consequences of a stage of a theory. Provisional theorems have a haphazard existence: what can be derived at one stage, might not be derivable in a later stage. So the status of a provisional theorem differs from that of a causal story. The syntax of the second language codes this difference. It introduces a “presumably” quantifier, denoted by $\exists$. Sentences (formulas) quantified by $\exists$ are provisional theorems at a stage of a theory if they follow from the premises at that stage.

Our arguments rely partly on auxiliary assumptions, which make certain analyses tractable. Because auxiliary assumptions have a different status from causal claims that are believed to be true in the world, we mark them with a different quantifier, $\forall$ (for “assumedly”). Their role in inference is the same as for formula quantified with $\forall$.

We assume here that the reader is familiar with the distinction of free and bound variables, and we use this to establish the following conventions.

1. The out-most quantifier of the formula, that is the quantifier whose scope is the whole formula, binds all the free variables of the formula. This allows us to omit the (sometimes long) lists of variables following these quantifiers.

2. If the quantifier whose scope is the whole formula is universal, then we omit the quantifier as well, but we still indicate its scope with square brackets.

Testing what follows from the premises in a stage of a theory in the nonmonotonic logic we use operates on representations of arguments in the form of “rule chains.” The links in these chains are strict rules, definitions, auxiliary assumptions, and causal stories. The chains start with the subject of the argument and terminate with the purported conclusion of the argument (the consequence to be derived). In nonmonotonic inference, different rule chains—each representing an argument embodied in the state of the theory—might lead to
opposing conclusions. The testing procedure determines whether any inference can be drawn at all and, if so, which one. Such testing requires standards for assessing whether a pair of relevant rule chains is comparable in specificity and determining specificity differences for comparable chains. In the case of this paper, the available premises and definitions all point in the same direction; we do not see any rule chains that point to opposing conclusions. Thus all that is required is that we establish a rule chain that connects the antecedent and consequent in a claimed theorem.

Definitions

Only some of the definitions needed for the formal analysis are set off as explicit definitions in the body of the paper. Here we include other needed definitions as well.

Schema.

Let \( f_i = \{f_1, f_2, \ldots, f_i\} \) be the indexed set of \( i \) features that are relevant for a schema. Each feature in the set has a range of possible values. We denote the set of possible values of feature \( f_j \) by \( r_j \) and a value for an object at a time point as \( f_j(x_t) \).

\[
\sigma_l : a \times t \rightarrow \mathcal{P}(r_1) \times \cdots \times \mathcal{P}(r_l); \quad \sigma_l(y, t) = \langle s_1, \ldots, s_l \rangle \equiv S_I
\]

where \( \mathcal{P}() \) denotes the powerset (set of all subsets of a set), \( s_I \) is the set of all the schema conforming values of the \( i \)th feature, and \( I \) is the total number of schema-relevant features. The schema \( \sigma(l, y, t) \) is defined provided that \( l \in I(k, y, t) \).

Type.

\[
ty : a \times t \rightarrow \mathcal{P}(l \times S), \text{ such that } (\langle l, S_I \rangle \in \ty(y, t)) \iff (\sigma(l, y, t) = S_I).
\]

Positively valued type.

\[
[pvt](l, y, t) \iff \exists S_I[\langle l, S_I \rangle \in \ty(y, t)]
\]

\[\wedge \forall x, x', y[(\mu_l(x, y, t) > \mu_l(x', y, t)) \rightarrow E[\bar{\alpha}(l, x, y, t)] > E[\bar{\alpha}(l, x', y, t)]]] \]

where \( \bar{\alpha}(l, x, y, t) \) is a function that tells the intrinsic appeal of the offering of producer \( x \) in type \( l \) to audience member \( y \) at time point \( t \).

Induction from a test.

Let \( S_I \) denote an indexed set of values of \( I \) features and \( F_J \) denote an indexed set of values of some subset of the relevant features: \( 0 \leq J < I \). We use the expression
$f_{i,x,t} \in s_i$ to represent the fact that the $i$th feature of the object $x$ has a value that complies with the schema $\sigma(l, x, t)$ at the time point $t$; and let $\sigma(l, y, t) = S_I$.

$$\text{INDUC}(\sigma(l, y, t), t_J) \rightarrow \forall i, j, x[[l \in I(x, y, t)) \land (j \in J) \land (i \in I \setminus J)$$

$$\land \left[ \mathbf{B}_{y} (f_{j,x,t} \in t_j) \rightarrow \neg \mathbf{B}_{y} (f_{i,x,t} \not\in s_i) \right] \land \left[ \mathbf{D}_{y} (f_{i,x,t} \in s_i) \right] \}}.$$

In that case, we refer to $t_J = \{t_j \mid j \in J\}$ as $y$’s test for judging conformity to the schema $\sigma(l, y, t)$, in notation $\text{TST}(\sigma(l, y, t), t_J)$, and we say that the test has $J$ items.

**Minimal test for induction.** The set of values of the $J$ features, $t_J$, is $y$’s minimal test for induction for the schema for $l$ at time $t$, in notation $\text{MT}(\sigma(l, y, t), t_J)$, iff (1) it is one of $y$’s tests for conformity with the schema; (2) it no more test features for the schema than any other of $y$’s tests; and (3) $y$ induces satisfaction of the schema $\sigma(l, y, t)$ on the untested features from this test.

**Taken for grantedness.**

$$[g(l, x, y, t) \equiv \begin{cases} (I - J)/I & \text{if } ((l, \sigma(l, y, t) = S_I) \in \text{TY}(y, t)) \land \text{MT}(\sigma(l, y, t), t_J) \\ 0 & \text{otherwise} \end{cases}]$$

$$[G(l, y, t) \equiv \sum_{x \mid l \in I(x, y, t)} g(l, x, y, t) / |[x \mid l \in I(x, y, t)]|. $$

**Concept.**

$$[\text{CONCEPT}(l, y, t) \equiv ((l, \sigma(l, y, t)) \in \text{ty}(y, t)) \land (G(l, y, t) > g \approx 1)].$$

**Flat schema.**

Let $\sigma(l, y, t) = S_I$ and let $p^+(l, x, y, t)$ denote the proportion of features values of $x$ on which $y$’s beliefs at $t$ are schema conforming, i.e.,

$$[p^+(l, x, y, t) = \frac{|[f \mid \mathbf{B}_{y} f_{i,x,t} \in s_i]|}{|[f \mid \mathbf{B}_{y} f_{i,x,t} \in s_i]| + |[f \mid \mathbf{B}_{y} f_{i,x,t} \not\in s_i]|}].$$

$$[\text{FLAT}(l, y, t) \equiv \exists x, x' [(p^+(l, x, y, t) > p^+(l, x', y, t))$$

$$\rightarrow E[\mu_{i(l)}(x, y, t)] > E[\mu_{i(l)}(x', y, t)]]].$$

**Code clash outside the minimal test codes.**
Let the set of features on which the schemas \( l \) and \( l' \) clash for the audience member and time point be denoted by \( \text{cl}^+(l, l', y, t) \), that is,

\[
[\text{cl}^+(l, l', y, t) = \{f_i \mid \forall t, x[(f_i \in S_I \cap S_{I'}) \rightarrow \neg((f_i, x, t) \in S_y)]\}];
\]

and let the set of features on which they do not clash be denoted by \( \text{cl}^-(l, l', y, t) \), that is,

\[
[\text{cl}^-(l, l', y, t) = \{f_i \mid \forall t, x[(f_i \in S_I \cap S_{I'}) \rightarrow ((f_i, x, t) \in S_y)]\}];
\]

Finally, let

\[
[((l, S_I(t)) \in \text{ty}(y, t)) \land ((l', S_{I'}(t) \in \text{ty}(y, t)) \land \text{MT}(\sigma(l, y, t), t_j) \land \text{MT}(\sigma(l', y, t), t_{I'}))].
\]

[\text{CLASH}(l, l', y, t) \leftrightarrow \forall j[(j \in j \cap j') \rightarrow (j \notin \text{cl}^+(l, l', y, t)] \land (\text{cl}^+(l, l', y, t) \neq \emptyset)].

**Type contrast.**

\[
c(l, y, t) = \frac{\text{card}(\mu_{l(t)}(y, t))}{|\text{supp}(\mu_{l(t)}(y, t))|}.
\]

**De-novo and de-alio entrants.**

\[
[\text{DE-NOVO}(l, x, y, t) \leftrightarrow (l \in I(x, y, t)) \land \forall t'[(t' < t) \rightarrow (l(x, y, t) = \emptyset)]];
\]

\[
e_n(l, y, t, t') = ||x \mid (t \leq u < t') \land \text{DE-NOVO}(l, x, y, u)||.
\]

\[
[\text{DE-ALIO}(l, x, y, t) \leftrightarrow (l \in I(x, y, t)) \land \forall s[(s < t) \rightarrow (l \in I(x, y, s))]]
\]

\[
\land \exists l', t' \left[\text{CLASH}(l, S_I, l', S_{I'}), y, t) \land \text{CONCEPT}(l', y, t) \land (I = I')
\land \text{MTST}(\sigma(l', y, t), t_{I'}) \land \forall s[(t' \leq s < t) \rightarrow (l' \in I(x, y, s)) \land
\land \forall j((j \in j') \rightarrow B_y(f_{j,x,t} \in t_{I'}))];
\]

\[
e_a(l, y, t, t') = ||x \mid (t \leq s \leq t') \land \text{DE-ALIO}(l, x, y, s)||.
\]

**Postulates**

**Auxiliary assumption 1.** Beliefs available for random samples of features.

Let \( \sigma(l, y, t) = S_I \).

\[
\exists! \pi \forall i [[i \in I] \rightarrow \text{Pr} \neg \exists u[\text{B}_y(f_{i,x,t} = v)] = \pi]]).
\]

25
Auxiliary assumption 2. Common probability of schema-conforming beliefs.
Let $\sigma(l, y, t) = S_I$.

$$\mathfrak{A} \exists \rho \forall i [(i \in I) \rightarrow \Pr[\mathbb{B}_y (f_{i,x,t} \in s_i) | \exists v[\mathbb{B}_y (f_{i,x,t} = v)] = \rho]].$$

Postulate. A type's expected taken-for-grantedness for an audience member normally increases (with some delay) monotonically with its contrast.

$$\mathfrak{A} \exists u \forall s [(0 < s < u) \land (c_{i(l)}(y, t + s) > c_{i(l')} (y, t' + s)) \land (G(l, t, t + s) = G(l', y, t' + s)) \rightarrow E\{(G(l, t, t + u) > E\{(G(l', y, t' + u))\}].$$

Nearest-Possible-World Conditions

$\Phi[t, t'] \leadsto$

1. $[(t \leq s \leq t') \rightarrow \exists S_{l(t,y)}[(l, S_{l(t,y)}(y, s)) \in ty(y, s) \land (S_I(l, y)(y, s) = S_I(l, y, s'))]];$$

2. $[\exists S_{l(t',y)}(l', y, t') \in ty(y, t')];$$

3. $[(t \leq s \leq t') \rightarrow FLAT(l, y, s) \land FLAT(l', y, s)];$$

4. within the period $[t, t']$ neither $\pi$ nor $\rho$ depends on the label, the time, the producer or the audience member.

$\Psi[t, t'] \leadsto$

1. $[\text{clash}(l, l', y, t') \land (0 < |\text{cl}^+(l, l', y, t')| > |\text{cl}^-(l, l', y, t')|)];$$

2. $[(t \leq s \leq t') \rightarrow (l \in I(x, y, s)) \land \exists S_J \forall j [\text{MTST}(l, y, s, t) \land (j \in J) \rightarrow \mathbb{B}_y (f_{j,x,s} \in s_j)];$$

3. $[(t \leq s \leq t') \rightarrow (l \in I(x, y, s)) \lor (\exists j [(j \in J) \rightarrow \mathbb{B}_y (f_{j,x,s} \in s_j)]);$$

4. $l' \in I(x, y, t') \cap I(x', y, t').$

Lemmas and Theorems

Lemma 1. Taken for grantedness and grades of membership.

$$\mathfrak{B} \mathfrak{I} \Phi[t, t'] \land (I(l, y) = I(l', y')) \land (g(l, x, y, t) > g(l, x', y', t')) \rightarrow E[\mu_{i(l)}(x, y, t)] > E[\mu_{i(l)}(x', y', t')].$$
Proof. Under the simplification stated in $\Phi(t)$ and the baseline probability model and the absence of induction, the expected ratio of beliefs that feature values fit the schemas for the two labels to positive beliefs about the relevant feature values are the same for the two situations being compared for any $I$ and $I'$ the expected ratio equals $\rho$. Given the restriction to a flat schemas, this implies that $E(\mu(l)(x, y, t)) = E(\mu(l')(x', y', t'))$ in the absence of induction. So the only systematic difference between these cases must be due to induction. In particular, if the expected number of inductions of schema satisfaction is greater for one situation than the other, then the expected grade of membership is higher for that situation.

Let the random variable that records the number of inductions be denoted by $in(l, x, y, t)$. By the law of total probability,

$$E[in(l, x, y, t)] = E[in(l, x, y, t) \mid \text{min. test for } l \text{ passed}] \cdot Pr\{\text{min. test for } l \text{ passed}\} + 0 \cdot (1 - Pr\{\text{min. test for } l \text{ passed}\}),$$

because no induction takes place if the minimal test is not passed. Under the baseline probability model stated in Auxiliary Assumptions 1 and 2, the probability that $x$ passes $y$'s minimal test for $l$ equals $(\pi \rho)^I$.

Because inductions can only apply to features outside the test code (of which there are $I - J$ for the label $l$ for which the audience member does not have a belief about the value of the feature. The probability of not having a belief on a feature is $1 - \pi$. So the expected number of inductions, conditional on passing the minimal test for $l$ is $(I - J)(1 - \pi)$. Thus $in(l, x, y, t) = (I - J)(1 - \pi) \cdot (\pi \rho)^I$. Similar calculations yield $E[in(l', x', y', t')] = (I' - J')(1 - \pi) \cdot (\pi \rho)^{I'}$. The rule chain supporting the theorem requires that the expected number of inductions for $l$ exceeds that for $l'$, which requires that $(I - J)(1 - \pi) \cdot (\pi \rho)^I > (I' - J')(1 - \pi) \cdot (\pi \rho)^{I'}$. Dropping the common multiplier $(1 - \pi)$ and setting $I = I'$, we check whether $(I - J)\kappa^I > (I' - J')\kappa^{I'}$, where $\kappa = \pi \rho$. after rearranging terms, we must show that

$$\frac{I - J}{I - J'} > \kappa^{I' - I}.$$

By Definition 5, $g(l, x, y, t) > g(l', x', y', t')$ yields $(I - J)I > (I' - J')I'$ and the antecedent states that $I = I'$. Together these conditions imply that $J' > J$. This latter inequality in turn implies that $(I - J)/(I' - J') > 1$ and $\kappa^{I' - I} < 1$ (because $\kappa = \pi \rho$ and the antecedent in the formula stating the lemma states that both $\pi$ and $\rho$ lie between zero and one). So the expected number of inductions is higher for $l$ the type for which the producer $x$ higher taken for grantedness in $y$'s view, which implies that $x$'s expected fit to $y$'s meaning of $l$ is higher than is the case for the other comparison.

\[\square\]
Theorem 1. Concepts versus types and grades of membership.

\[ \mathbb{P} [\Phi[s, s'] \land \Phi[t, t'] \land (I(l, y) = I(l', y')) \land \text{CONCEPT}(l, y, t) \land \neg\text{CONCEPT}(l', y', t') \]
\[ \rightarrow E[\mu_{i(l)}(x, y, t)] > E[\mu_{i(l')}(x', y', t')] \].

Proof. The definition of a concept as a pair of label and schema for which an audience member has a very high level of taken-for-grantedness tells that the relative size of the minimal test for \( l \) is smaller than that for \( l' \). With this inequality granted, the argument chain behind Lemma 1 applies.


Let the conditions stated in the preamble to the foregoing theorem hold and let \( \text{PVT}(l, y, t) \land \text{PVT}(l', y', t'). \)

\[ \mathbb{P} [\Phi[l, l'] \land (I(l, y) = I(l', y')) \land \text{CONCEPT}(l, y, t) \land \neg\text{CONCEPT}(l', y', t') \]
\[ \rightarrow E[\tilde{\alpha}(l, x, y, t') > E[\tilde{\alpha}(l', x', y', t')]. \]

Proof. The rule chain linking the antecedent and consequent results from application of the chain rule to the (rule chain supporting) Theorem 1 and the definition of a positively valued type (Definition 3).

Theorem 2. Typecasting and grades of membership.

\[ \mathbb{P} [\Phi[t, t'] \land \Psi[t, t'] \land (t' > t) \rightarrow E[\mu_{i(l)}(x, y, t)] > E[\mu_{i(l')} (x', y', t')] \]
\[ \land E[\mu_{i(l)}(x, y, t')] < E[\mu_{i(l')} (x', y', t')]. \]

Proof. In the absence of induction, the expected fit of both producers is the same for each label under the assumptions stated in the definition of \( \Psi[t, t'] \), because the audience member’s schemas for the labels are flat and the probability that a schema-relevant feature will be observed is the same as is the probability that a positive belief will be one of schema conformity for each label. Induction can produce both increased fit (when the feature value induced fits the schema) and reduced fit (when the induction goes the other way).

In the case of the first term in the consequent (fit to \( l \)), the result follows from the assumption that the audience member believes that \( x \) passes the minimal test at all points in the interval and no information is available about such a belief for the second producer, \( x' \). It then follows that induction will normally increase the fit to \( l \) for one producer \( (x) \) but not the other \( (x') \). Under the assumption that clashes outnumber non-clashes, the expected net effect of induction is to reduce the fit of \( x \) to the second label \( (l') \) relative to that of \( x' \).

\[ \mathbb{P} \left[ \Phi(t, t') \land \Psi(t, t') \land (t' > t) \land \text{PVT}(l, y, t) \land \text{PVT}(l', y, t') \right] \\
\rightarrow E[\tilde{a}(l, x, y)] > E[\tilde{a}(l', x', y', t')] \land E[\tilde{a}(l, x, y)] < E[\tilde{a}(l', x', y', t')]. \]

Proof. The rule chain linking the antecedent and consequent results from application of the cut rule to the (rule chain supporting) Theorem 2 and the definition of a positively valued type (Definition 3).

Theorem 3. De-novo entrants have higher grade of membership.

\[ \mathbb{P} \left[ \Phi(t, t') \land \Psi(t, t') \land \text{DE-NOVO}(l, x, y, t') \right] \\
\rightarrow E[\mu(l, x, y, t')] > E[\mu(l', x', y', t')]. \]

Proof. According to Definition 10, a producer is a de-novo entrant in a label if the audience member applies the label at that time point and does not apply any label to the producer at any earlier time point. A producer is a de-alio-entrant (from a clashing concept) if the audience member applies the focal label to the producer and has earlier applied to it the label of a clashing concept. In such a comparison, the rule chain that supports Theorem 2 applies; and this rule chain yields the conclusion.

Corollary 3. De-novo entrants have higher expected appeal.

\[ \mathbb{P} \left[ \Phi(t, t') \land \Psi(t, t') \land \text{DE-NOVO}(l, x, y, t') \right] \\
\rightarrow E[\tilde{\alpha}(l, x, y, t')] > E[\tilde{\alpha}(l', x', y', t')]. \]

Proof. The rule chain that supports this implication relies on the application of the cut rule to the (rule chain behind) Theorem 3 and Definition 3.

Lemma 2 De-novo entrants contribute more to expected contrast.

Let the number of entries in two labels be the same over the relevant period \( e_n(l, y, t, t') + e_a(l, y, t, t') = e_n(l', y', t, t') + e_a(l', y', t, t') \).

\[ \mathbb{P} \left[ \Phi(t, t') \land \Psi(t, t') \right] \land (e_n(l, y, t, t') > e_n(l', y', t, t')) \land (c(l, y, t) \geq c(l', y', t)) \\
\rightarrow E[c(l, t, t')] > E[c(l', t, t')]. \]
Proof. This is an immediate implication of Theorem 3, which tells that each de-novo entrant has higher expected GoM in the audience member’s type than does a de-alio entrant. Definitions 11 and 12 tell that an entrant with higher GoM increases contrast more than does one with lower GoM, which implies that a de-novo entrant adds more to contrast. Addition over entries preserves this inequality, given the stipulation that the number of de-novo entries is at least as great as the number of de-alio entries.

\[ \text{Theorem 4. McKendrick and Carroll 2001.} \]

Let the condition stated in the preamble to Lemma 2 hold.

\[ \mathbb{P} \{ \Phi[t, t'] \land \Psi[t, t'] \land (e_n(l, y, t, t') > e_n(l', y', t, t')) \land (c(l, y, t) \geq c(l', y', t)) \rightarrow \exists u[(u \geq t') \land (E[G(l, y, u)] > E[G(l', y', u)])] \} . \]

**Proof.** This theorem follows from a cut rule applied to (the rule chain supporting) Lemma 2 and Postulate 1.
References


