

## Conceptual Spaces and the Consequences of Category Spanning

Balázs Kovács,<sup>a</sup> Michael T. Hannan<sup>b</sup>

a) Università della Svizzera italiana; b) Stanford University

**Abstract:** A general finding in economic and organizational sociology shows that objects that span categories lose appeal to audiences. This paper argues that the negative consequences of crossing boundaries are more severe when the categories spanned are distant and have high contrast. Available empirical strategies do not incorporate information on the distances among categories. Here we introduce novel measures of distance in conceptual space and derive measures for typicality, category contrast, and categorical niche width. Using the proposed measurement approach, we test our theory using data on online reviews of books and restaurants.

**Keywords:** category spanning, similarity, conceptual space, online ratings, restaurants, books

RECENT research has reopened the classic sociological problem of how social categorization shapes social life (Durkheim and Mauss, 1969 [1903]; Zuckerman, 1999; Hsu, 2006; Kovács and Hannan, 2010; Koçak et al., 2014). Much of this research works in the spirit of what Garfinkel (1968) called breaching experiments. It seeks to understand the role of categorical boundaries by examining what happens to actors who ignore them. In this context, ignoring boundaries means claiming membership in multiple categories or adopting feature values that cause audience members to assign multiple categories.

The literature in cognitive science on which we build examines mental representations and their mappings to objects in the world. Mental representations are generally called concepts, and the sets of objects to which a concept refers are called categories. Much sociological work has used the term category to refer to both the mental representation and its extension, its mapping to objects. Because this usage impedes analysis, we reserve the term category to its classic meaning. Sociological analysis treats concepts that have a social standing, some degree of consensus about meaning in an audience. In many situations of interest (including the ones we study), the concepts refer to what can be considered broadly to be genres.<sup>1</sup>

Category-spanning objects have patterns of feature values that fit more than one genre (or other social concept) and, sometimes, their producers claim membership in more than one genre at one time or over time in a sequence of affiliations. Sociological studies show that conceptual boundaries do prove consequential in diverse domains: combining genres generally generates some form of direct or indirect devaluation (for reviews, see Hannan (2010) and Negro et al. (2010b)).

The line of research on boundary crossing has not reached its full potential because it has not yet considered the structure of conceptual spaces. In the interest of tractability, researchers have treated all pairs of social concepts alike, meaning that all kinds of spanning are expected to have the same consequences. This assumption

**Citation:** Kovács, Balázs, and Michael T. Hannan. 2015. "Conceptual Spaces and the Consequences of Category Spanning." *Sociological Science* 2: 252-286.


**Received:** July 21, 2014

**Accepted:** September 24, 2014

**Published:** May 13, 2015

**Editor(s):** Olav Sorenson

**DOI:** 10.15195/v2.a13

**Copyright:** © 2015 The Author(s). This open-access article has been published under a Creative Commons Attribution License, which allows unrestricted use, distribution and reproduction, in any form, as long as the original author and source have been credited. 

lacks credulity. Indeed it is amazing that research that builds on it has managed to find clear patterns of results. Perhaps this is testimony to the strength of the boundaries.

Nonetheless, continued development of this line of theory and research demands that this simplifying assumption be made problematic and that strategies of conceptualization and measurement be devised for addressing more credible accounts of conceptual structure. This article builds models that take account of distances in conceptual space and proposes an empirical strategy for using commonly available data to estimate them. Before turning to these matters, we briefly sketch the range of sociological applications for which our proposed models and methods have relevance.

Sociological research has addressed the consequences of combining genres for diverse kinds of entities. For instance, research on *individuals* in markets finds that bridging lowers: the probability that film actors will gain additional roles (Zuckerman et al., 2003), the odds of winning a bid for contract work (Leung, 2014); the productivity and earnings of researchers in linguistics and sociology (Leahey, 2007); the odds of receiving funding in an online market for lending (Leung and Sharkey, 2014); and the odds of completing a sale on eBay (Hsu et al., 2009).

Research on *organizations* finds that mixing genres reduces: the probability that listed firms receive coverage from stock market analysts, which in turn diminishes valuations (Zuckerman, 1999); ratings of feature films by critics and general audiences, and also box-office revenues (Hsu et al., 2009); ratings of restaurants by critics (Rao et al., 2005) and by the general audience (Kovács and Hannan, 2010); critical evaluations and prices of elite Italian wines (Negro et al., 2011; Negro and Leung, 2013); sales of software products (Pontikes, 2012a);<sup>2</sup> and the ability of terrorist organizations to generate lethal actions (Olzak and Roy, 2013). In addition, work underway explores the implications of spanning for the success of social movements and of terrorist organizations as well as for the technical impact of discoveries (patents) and patent classes (Wezel et al., 2014).

Finally, some research examines the effects of spanning for the *genres themselves*. Pervasive spanning weakens boundaries, which lowers the penalties for spanning (Rao et al., 2005). But it also reduces the value of membership in the sense that appeal to the audience falls even for those who do not stray over the boundary (Negro, Hannan, and Rao, 2010a, 2011).

Although this style of research has produced useful new knowledge, it could yield sharper and deeper insight by changing how we assess multiple-category memberships. Understanding the consequences of moving across boundaries requires analysts to incorporate information about the *distances* among concepts and categories and the strength of the boundaries, which has not yet been done. Audience members regard domains such as cuisine, films, software, and scholarly disciplines as populated by genres arrayed over a (largely shared) conceptual space with similar ones arrayed close to one another and dissimilar ones standing at large distances. Genres also differ considerably in the *clarity of their boundaries*, the degree to which their memberships stand out in the domain. Knowledge of the typography of such a sociocultural space provides essential information about the audience's conceptual structure (Pontikes and Hannan, 2014).

The first part of the article argues that objects that bridge distant concepts have two problems: they are more difficult to interpret than those combining neighboring ones,<sup>3</sup> and it is less likely that the objects meet the expectations of distant concepts,

In designing a new research strategy, we seek to represent conceptual distances in a way that plausibly maps to what audiences see. We build on research in social cognition and computational linguistics to propose a simple co-occurrence approach for building a representation of conceptual structure. A key premise holds that concepts whose instances tend to co-occur (in the sense that entities often share the labels) lie close in the sociocultural space, while those that rarely co-occur are distant (Church and Hanks, 1990). Therefore we use the pattern of co-occurrences for calculating distances.

The second step uses these distances to calculate typicalities (or graded memberships (Hannan et al., 2007)).<sup>4</sup> The current approach to assigning typicalities assumes that objects that get classified into multiple categories are unlikely to be typical members of each (Hsu et al., 2009). We claim that this pattern gets exaggerated when the concepts lie far apart. For example, a scholar whose work gets labeled as <sociology> and <genetics> likely publishes research atypical of each discipline. But this might not be the case when the concepts are close. To continue the example, a scholar whose work is tagged as <sociology> and <gender studies> can plausibly produce research that fits well in each. We want to allow for such differences in assessing the effect of categorical overlaps.

Based on this distance-based approach to calculating typicalities, we propose measures of niche width in the space of categories and of categorical contrast that also make adjustments for distances in the conceptual space. The section of the paper that outlines these developments has a methodological flavor.

After explaining the new approach, we report tests of our argument. A prior study applied part of this argument (but ignored distance) to reviews of restaurants in San Francisco (Kovács and Hannan, 2010). Here, we reexamine these issues in light of the structure of the conceptual space. We use greatly expanded data. These data contain categorizations of the objects in one or more genres as well as consumers' assessments of restaurant quality.

We also test the implications of our arguments using online reviews of books. This application is especially appealing because the site records label assignments by audience members, not site curators as has been the case in previous research using online reviews.

This empirical part of our study has both substantive and methodological goals. We want to learn whether the consequences of spanning depend on: 1) the distances involved and 2) fuzziness. The methodological goal involves learning whether bringing the conceptual structure into the picture improves over the approach used previously.

In the Discussion, we relate our theoretical and methodological approach to a broader set of sociological domains, such as cultural sociology and the sociology of consumption, and describe the importance of taking conceptual spaces into account in studies of various sociological topics, such as that of omnivore consumers.

## Genre Spanning and Appeal

As we explain below, we infer distances in conceptual space based on the relative frequency of co-occurrence of the objects that bear the associated labels—how common it is for agents playing the producer role to be jointly categorized in each pair of genres. This approach places genres close to each other if their categories commonly overlap and far apart if they rarely (or never) overlap. Co-occurrence can be rare for three broad reasons, and each has given rise to sociological theories of the consequences of spanning.

The first reason concerns *skills and learning*. The skills involved in gaining membership in a pair of categories might be so different (or, especially in the case of organizations, so difficult to integrate) that few agents attempt to acquire mastery of both or succeed in doing so. This is the jack-of-all-trades problem (Hsu, 2006). Because spanning makes it hard to develop expertise in any genre, spanners generally accumulate less skill for any one than do specialists. Appeal to an audience increases with the level of genre-specific skills. So those that attempt (certain) combinations will appeal less to the audience.

Second, audiences might believe that jacks-of-all trades are generally inferior to specialists (and, according to the argument just sketched, this belief might have a rational basis). According to typecasting arguments, audiences use generalism as an indicator of low skill by invoking the default that the master-of-none rule applies (Phillips and Zuckerman, 2001; Hsu et al., 2011). Even actors who do manage to develop high skill in multiple genres (by dint of superior overall ability or some kind of resource advantage) have difficulty convincing audiences of this. Such categorical attributions generally make generalists less appealing.

Third, audience members usually find it hard to make sense of objects whose characteristics cause them to be assigned to multiple genres. As mentioned above, objects that get assigned to several (widely separated) genres have feature values that make them dissimilar to the objects that specialize in each. Hannan et al. (2007) argue that audiences generally find more appealing the offers of the more typical object in a category. Again, this argument implies that objects whose genre spanning causes difficulties of interpretation will find it difficult to appeal to the audience.

We do not regard the three arguments as rivals. It seems likely that all three operate in most empirical settings; indeed they are likely intertwined. For instance, an absence of certain co-occurrences due to the technical difficulty of mastering the requisite skills causes the combination to be unfamiliar to the audience and likely confusing when they do encounter it. Therefore the three mechanisms are difficult to discriminate empirically in the absence of natural or controlled experiments (e.g., Leung and Sharkey (2014)). We do not attempt to make such a discrimination with our nonexperimental data. Instead we concentrate on the importance of bringing the conceptual space into the picture in analyzing any of these arguments. We believe that the typography of this space shapes all three explanations: distant genres are harder to master; spanning distant genres is more likely evoke in audiences the impression of dilettantism; and spanning distant genres is more likely to cause confusion.

### Contexts Lacking a Single Categorical Focus

Making precise arguments about combination and appeal requires attention to the setting in which audience members encounter objects. Sometimes audience members see producers portraying different roles as instances of a sequence of different labels. For instance, some studies analyze the success of sellers in online markets, where sellers make offers in one genre but buyers can see histories of participation over genres (Hsu et al., 2009; Leung, 2011). Such contexts provide a clear focal genre for evaluating offers.

Many other studies collect overall assessments of objects without knowing the evaluators' conceptual focus. For instance, Zuckerman (1999) observes decisions by analysts to cover or not cover firms as entities, not as a set of business segments, each to be considered separately. Hsu (2006) observes what ratings filmgoers give to a film, but not what they think of a film as a drama, comedy, horror, musical, and so forth. In these cases researchers observe *overall* appeal, not appeal as an instance of any of the relevant genres.

The difference we highlight concerns the mode in which the audience members typically encounter an object. If these encounters are specialized to a genre context, the context presumably shapes the audience's focus and ties their assessments of appeal to it. Otherwise, audience members likely make assessments that consider *all* of the assigned labels. Indeed it might be very hard for audience members to parse out their separate reactions with respect to each label. So the analysis must be modified to deal with the difference between these two generic situations.

The view that categorical memberships shape expectations about objects suggests that spanning causes problems because audience members find it hard to make sense of the combinations. According to our reading of the wave of research on categories and markets, audience members prefer objects that they find easy to grasp and avoid dealing with hard-to-interpret objects and devalue them.

Here we specialize the arguments to contexts that do not supply a genre focus. That is, we assume that audience members consume and evaluate particular offers (restaurant meals and books in our empirical studies) and that they try to make sense of their producers and evaluate their skills in light of their conceptual inventories. We propose that investigating such issues requires attention to the distances among concepts, specifically the distances among an audience member's *schemas*.

### Fit to Schemas

Schemas are cognitive representations of patterns of attributes and/or social ties. In the context of concepts and categories, schemas tell what it means to be a full-fledged member or instance (Rosch, 1975). Building on work in cognitive psychology, attention to schemas has helped sharpen theory in cognitive and cultural sociology (e.g., DiMaggio 1997; Cerulo 2006; Vaisey 2009), as well as in organizational and economic sociology (Hannan et al., 2007). Here we build on the latter work, which formalizes and specifies the general arguments in cognitive psychology and applies them to sociological questions. Hannan et al. (2007) assert that organizational schemas are sets of profiles of feature values, where each profile identifies a prototype. Specifically, a schema is a subset of the feature space that

contains the prototypes (Pontikes and Hannan, 2014). That is, schemas are the sets of profiles of feature values that yield full membership.

As an organization's fit to a schema can be partial, Hannan et al. (2007) developed a fuzzy-membership approach: an object that fits fully the constraints expressed by an audience member's schema for a label has full grade of membership—is fully typical, in that agent's meaning of the label. Partial fits produce partial memberships. In other words, fit to a schema tells the degree of typicality. Full membership means possessing a profile of feature values that falls in the schema by matching one of the prototypes (distance from the schema is zero). Otherwise, the degree of typicality of an object as an instance of a concept is inversely proportional to the distance of its profile of feature values from the closest prototype in the schema (Pontikes and Hannan, 2014). (This is a classical point-to-set distance: the minimum of the distances of the point from the members of the set.)

### Categorization and Typicality

Issues of partiality pertain to the categorization of objects as instances of concepts. As we noted at the outset, a category is a set of objects that get classified as belonging to a concept. <Bureaucracy> is a concept; the set of objects that an agent regards as bureaucracies is a category. The main line of research on categorization regards the process of assignment as a probabilistic function of typicality. In formal terms, an object  $x$ 's typicality as an instance of the concept  $l$  to an (unspecified) audience member, in notation  $\theta(l, x)$ , is inversely proportional to the distance,  $\delta$ , between its profile of feature values,  $f_x$ , and the audience member's schema for the concept,  $\sigma_l$ . Because it is consistent with experimental evidence, we build on Hampton's (2007) threshold model of categorization using the following negative-logistic form (see Pontikes and Hannan (2014)).

$$\theta(l, x) = \frac{1}{1 + a \exp(b \delta(f_x, \sigma_l))}, \quad a, b > 0. \quad (1)$$

Most sociological work on typicality has focused on label assignments—not profiles of feature values—because labels are generally easier to observe from available records. Below we describe how label assignments have been used to measure grades of membership, and we analyze the conditions under which the common approach coincides with this feature-based approach. Before doing so, we discuss why distances in conceptual space matter.

The sociocultural distance between a pair of concepts depends on the distance in the feature space between their associated schemas (Pontikes and Hannan, 2014). If the schemas for a pair of concepts lie far apart in the sociocultural space, an offer that gets labeled as an instance of both cannot fit either schema well. Any offer that partly fits a pair of distant schemas must be a very atypical instance of the labels associated with those schemas.

Based on this reasoning, we propose that an offer's typicality in any label falls with: 1) the number of labels used to describe it; and 2) the distances among the schemas associated with those labels. When many (distant) labels are applied to an object, atypicality characterizes the object as a whole: not only is it atypical for some



label, it is atypical for *all* of the applied labels. That is, its *overall* typicality is low. We develop the implications of this idea in the context of the new approach below.

## A New Approach

Empirical progress on issues of spanning has been rapid (Hannan, 2010; Negro et al., 2010b). This is due partly to the proliferation of archives and websites that assign categorical memberships to objects and also provide audience reactions. For instance Pontikes (2012a) analyzes the relationship between the positions of software producers in “knowledge space” (induced from patterns of patent citations) and claims to membership in product types as coded from press releases; Hsu (2006) and Hsu et al. (2009) analyze data drawn from websites that provide reviews by professional critics and members of the general audience of films assigned to one or more genres; Hsu et al. (2009) analyze producers that affiliate (by listing products for sale) with one or more of eBay’s product classes; and Carroll et al. (2010) code producers of tape drives as participating in various technical formats; and Leung (2014) analyzes the effects on employment chances of spanning occupations in job histories. Researchers have devised ways to use these data to characterize the strength of producers’ association with genres and to relate patterns of affiliation to the audience reaction.

The current empirical strategy for relating label assignments to typicalities does not pay attention to the distances spanned. We propose generalizations of the measures used in prior research; the new measures use information on the structure of *conceptual spaces*.

Multiple-genre classification relates directly to the nature of boundaries. Until fairly recently sociologists, following the so-called classical perspective on concepts, ignored the possibility that the entities assigned a concept label might differ in the degree to which they typify the concept, the degree to which they “belong.” However, a long tradition of research in cognitive psychology shows that familiar concepts such as <fruit> and <furniture> have an *internal structure*: apples and oranges are viewed as typical fruits and olives and pineapples as atypical fruits and so forth (Rosch, 1975; Rosch and Mervis, 1975; Hampton, 2007; Verheyen, Hampton, and Storms, 2010). One useful way to represent such internal structure views labels as referring to fuzzy sets—sets with partial membership (Hannan et al., 2007). Our research implements this view.

Recent research attempts to construct meaningful typicalities from sparse data containing label assignments but not feature values and schemas.<sup>5</sup> The now-common study design obtains assignments to a predetermined list of genre labels. In most settings studied, including one of ours, a market intermediary (e.g., managers of publications or websites that post reviews) assigns the labels. The analyst does not know how audience members would apply the labels. This means that using such data to test arguments stated at the level of the audience member, as above, requires an assumption that the audience uses the language in a homogenous manner, that they associate similar schemas with the labels of the domain.

Suppose a domain language contains a set of labels for objects, which we denote by  $L$ . The label assignments made to an object  $x$  during an (unspecified) interval is

the tuple:

$$\mathbf{I}(x) = \langle n(1, x), n(2, x), \dots, n(L, x) \rangle,$$

where each element in the tuple records how many assignments are to the particular label: i.e.,  $n(i, x)$  denotes the number of times that the label indexed by  $i$  has been applied to  $x$  (during an unspecified period). (Throughout we denote sets with bold font.) What we call the *general label function* gives, for each label used in the domain, the proportion of the assignments of that label that are made to the entity  $x$ :

$$p(i, x) = \frac{n(i, x)}{\sum_{j \in \mathcal{L}} n(j, x)}. \quad (2)$$

This definition allows for the possibility that the observable data contain multiple label assignments. These can be multiple assignments by the same agent, as in the case of firms' assigning labels to themselves in a series of press releases during the unspecified period (the application studied by Pontikes (2008)) or assignments by multiple agents during the period. For instance, goodreads.com reports that reviewers of Mario Puzo's novel *The Godfather* categorized it as <mystery/crime> (448 times), <classics> (418), <thriller> (188), <historical fiction> (104), <mystery> (101), and a set of other labels with less frequent assignment.

Prior research, with the exception of Pontikes (2012b,a), is tuned to contexts in which some agent, usually a market intermediary such as a website curator or regulator, assigns labels or provides a fixed set of labels from which a seller can choose. Crucially labels are assigned but there are no weights that tell how well the label applies. So we learn from IMBD.com, say, that the Francis Ford Coppola's film version of *The Godfather* is classified simply as <crime> and <drama>.

What we call a *binary* label function tells for each label whether it has been applied at all to the producer/offer during the unspecified interval.<sup>6</sup> In this case, the data on label assignments can be represented as a vector that assigns to an object a value (say one) for each label assigned and a value (zero) for those that are not. A binary label function has the form:

$$l(i, x) = \begin{cases} 1 & \text{if } p(i, x) > 0; \\ 0 & \text{if } p(i, x) = 0. \end{cases} \quad (3)$$

A binary label assignment vector for an object  $x$  is an  $L$ -long vector of ones and zeros. We denote the number of positive binary assignments as  $l_x$ .

A basic analytical question asks how to use these two kinds of label profiles to make inferences about memberships in the concepts/genres that correspond to the labels. We next consider some answers to this question.

### A Discrete Approach

Much recent work follows what we call the discrete approach (for instance, Pontikes, 2008; Hsu et al., 2009; Pontikes, 2012a; Carroll et al., 2010; Negro et al., 2010a; Kovács and Hannan, 2010; Negro et al., 2011). It uses only the number of labels applied to the object, in notation  $l_x$ . This strategy is discrete, because it does not use any metric information about the conceptual space, specifically the distances among



the concepts. That is, it does not adjust for the fact that some pairs are closer than others.

The discrete approach works as follows. In the first (largely implicit) step, which we discussed above, the analyst assumes that, because the schemas for the various labels differ (they impose different constraints), objects with only one label generally fit better the schema for that label than those assigned two labels. For instance, a film classified as <Comedy> and <Horror> likely lacks the typical features of either genre. Similarly a restaurant labeled as <Mexican> and <Thai> can hardly typify either label.<sup>7</sup> The reasoning then makes a similar assertion about two-label versus three-label entities, and so forth.

Following this reasoning, an object's typicality as an instance of any label (interpretability) generally declines with the number of labels it bears (at the same level of label specificity, with no label nested within another). In particular this reasoning suggests that the typicality (a grade-of-membership function) in any assigned label decreases monotonically with the number of labels assigned:  $\theta(i, x) = g(l_x)$ , with  $g'(l_x) < 0$  subject to the condition that  $g$  lies in the unit interval:  $0 \leq g(l_x) \leq 1$ . Hsu et al. (2009) proposed the following functional form for relating binary label assignments and typicality. Their discrete/binary measure,  $\theta_{DB}$ , has the form:

$$\theta_{DB}(i, x) = \frac{l(i, x)}{\sum_{j \in \mathcal{L}} l(j, x)}. \quad (4)$$

For example, if three binary labels are applied to an entity, then its membership in each of these labels is 1/3, and its membership in each of the other labels is set to zero.

The second case combines a generalized label function with the discrete analytic strategy. This yields the *discrete/general* measure of typicality:

$$\theta_{DG}(i, x) = \frac{p(i, x)}{\sum_{j \in \mathcal{L}} p(j, x)}. \quad (5)$$

For example, if reviewers apply the label A to an object 8 times and apply labels B and C each one time, then  $\theta_{DG}(A) = 0.8$  and  $\theta_{DG}(B) = 0.1 = \theta_{DG}(C)$ .

### A Metric Approach: Incorporating Distance

To build metric measures we define *distance in label space* (as contrasted with  $\delta$ , which gives distances in a space of feature-values profiles). Let  $d(i, j)$  be a metric that tells the distance between the labels  $l_i$  and  $l_j$ . (We discuss below how we calculate this distance from the pattern of co-occurrences.)

The *metric/binary* analog to the standard measure in Equation (4) is

$$\theta_{MB}(i, x) = \frac{l(i, x)}{l(i, x) + \sum_{j \in \mathcal{L}} l(j, x) d(i, j)}. \quad (6)$$

For the *metric/general* case, we define typicality as

$$\theta_{MG}(i, x) = \frac{p(i, x)}{p(i, x) + \sum_{j \in \mathcal{L}} p(j, x) d(i, j)} = \frac{1}{1 + \sum_{j \in \mathcal{L}} \left( \frac{p(j, x)}{p(i, x)} \right) d(i, j)}. \quad (7)$$

This definition, like the discrete one (Equation (4)), sets  $\theta(i, x) = 1$  if  $i$  is the only label assigned to  $x$  during the interval, and it sets  $\theta(j, x) = 0$  for  $j \neq i$  in such cases. The addition of each label lowers typicality, but it does so much more when the added label lies far from the others.

**Remark 1.** *How do these intuitively motivated measures relate to threshold models for grades of membership based on distance from the applicable schema in feature space? Consider this issue for the most general measure,  $\theta_{MG}(i, x)$ . Comparing Equations 1 and 7 shows that the grade of membership in a label based on distance between the object's position in feature space from the applicable schema equals the grade of membership in that label based on label assignments ( $\theta(i, x) = \theta_{MG}(i, x)$ ) if and only if*

$$a = \frac{1}{p(i, x)}; \quad (8)$$

$$\delta(f_x, \sigma_l) = (1/b) \ln \left( \sum_{j \in \mathcal{L}} p(j, x) d(i, j) \right). \quad (9)$$

Making the connections between the two spaces shows clearly what the key intuitions imply. The key constraint is that the addition of inter-label distance decreases the fit of object to an agent's schema for a focal label at a decreasing rate. In other words, once an object has been assigned several distant labels, the addition of another has a smaller effect on decreasing typicality than if the same label were assigned to an object with no categorical affiliations other than the focal one. We cannot demonstrate that this relation holds, because we do not have data on feature values and schemas. But we assume that it holds at least roughly. Checking this assumption is an important issue for future research.

## Distance from Co-occurrence

The relatedness of concepts gets reflected in their tendency to co-occur in systems of classification (Gärdenfors, 2004; Widdows, 2004). For example, if <Western> films also tend to be classified as <drama>, then these labels have more similar meanings than pairs that do not tend to co-occur, e.g., <Western> and <comedy>. Such a frequentist approach enables researchers to map out the relationships among social concepts as we show below. The procedure is based on the assumption that co-occurrence maps to similarity.

A basic intuition, backed by research in cognitive psychology, holds that similarity and distance are inversely related. Following the foundational work of Shepard (1987) (see also Tenenbaum and Griffiths 2002; Chater and Vitányi 2003), we posit a negative exponential relationship between perceived sociocultural distance and similarity:

$$\text{sim}(\mathbf{i}, \mathbf{j}) = \exp(-\gamma d(i, j)), \quad \gamma > 0. \quad (10)$$

We use a simple and widely used measure of category similarity due to Jaccard (1901).<sup>8</sup> The Jaccard similarity of a pair of labels amounts to a simple calculation on their extensions.<sup>9</sup> Let  $\mathbf{i}$  denote the extension of  $l_i$ , that is,  $\mathbf{i} = \{x \mid l_i \in \mathbf{I}_x\}$ . Then the similarity of labels  $l_i$  and  $l_j$  can be defined as the ratio of the number of objects categorized as both  $l_i$  and  $l_j$  to the number categorized as  $l_i$  and/or  $l_j$ . Formally, if  $|\mathbf{i} \cap \mathbf{j}|$  denotes the cardinality of the set of objects categorized as both  $l_i$  and  $l_j$ , and  $|\mathbf{i} \cup \mathbf{j}|$  denotes the cardinality of the set of objects categorized as  $l_i$  and/or  $l_j$ , then

$$J(\mathbf{i}, \mathbf{j}) = \frac{|\mathbf{i} \cap \mathbf{j}|}{|\mathbf{i} \cup \mathbf{j}|}. \quad (11)$$

This index takes values in the  $[0, 1]$  range, with 0 denoting perfect dissimilarity and 1 denoting perfect similarity. For example, the dataset on restaurants analyzed below contains nine restaurants labeled as <Malaysian> and eleven <Singaporean>. Four of these restaurants receive both labels. Thus the similarity of <Malaysian> and <Singaporean> in these data is  $4/(9 + 11 - 4) = 0.25$ .

### Categorical Niche Width

The concept of the width of a categorical niche provides a useful way to analyze typicality in the context of multiple-category memberships.<sup>10</sup> Hsu et al. (2009) defined the *width* of an object's niche in conceptual space as a way to summarize the differences among market participants in the degree to which they specialize in terms of categorical memberships. As we noted above, a genre specialist belongs to only one category. A genre generalist is a partial member of several.

The category-membership niche, defined in terms of label-based typicalities, is a tuple:  $\theta_x = \langle \theta_x(1), \dots, \theta_x(L) \rangle$ . Hannan et al. (2007) defined the width of a fuzzy niche using the Simpson (1949) index of dissimilarity applied to a tuple of memberships (typicalities). When applied to categorical niches, the Simpson index yields the following measure of niche width, where the subscript DG stands for the discrete/general case:

$$W_{DG}(\theta_x) = 1 - \sum_{i \in \mathcal{L}} \theta_{DG}(i, x)^2. \quad (12)$$

The measure of niche width proposed by Hsu et al. (2009) can be shown to be a special case of Equation (12) applied to binary label functions (Equation (4)). In this case, niche width of an object depends only on the number of distinct labels assigned to it, which we have denoted by  $l_x$ . This discrete/binary measure of niche width can be written:

$$W_{DB}(\theta_x) = \frac{l_x - 1}{l_x}. \quad (13)$$

Unlike the discrete approach, our proposal does not constrain the sum of squared typicalities to lie in the unit interval. As a result, niche width calculated as a Simpson index can be negative, which does not make sense. As we see it, new thinking about niche width is needed once distance enters the picture.

Our proposed measures depend on the the total distance among the labels assigned to a market participant for both general and binary cases:

$$D_x = \begin{cases} \sum_{i \in L} \sum_{j \in L} p(i, x) p(j, x) d(i, j) & \text{if labels are general;} \\ \sum_{i \in L} \sum_{j \in L} l(i, x) l(j, x) d(i, j) & \text{if labels are binary.} \end{cases} \quad (14)$$

In devising a measure of niche width, we want to satisfy four desiderata. The measure should be non-negative and have a minimal value of zero (if an agent gets assigned to a single label), and increase with the number of labels assigned and with the distances among them. Moreover, we want the measure to reflect the intuition that a given increment of distance spanned has a bigger impact on niche width at low levels of total distance spanned than at higher levels. In other words, if a producer has a narrow niche and adds a label at a fixed distance, its niche width increases more than if it initially had a broad niche. Specifically we want a measure that implements the intuition that niche width grows with increasing  $D_x$ , holding constant the number of labels applied (i.e.,  $\partial W / \partial D_x > 0$ ), but this effect is weaker at higher levels of total distance ( $\partial^2 W / \partial D_x^2 < 0$ ). With this definition, niche width shrinks as the number of labels applied increases, when total distance is held constant (i.e.,  $\partial W / \partial l_x < 0$ ).

There might be many alternatives that meet our desiderata. We propose a particularly simple metric/general measure:

$$W_{MG}(\theta_x) = 1 - \left( \frac{1}{1 + D_x} \right). \quad (15)$$

When the labels are binary,  $D_x$  will almost surely be much larger than the comparable sum for the general label assignments; even a single assignment of many labels causes the binary function to equal one when each assignment is low in relative frequency. This is because, as we explained above, binary measurements are insensitive to the difference between commonly applied and rarely applied labels, so we need to adjust for this difference. We experimented with a variety of cases that resemble observations in our data and got measures of niche width that fit the abovementioned intuitions when we convert the total distance into an average distance per label (minus one). We use that specification in our empirical analysis below for the binary case (restaurants), with  $\bar{D}_B(x) = D_B(x) / (l_x - 1)$ . That is, we use the following measure of niche width for the metric/binary case in our analysis:

$$W_{MB}(\theta_x) = 1 - \left( \frac{1}{1 + \bar{D}_x} \right) \quad \text{if } l_x > 1, \quad (16)$$

and it equals zero if only one label is applied ( $l_x = 1$ ).

We obtain very similar results with the normed measure in Equation (16) and with the unformed one (Equation (15)). Whether the normed distance better matches sociological intuitions in other applications is an open question.

### Category Contrast

Contrast refers to the degree to which a set stands out from the background, the clarity of its boundary. Hannan et al. (2007) defined the contrast of a category as the average typicality among those with the label who have a nonzero typicality. Using our notation for label functions and extensions, we have

$$C(l_j) = \sum_{x \in \mathbf{o}} \theta(j, x) / |\mathbf{o}|, \quad (17)$$

where  $\mathbf{o}$  denotes the set of objects to be classified and  $|\mathbf{o}|$  denotes the number of objects in the set.

Obviously the discrete approach tends to understate the contrasts of categories that lie close to others in the conceptual space and overstates contrasts for categories that overlap distant ones. In our empirical analysis below, we show that adjusting for distances in calculating contrasts does indeed make a difference.

### Categorical Niche Width, Contrast, and Intrinsic Appeal

The rest of the article illustrates and evaluates the new approach by using it to analyze the effects of categorical specialism/generalism on appeal to an audience. We adopt the standard notion that audiences seek to interpret and evaluate the offers of producers on a market. (In the two empirical analyses presented below, the offers are restaurant meals and books.) We focus on the interpretability of an offer. As we suggested above, offers whose feature values fit partially to multiple genres are hard to interpret. Moreover, this problem is compounded when the genres spanned are distant. We gain empirical leverage by connecting interpretability to appeal to the audience. We develop the implications of this intuition in terms of niche width and contrast.

Objects that get associated with multiple labels are neither fish nor fowl. Such generalism can be expressed well in terms of the *width of the categorical niche*. Objects associated with a single label—specialists—have a niche width of zero. Niche width increases with the number and diversity of categorical affiliations. We propose that an object's interpretability declines with its categorical niche width.

Following Hannan et al. (2007), we focus on genres with positive valuation: those where the intrinsic appeal of the object for audience members with typical tastes increases with the fit to an applicable schema. A testable implication of the overall argument is<sup>11</sup>

**Hypothesis 1.** *The intrinsic appeal of an offer decreases with its categorical niche width.*

The second part of the argument concerns a category-level variable: contrast. (This is the part of the argument that bears most distinctly on interpretation rather than skill.) According to Hannan et al. (2007), categories with very fuzzy boundaries exert less social power than crisper ones. A relatively crisp category stands out from the social background and likely serves as a basis of enduring expectations about those who bear the genre label. The main intuition holds that membership in a high-contrast category conveys greater advantage than membership in a fuzzy

one. This intuition has been found to hold in diverse contexts, as we noted in the introduction.

Negro, Hannan, and Rao (2011) argue that pervasive spanning by members of a category lowers its contrast and thereby reduces the appeal of all of its members. This process likely works in two ways. One way involves the relationships among the categories. Fuzziness implies a loss of distinctiveness relative to the other categories, raising questions about what comparisons are appropriate for the members of a category. With increasing fuzziness, clusters of objects become less salient and elicit lower attention. Previous research shows that comparisons become more difficult; audience members have trouble using distinct descriptors and develop attitudes of reserve, strangeness, and even aversion or repulsion (Griswold, 1987). Negative evaluations are more common, and audience members claim previous judgments were too generous or neglected important differences.

A second way involves the loss of agreement about the meaning of the genre (Hannan et al., 2007). Declining contrast means that the set of objects to which some audience members apply a label share fewer schema-relevant feature values. Such a situation sparks disagreement about about which objects merit a particular label and what the label means. According to Simmel (1978 [1907]), the loss of distinctiveness “hollows out the core of things.” Hence, low-contrast categories lack intrinsic appeal relative to higher-contrast ones.

A challenge is to translate these intuitions about the advantages of affiliation with a high-contrast category to the context of spanning. If contrast matters, which category’s contrast? After considering a number of possible answers to this question, Kovács and Hannan (2010) concluded that two facets of this matter need to be considered: 1) whether the object is associated with any high-contrast category; and 2) how much the highest-contrast membership dominates the others.

If belonging to any high-contrast category makes it easier to make sense of social actors or artifacts, then having at least one membership in a high-contrast category will be beneficial. We posit that the interpretability of an object increases with the maximum of the contrasts of the categories assigned.

But membership in two or more high-contrast categories likely confuses the audience (Kovács and Hannan, 2010). For instance, the combination of the high-contrast <Soul food> and <Thai> would be more difficult to interpret than the combination of the low-contrast <new American> and <Asian fusion>. So it is clearly not enough to analyze only maximum contrast. We need to know the consequences of having second (and third, and fourth) high-contrast memberships. We reason that, net of the effect of the maximum-contrast membership, having another high-contrast membership will confuse the audience and thereby reduce intrinsic appeal.

Given assignment of multiple labels, interpretability is low when a label other than the one with maximum contrast also has high contrast. We use the term *secondary contrast* to refer to the next-to-maximal contrast. In this case we posit that the interpretability of an object decreases with its secondary contrast. Again we have a testable implication of the argument:

**Hypothesis 2.** *The intrinsic appeal of an offer*

**a.** *increases with the maximum contrast of the set of categories assigned.*



b. *decreases with the secondary contrast of the set.*

## Empirical Applications: Reviews of Books and Restaurants

We turn now to empirical application to books and restaurants. We take advantage of the upsurge in interest and involvement in websites that publish critical reviews by general audience members. The hypotheses tested, the theoretically relevant measures used, and the method of estimation used is the same for both studies. So we begin by describing these matters. Our theory pertains to intrinsic appeal, and the ratings reflect actual appeal, which, according to Hannan et al. (2007), depends on intrinsic appeal and engagement. It seems clear that book publishers and restaurant staff have engaged the audience, but we do not know much about the intensity of engagement. We assume that actual appeal is proportional to intrinsic appeal. (In the case of restaurants, we do control for engagement outside the food domain, which likely serves as the major source of variation in engagement as a restaurant.) We assume that variations in intrinsic appeal overwhelm variations in engagement so that actual appeal is proportional to intrinsic appeal.

### Measurement

*Dependent Variable.* Both data sources, goodreads.com and yelp.com, allow registered users to submit one or more reviews of books (for Goodreads) or any producer/product (for Yelp) and to record summary ratings, the outcome variable in our analyses. These ratings range from one to five stars (the highest rating) on both sites. The distribution of ratings of books is as follows: the modal rating is four stars (37 percent of the reviews), followed by five stars (31.5 percent), three stars (21 percent), two stars (7.5 percent), and one star (3 percent). Most restaurant reviews give three or more stars: the mode of the distribution is four stars, and the mean is 3.7.<sup>12</sup>

*Covariates Based on Categorical Dimensions.* In terms of the notation, introduced above, we calculate typicalities using  $\theta_{DB}$  and  $\theta_{MB}$  with the Yelp data (because we have only binary labels). With the goodreads.com data we can calculate all four variations. In our analysis we concentrate on  $\theta_{DB}$  and  $\theta_{MG}$ , which allows us to compare the full-blown new approach to the standard used in the earlier studies. Typicality measures do not appear directly as variables in the stochastic specifications we estimate. Rather they form the basis for calculations of niche width and contrast.

We conduct analyses using the three alternative measures of niche width: the purely discrete measure used in previous research  $W_{DB}$  and our proposed alternatives:  $W_{MG}$  (for book reviews) and  $W_{MB}$  (for restaurant reviews).

In analyzing the restaurant reviews, we can also pay attention to another aspect of the niche. Any restaurant that engages outside the very broad <food> class has a very wide niche. We constructed a dummy variable to capture this dimension of niche width, labeled “any non-food genre,” which equals one for restaurants

with an assignment to a non-food genre, such as <art gallery> or <gas station>, and equals zero otherwise. We do not have a comparable situation for books.

What we call the maximum contrast of an offer is the maximum of the contrasts (average typicality in a label for those offers with positive typicality) over the labels assigned to it; and its secondary contrast is set to the value for the next-highest contrast label assigned. Of course, the values of these two variables depend on the method used for measuring typicality.

The distributions of these theoretically relevant variables can be found in Table 1. Note that we discriminate among the various combinations of measurements of interest. Note that there are large and systematic differences by type of measurement, especially for books.

We report these distributions for all cases and separately for cases with multiple labels. For books, the distributions are quite similar for all books and only those assigned to multiple genres. But this is not so for our sample of restaurants, because three quarters bear only one genre label. On average, secondary contrast and niche width are much greater in the sample of specialists than for the whole sample. As we will see below, this has a big impact on the results.

We illustrate the implications of adjusting for distance for particular cases, using the data on restaurants. Table 2 shows the implications for measuring contrast. The discrete approach assigns a typicality of 0.5 in each label to a restaurant categorized as both <Spanish> and <Basque>. What happens under the alternative that pays attention to distance spanned (with  $\gamma = 1$ )? Because the similarity of <Spanish> and <Basque> in the combined data equals 0.48, the distance between these genres equals  $-\ln(0.42)/1 = 0.73$ . So our measurement strategy assigns typicality of  $1/(1+0.73)=0.54$  in both labels. What about dissimilar labels? Consider the pair <Chinese> and <French>, whose similarity is 0.0007. Using the Shepard transformation, this gives a distance 7.2, which means that a restaurant in the intersection has a typicality of  $1/(1+7.2)=0.12$  in each genre according to our (binary, metric) measure. Recall that the binary, discrete measure assigns a typicality of 0.5 in each genre to such cases.

We expect that incorporating distance in calculation of typicality would increase or keep constant the contrasts of such highly overlapping pairs as <Indian>/<Pakistani> and <Japanese>/<sushi bar>. Because some members of these categories do not restrict their memberships to the pairs, contrasts generally rise even for them when we adjust for distances, as can be seen in Table 2. Only for <Pakistani> does contrast rise. However, it declines only slightly for others such as <Chinese> and <Italian>. At the other extreme, contrast falls considerably with the distance correction for <Asian fusion>, <barbecue>, and <vegetarian>. What matters here are not the exact magnitudes of changes in contrasts (as these depend on the specific distance measure used), but that incorporating distance alters the ordering in terms of contrast.

Adjustments for distance also affect measures of niche width. Table 3 compares the two measures for some restaurants with three label assignments. Of course,  $W_{DB}$  does not discriminate cases with the same number of labels assigned; however, the distance-based measures do. The combinations in the first rows of Table 3 span a very considerable distance, and they receive high values for  $W_{DB}$ . However, the

**Table 1:** Distributions of theoretically relevant variables with alternative measurements

Variable	Measure	Mean	S.D.	Min.	Max.
All Books					
Maximum contrast	$\theta_{DG}$	0.241	0.076	0.041	1.000
Maximum contrast	$\theta_{MG}$	0.553	0.042	0.363	1.000
Secondary contrast	$\theta_{DG}$	0.189	0.080	0.008	0.626
Secondary contrast	$\theta_{MG}$	0.525	0.053	0.310	1.000
Niche width	$W_{DG}$	0.742	0.127	0.000	1.000
Niche width	$W_{MG}$	0.428	0.080	0.000	0.609
Books with multiple genre assignments					
Maximum contrast	$\theta_{DG}$	0.241	0.076	0.041	0.842
Maximum contrast	$\theta_{MG}$	0.553	0.042	0.363	1.000
Secondary contrast	$\theta_{DG}$	0.189	0.080	0.008	0.626
Secondary contrast	$\theta_{MG}$	0.525	0.053	0.310	1.000
Niche width	$W_{DG}$	0.744	0.120	0.025	1.000
Niche width	$W_{MG}$	0.432	0.069	0.054	0.609
All Restaurants					
Maximum contrast	$\theta_{DB}$	0.784	0.091	0.417	1.000
Maximum contrast	$\theta_{MB}$	0.695	0.112	0.240	1.000
Secondary contrast	$\theta_{DB}$	0.231	0.320	0.000	0.879
Secondary contrast	$\theta_{MB}$	0.192	0.270	0.000	0.805
Niche width	$W_{DB}$	0.185	0.255	0.000	0.750
Niche width	$W_{MB}$	0.292	0.402	0.000	0.946
Restaurants with multiple genre assignments					
Maximum contrast	$\theta_{DB}$	0.759	0.082	0.417	0.975
Maximum contrast	$\theta_{MB}$	0.670	0.095	0.240	0.961
Secondary contrast	$\theta_{DB}$	0.659	0.101	0.400	0.879
Secondary contrast	$\theta_{MB}$	0.547	0.119	0.138	0.805
Niche width	$W_{DB}$	0.528	0.063	0.500	0.750
Niche width	$W_{MB}$	0.833	0.102	0.505	0.946

restaurants in the bottom rows, which also bear three labels, combine closer genres, and the distance-based measures of niche width are lower.

*Controls.* In both studies we control for variation among reviewers in engagement in the genre and the website using (the natural log of) the number of reviews posted for restaurants, following Hsu et al. (2009). We refer to this variable as the reviewer's *activism*. We also control for the book's/restaurant's prominence (on the website), measured as the natural log of the number of reviews it receives, and we include the date of a review to control for secular trends in appeal. Finally, we control for price levels. We have been able to collect actual retail prices (taken in May 2013 from the sites amazon.com and barnesandnoble.com) for 78 percent of the books, accounting for 85 percent of the reviews. There is little price variation in the books

**Table 2:** Contrasts of selected restaurant genres constructed using alternative measures of grade of membership: discrete/binary and the metric/binary

Genre:	Measure of category contrast based on	
	$\theta_{DB}$	$\theta_{MB}$
Asian fusion	0.739	0.596
Barbecue	0.733	0.598
Chinese	0.899	0.840
Gastropub	0.669	0.490
Indian	0.693	0.661
Italian	0.796	0.709
Japanese	0.715	0.669
Modern European	0.591	0.436
Pakistani	0.513	0.543
Steakhouses	0.739	0.605
Thai	0.926	0.885
Vegetarian	0.627	0.468

reviewed. However, there is considerable variation in prices of restaurant meals. In the yelp.com data, price can take four ordered values: coded 1 (for “cheap”) to 4 (“splurge”).

### Estimation

We estimate ordered logit specifications by maximum likelihood with random effects for reviewer<sup>13</sup> to assess the effect of the theoretically relevant variables on appeal, because the outcome variable (number of stars) is discrete and ordered. The stochastic specification used is

$$A_{it}^* = \mathbf{x}_{it}'\boldsymbol{\beta} + v_i + \epsilon_{it},$$

where  $\mathbf{x}_{it}$  denotes a time-varying vector of covariates,  $\boldsymbol{\beta}$  denotes a vector of parameters,  $v_i$  are independently and identically distributed with  $N(0, \sigma_v)$ , and  $\epsilon_{it}$  has a logistic distribution with mean zero and variance equal to  $\pi/3$ , and

$$A_{it}^* = i \text{ if } \delta_{i-1} \leq A^* < \delta_i,$$

where  $\delta_i$  ( $i = 1, \dots, I$ ) are cut points  $\delta_0 = -\infty$ , and  $\delta_I = \infty$ .

## Study 1: Books

In the first study we explore the consequences of genre spanning in the setting of online reviews of books. This domain and the data source we use offer two advantages over restaurants, which makes the book domain especially useful for

**Table 3:** Examples of alternative calculations of categorical niche width for restaurants with assignments to three labels using the binary-discrete approach and binary-metric approach

Labels applied:	$W_{DB}$	$W_{MB}$
Seafood; Food Stands; Sandwiches	0.667	0.942
Mexican; American (New); Burgers	0.667	0.940
Chinese; Vegetarian; Kosher	0.667	0.932
Pizza; Burgers; Sandwiches	0.667	0.930
French; American (New); Brasseries	0.667	0.922
Indian; Buffets; Pakistani	0.667	0.901
Middle Eastern; Halal; Mediterranean	0.667	0.901
Mediterranean; Middle Eastern; Turkish	0.667	0.886
Middle Eastern; Greek; Mediterranean	0.667	0.852
Singaporean; Malaysian; Indonesian	0.667	0.829
Spanish; Tapas; Bars; Basque	0.667	0.792

our purposes. First, evaluations of restaurants are affected by variations in service quality, which introduces noise in the relation between spanning and appeal. This does not seem to be an issue in the case of books. Second, the data source we use for book reviews allows us to construct generalized label functions based on categorizations made by the audience members, not the site curators.

Our data come from the book-review website, goodreads.com. This is the largest such website with more than a million reviewers and more than ten million ratings of books. The website, founded in 2007, provides a free interface for reviewers to create a profile, review books, and connect to other registered users. Importantly, Goodreads.com provides information about how reviewers categorize books. Each reviewer can assign one or more labels to a book but assigning them to a labeled “bookshelf.” These labels are freely chosen by the reviewer. They could be those already used by others, or they could be idiosyncratic. Goodreads.com aggregates these individual tags by book and lists for each book the ten most commonly applied tags and the number of reviewers who applied that tag to the book. For example, *Venetia Kelly's Traveling Show* was classified as <historical fiction> (29 times) <Ireland> (18), <adult fiction> (3), <Irish literature> (2), and <mystery> (2). Note that these tags can have partially overlapping meanings and do not typically form a neat and nested classification hierarchy. Table 4 contains additional examples.

This more refined information about labeling obviously allows more accurate assessments of typicalities. A book whose tags are half <mystery> and half <romance> is less typical as a mystery and more typical as a <romance> than the one categorized as 80/20.

In April 2013, we selected all 2,075 English-language books that were first published in February and March 2010 from the records available on goodreads.com.<sup>14</sup> We downloaded all the reviews of these books: 620,594 reviews posted by 111,185 reviewers. Readers who post reviews on goodreads.com appear to provide a representative sample of the reading public in the United States: the average age (38.4

years) and the female dominance (73 percent of the registered users are female) match the demographic distribution of fiction readers (Griswold et al., 2005).

The count of reviews has a skewed distribution: the average number of reviews per book is 299; 25 percent get ten or fewer reviews, 40 percent get more than 100, and 10 percent of the books get more than 1,000. On average, reviewers (of any of the books in our sample) posted 5.6 reviews of the 2,075 books; 29 percent reviewed only one book in our sample, 16 percent reviewed two books, 19 percent reviewed more than ten, and about 2 percent posted more than 25 reviews. Table 1 reports descriptive statistics for the theoretically relevant variables.

Table 4 illustrates the niche widths of a few randomly selected books as calculated using the two alternative measures of grade of membership. Take for example *A Vampire's Mistress*, labeled four times as <paranormal-romance> and twice as <harlequin>. Suppose we collapsed the label assignment to binary labels. Then  $W_{DG} = 1 - ((4/6)^2 + (2/6)^2) = 0.44$ . In contrast, the metric approach indicates a much wider niche because the two genres assigned are dissimilar in the sense that they rarely appear together. In other cases, metric niche widths fall below the discrete ones. Such is the case with *Sword of My Mouth*, which is classified as <comics>, <graphic novels>, <dystopia>, and <fantasy>, which are close to each other.

Applying the new conceptualization and measurement strategy requires that we supply a value for the free parameter  $\gamma$  that relates distance and similarity in Equation (10). In the analyses reported below we use  $\gamma = 2$  because this provided the best model fit, but we note that the empirical patterns are quite robust for the choice of gamma.

## Results

We ask two questions about the empirical results. First, do these data support our hypotheses about typicality, interpretability, and appeal? Second, does the new conceptualization, based on the structure of a conceptual space, make a substantive difference?

We see in Table 5 that adopting the distance-based approach improves the fit greatly: the difference in log-likelihoods between the discrete approach and the distance-based approach is 498. In addition there is a slight improvement in terms of (mean squared) prediction error.

All three hypotheses are supported with both approaches to measurement. The effect of maximum contrast on appeal is positive and significant. The effect of secondary contrast (the next-higher contrast) is negative and significant. And the effect of niche width is negative and significant. This is the case both when the covariates are measured using the discrete approach and when they are based on metric measures (with generalized label functions). According to these results, books that hew to genre conventions and eschew themes associated with other genres, especially crisp ones, have higher appeal to this general audience. Net of these effects, books that get associated with many dissimilar genres fail to appeal.

Beyond these qualitative implications, choice of measurement strategy has a very big impact on the estimated magnitudes of the effects. Both contrast effects



**Table 4:** Examples of alternative calculations of categorical niche width for books using the discrete-binary approach and metric-general approach

Title	Genre weights	$W_{DG}$	$W_{MG}$	Genres
A Vampire's Mistress	4;2	0.444	0.493	paranormal-romance; harlequin
Autobiography of a Sadhu	3;3;2	0.656	0.558	spirituality; autobiography; Hinduism
Dimanche and Other Stories	20;7;5;3;2	0.644	0.557	short stories; France; historical fiction
Scryer's Gulch Episodes 1–10	5;4	0.494	0.448	literary fiction; World War II
The Pendant	6;3;3	0.625	0.567	fantasy; steampunk
Signed Dealed and Delivered	4;3	0.490	0.393	historical fiction; Italy; medieval
Container Gardening For Dummies	7;3;2	0.569	0.505	biography; music
Obama Zombies	27;10;2;2;2;2;2	0.645	0.470	gardening; nonfiction; how-to
				politics; nonfiction; science;
				literary fiction; essays
Catherine's Gift	3;3	0.500	0.245	collections; biography; autobiography
Angelina and the Royal Wedding	5;2	0.408	0.203	biography; nonfiction
Sword of My Mouth	9;8;2;2	0.653	0.418	picture books; childrens
Addicted to Her	6;5;3;2	0.711	0.317	comics; graphic novels; dystopia; fantasy
Shadow of the Sun	35;19;11;5;5;3;2	0.723	0.342	young-adult; teen; romance; realistic fiction
				paranormal; fantasy; young-adult; romance
				suspense; paranormal-romance; magic
Wicked Games	5;3	0.469	0.224	chick-lit; romance
Chasing the Dead	7;3;2	0.569	0.264	crime; mystery; suspense
The Serialist	9;5;4	0.623	0.253	mystery; mystery-thriller; crime
Captain America	70;46;20;9;4;4;2	0.686	0.308	comics; graphic novels; Marvel; superheroes
				graphic-novels-comics; comic book; time travel

**Table 5:** Effects of categorical contrast and niche width on the appeal of books: ML estimates of ordered-logit specifications with reviewer random effects

	Niche width measure	
	$W_{DG}$	$W_{MG}$
Maximum contrast	2.25 (0.090)	7.47 (0.180)
Secondary contrast	-0.994 (0.076)	-1.56 (0.128)
Niche width	-0.924 (0.058)	-0.853 (0.106)
Book prominence	0.527 (0.005)	0.486 (0.005)
ln(Price)	0.071 (0.009)	0.105 (0.009)
ln(pagecount)	-0.027 (0.005)	-0.015 (0.005)
Review date (in years)	-0.114 (0.008)	-0.115 (0.008)
Reviewer activism	-0.166 (0.017)	-0.197 (0.017)
Cut points	-6.32 -3.29 0.023	-3.23 -0.217 3.11
$\hat{\sigma}_v^2$	4.10	7.19
Observations	530,296	530,296
Mean sq. pred. error	1.086	1.078
Log-likelihood	-471,422	-470,924

*Note:* Figures in parentheses are standard errors; all effects are significant at the 0.001 level.

are much larger in absolute value with the metric measurements: the effect of maximum contrast is three times larger with the metric measurement, and the effect of secondary contrast is roughly 50 percent larger in absolute value. On the other hand, the estimated effect of niche width is only about 20 percent as large with the metric measure. So adjusting for distances appears to have heightened the importance of contrast relative to niche width.

The effects of the control variables are in line with expectations. Popular books receive higher ratings, longer books get lower ratings, and more expensive books get higher ratings. Activist reviewers are more critical and give lower ratings.

## Study 2: Restaurants

The domain of restaurants contains many broadly understood genres (Carroll and Wheaton, 2009).<sup>15</sup> We draw our data from yelp.com, which categorizes restaurants in 78 genres. These genre labels appear prominently on the site.<sup>16</sup> A call for reviews of restaurants for some location yields a screen with the list of 78 genre labels. Clicking through to a label produces a list of establishments shown by name, address, neighborhood, and a set of categorizations.

Our data include all the organizations in San Francisco categorized in at least one genre in the restaurant domain. Restaurants receive very frequent reviews; and they are distributed over a broad diversity of genres. Some labels concern various ethnic/national cuisines, e.g., <Basque>, <soul food>, and <Thai>. Others refer to the mode of service, e.g., <buffet> and <food stand>. Still others pertain to the key ingredient(s) or dishes, e.g., <burgers> and <seafood>, and some refer to food codes, e.g., <halal> and <vegan>. Some restaurants also get classified in non-food genres such as <art gallery> and <bowling alley>.

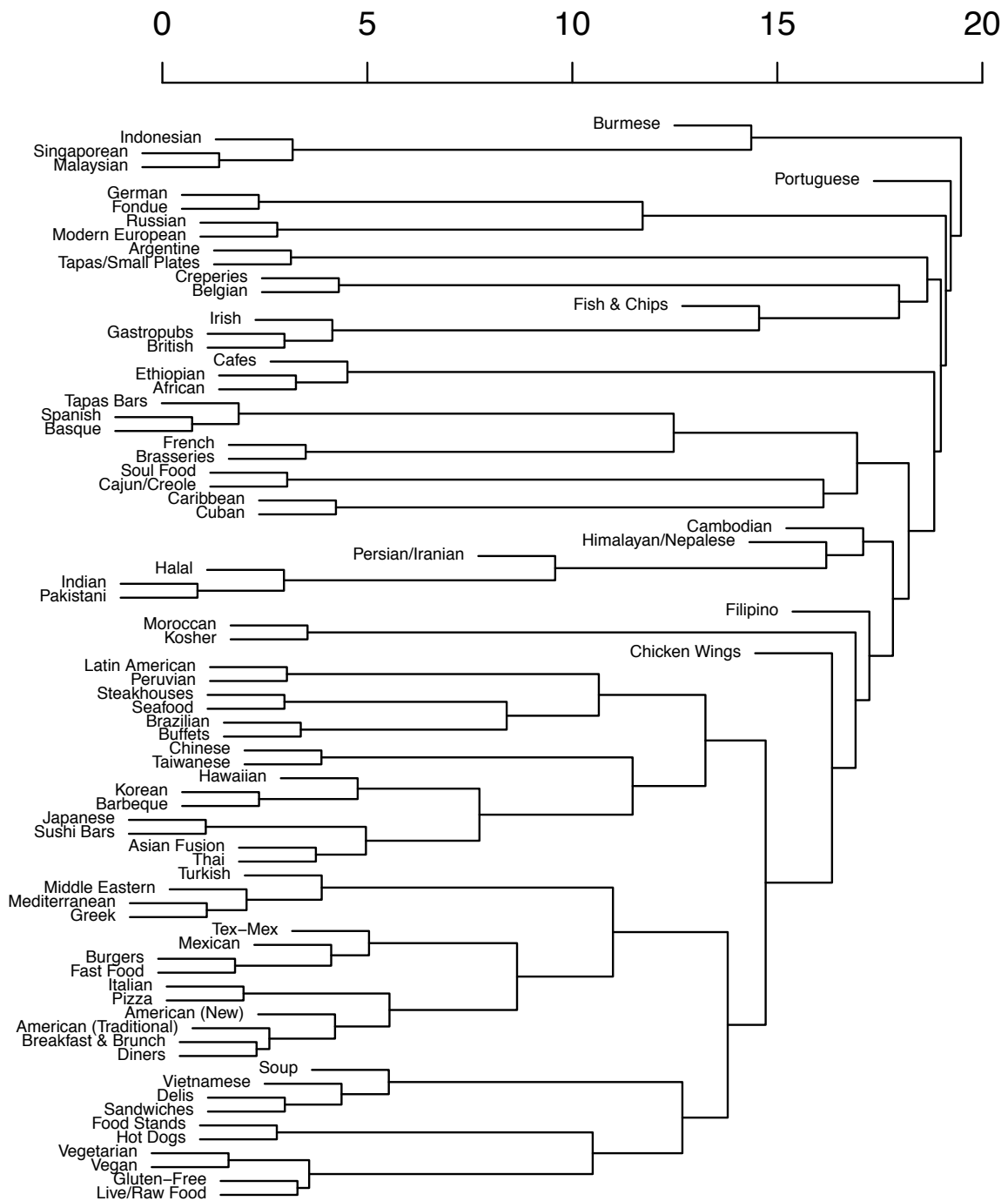
We analyze reviews posted between October 2004 and September 2011. The sample contains 767,268 reviews written by 59,473 reviewers about 3,976 producers. This website encompasses a broader audience than most food and gourmet magazines and media outlets (Johnston and Baumann, 2007).

### Genre Classifications

Most restaurants (73 percent) are assigned to only one restaurant genre. About a quarter (24 percent) are assigned two, and roughly 3 percent get three or more. The most common labels are <Mexican> (1,907 instances), <Chinese> (1,205), <Japanese> (1,024), and <pizza> (997). An overall view on the similarity structure can be seen in Figure 1, which shows the result of an agglomerative hierarchical clustering of the food labels (not just the restaurant labels in the combined data).<sup>17</sup> This classification uses the average clustering method applied to all labels that contain five or more members, in which each step joins two clusters if the average distance between the members of the clusters is smaller than any other possible combinations of clusters at that point of the classification.<sup>18</sup> The numbers at the top of the figure indicate the average distance between the members in that joint cluster. Consider the branch at the bottom of this figure. At the first level it combines <tea rooms> with the set consisting of <Moroccan>, <Turkish>, <Middle Eastern>, <Greek> and <Mediterranean>. At the next branching point, it breaks out <Moroccan>, and so forth. The main branch ends with the pair <Greek> and <Mediterranean>, which are very close according to this analysis.

## Results

Our analysis of restaurant reviews is complicated by the fact that nearly three quarters of the restaurants in the sample get categorized in only one restaurant genre. These genre specialists do not provide much relevant information because secondary contrast is not defined for them and distance does not enter the pic-



**Figure 1:** Hierarchical clustering of the food genres (“average distance” method—see text)

ture for producers associated with only one label. So our proposed revision to conceptualization and measurement will not make a difference in such cases.

Table 6 reports estimates of comparable specifications that use different measurements of niche width, i.e., for the discrete approach (where  $NW$  refers to  $W_{DB}$ ), and for the metric approach with binary label assignments (where  $NW$  refers to  $\tilde{D}_{MB}$ , the normed measure given by Equation 16). (Estimates for control variables and cut-points can be found in the online supplement.) In both specifications, the effect of maximum contrast is positive and significant, as predicted.<sup>19</sup> (We do not estimate an effect of secondary contrast because it is not defined for the many specialists.) But the effect of niche width is also positive and significant, which runs against hypothesis 1. In earlier research with a subset of these data (with discrete measures), Kovács and Hannan (2010) found that the effect of niche width varies greatly with price, so we report estimates with such interactions in. In each case there is a very substantial and significant improvement in fit over the models with a single parameter for the niche width effect. Nonetheless we get a difficult-to-interpret non-monotonic pattern of niche width by price. These results do not provide support for hypothesis 1.

We think that the predominance of specialists obscures the core issues, for the reasons stated above. So we also estimated specifications parallel to those just discussed for the multiple-category restaurants only. We calculated likelihood-ratio tests of joint hypothesis that the single-category cases can be pooled with the others (that there are no interactions between having  $NW = 0$  and the other covariates). Specifically we compared the fit of the pooled model in column (5) with the fits of the model in column (7) and a comparable model for specialists with the same covariates as in column (5). The likelihood-ratio test statistic, which is distributed as chi-square with 10 degrees of freedom, equals 756 with  $p < 0.001$ . The pooling hypothesis is rejected decisively.

Given the result of the test just discussed, the most relevant results for our purposes come from comparisons of the estimates reported in columns (4) and (8). Again the metric estimates fit better than the discrete: the difference in log-likelihoods is 483. In this case the metric measure also provides a model with a slightly lower prediction error (1.219 versus 1.217).

The effect of maximum contrast is again positive and significant in both sets of estimates. Now we can also estimate the effect of secondary contrast. It turns out to be positive and significant (opposite the prediction in hypothesis 2b) in column (4), but negative and not significant in column (8). So hypothesis 2b is not supported.

The biggest differences concern the effects of niche width by price level. In column (4), the effect is positive at all four price levels, and the effect is significant for three of them. In sharp contrast, the metric measures produce a monotonic pattern: the effect of niche width is positive for “cheap” restaurants and negative for all higher price levels, and the effect gets more negative at each higher price level. Here the metric measurement yields a much more interpretable pattern: niche width matters, and the way it matters depends on price. At the “cheap” level, consumers appear to value a broad niche, consistent with the popularity of “food courts” in many takeaway food venues. However, as the stakes rise, consumers place more value on genre focus. Indeed this effect is particularly strong at the

**Table 6:** Effects of contrast and niche width on the appeal of restaurants: ML estimates of ordered-logit specifications with reviewer random-effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measure of NW	$W_{DB}$	$W_{DB}$	$W_{DB}$	$W_{DB}$	$W_{MB}$	$W_{MB}$	$W_{MB}$	$W_{MB}$
Price level 2	-0.277 <sup>†</sup> (0.005)	-0.291 <sup>†</sup> (0.006)	-0.221 <sup>†</sup> (0.010)	0.203 <sup>†</sup> (0.079)	-0.281 <sup>†</sup> (0.005)	-0.281 <sup>†</sup> (0.006)	-0.216 <sup>†</sup> (0.010)	1.990 <sup>†</sup> (0.074)
Price level 3	-0.135 <sup>†</sup> (0.007)	-0.193 <sup>†</sup> (0.009)	0.010 (0.012)	0.763 <sup>†</sup> (0.102)	-0.138 <sup>†</sup> (0.007)	-0.181 <sup>†</sup> (0.009)	0.003 (0.012)	2.58 <sup>†</sup> (0.092)
Price level 4	0.366 <sup>†</sup> (0.014)	0.430 <sup>†</sup> (0.016)	0.162 <sup>†</sup> (0.030)	0.349 (0.248)	0.366 <sup>†</sup> (0.014)	0.444 <sup>†</sup> (0.016)	0.166 <sup>†</sup> (0.030)	3.86 <sup>†</sup> (0.301)
NW	0.041 <sup>†</sup> (0.009)		0.632 <sup>†</sup> (0.058)		0.014 <sup>†</sup> (0.005)		0.073 <sup>*</sup> (0.037)	
NW × price 1		-0.054 <sup>†</sup> (0.017)		1.38 <sup>†</sup> (0.130)		-0.008 (0.010)		2.23 <sup>†</sup> (0.075)
NW × price 2		0.044 <sup>†</sup> (0.011)		0.563 <sup>†</sup> (0.072)		0.004 (0.007)		-0.442 <sup>†</sup> (0.048)
NW × price 3		0.241 <sup>†</sup> (0.020)		-0.071 (0.144)		0.125 <sup>†</sup> (0.012)		-0.877 <sup>†</sup> (0.080)
NW × price 4		-0.590 <sup>†</sup> (0.060)		1.01 <sup>*</sup> (0.447)		-0.401 <sup>†</sup> (0.036)		-2.11 <sup>†</sup> (0.334)
Max. contrast	0.268 <sup>†</sup> (0.024)	0.261 <sup>†</sup> (0.024)	0.228 <sup>†</sup> (0.055)	0.200 <sup>†</sup> (0.055)	0.150 <sup>†</sup> (0.019)	0.137 <sup>†</sup> (0.019)	0.279 <sup>†</sup> (0.045)	0.240 <sup>†</sup> (0.046)
Sec. contrast			0.129 <sup>†</sup> (0.044)	0.152 <sup>†</sup> (0.045)			0.039 (0.035)	-0.053 (0.035)
Sample	Full	Full	$W_{DB} > 0$	$W_{DB} > 0$	Full	Full	$W_{MB} > 0$	$W_{MB} > 0$
Observations	766,229	766,229	268,996	268,996	766,229	766,229	268,996	268,996
No. of reviewers	59,441	59,441	40,938	40,938	59,441	59,441	40,938	40,938
Mean sq. pred. error	1.242	1.242	1.219	1.219	1.242	1.242	1.219	1.217
Log-likelihood	-1,080,027	-1,080,027	-377,918	-377,890	-1,080,471	-1,080,182	-377,974	-377,407

Note: Standard errors in parentheses.  
<sup>†</sup>  $p < 0.01$ ; \*  $p < 0.05$ .



“splurge” level. It remains to be seen if other research finds a similar pattern. For present purposes what matters most is that getting this pattern requires taking conceptual distances into account.

Popular and more expensive restaurants and food stands receive higher ratings; activist reviewers give lower ratings. Surprisingly, being categorized with a non-food label has a small positive effect on ratings. This result runs against the spirit of the category-spanning story. It is worth further investigation.

## Discussion

We began with a general pattern emerging in contemporary research: objects categorized in multiple genres suffer diminished appeal in the eye of audience members. We argued that better understanding of the consequences of spanning requires consideration of the structure of the underlying conceptual space. We made two arguments. First, the sociocultural distances between the genres being combined affects evaluations: the less similar are the spanned genres, the more confusing is the identity of the object. Such confusion lowers appeal to audience members. Second, the contrasts of categories also influence the consequences of combination. Because high-contrast categories come with stronger expectations and norms, we expect that combining high-contrast categories leads to more confusion and therefore to lower appeal to audience members.

Studying these propositions empirically required rethinking how multiple-category membership influences objects' appeal to audience members. We took as a starting point the approach of Hsu et al. (2009), which by now has become a widespread approach to the study of multiple memberships. This approach assumes that, because the schemas for the various labels differ, organizations that get assigned only one label generally fit to that genre schema better than do those that get assigned two or more labels. Therefore, Hsu et al. (2009) argue that typicality in each genre falls as more labels get assigned. They propose that typicalities can be measured simply in terms of the count of labels assigned.

Our goal to take conceptual space into account led us to rework this approach. We proposed that distances in conceptual space be built directly into the categorization function in a way that yields lower typicalities for objects that combine more distant concepts. We then use the new measure of typicality to calculate contrast. Thus we arrive at a measure of contrast that builds on the agents' conceptual structure of the domain. We also introduced two alternative measures to niche width that incorporate the similarity structure of the concepts.

In the empirical part of the article we analyzed customers' evaluations of books and restaurants. We found strong support for two theoretical propositions in the book setting and mixed support in the restaurant setting. Importantly, we found that the proposed measures provide better model fits than the approach used previously, suggesting that they provide a more precise description of the data. In other words, incorporating information about the structure of conceptual spaces indeed leads to a better understanding of the consequences of genre spanning.

Our proposed approach could be useful in diverse empirical settings. Organizational examples include law firms (some practices are closer to others); wineries

(looking at different blends of varieties to calculate distances among varieties); movies (certain movie genres are closer to others); and financial organizations such as hedge funds (some stocks and financial instruments are closer to each other than others, thus hedge funds differ in their focus). Other possible applications include genre combination in work (Leahey, 2007) innovation (Wezel et al., 2014), and culture (Goldberg et al., 2014b).

More generally, we believe that taking conceptual distances into account matters not only for issues in organizational and economic sociology but for other domains of sociological research as well. For example, in cultural sociology, the question of “omnivore” consumers has attracted considerable interest (e.g., Peterson (1992); Goldberg (2011); Goldberg et al. (2014b)), but extant research has not taken distance in conceptual space into account when measuring omnivorousness. Rather, researchers typically take a list of the genres consumed by the actor and measure omnivorousness with the number of genres consumed and/or appreciated. The results in the current article suggest that such practice might lead to erroneous conclusions by failing to take distances into account. For instance, a person who consumes <opera> and <blues> is more omnivore than a person who consumes <opera> and <chamber music>. Yet a measure that only counts the number of genres consumed will miss this distinction. The framework developed in the current article suggests using an approach that incorporates genre distances into the measure of omnivorousness. A possible approach could be to measure omnivorousness with the niche width of the consumption portfolio of the consumer. We leave for future research the exploration of this idea and its consequences for cultural sociology.

A potential limitation of the current article is the assumption that audience members share the same cognitive schemata and conceptual space. For instance, in the case of restaurants, we have assumed that the <Italian> and <Mexican> cuisines are perceived to be similar to the same extent by all restaurant patrons in San Francisco, and, as a consequence, restaurants that span these two cuisines would similarly confuse patrons. Relatedly, we assumed that audience members perceive the same contrasts. These assumptions are likely to be only partially true (Goldberg, 2011). Different audience members have different tastes and preferences, and they might have been exposed to a different set of genres and genre combinations. While we do not address this possible diversity in this article, future research can build on our approach in investigating the implications of heterogeneity in audience members’ perception of conceptual structure. Ideally, one would build a dynamic understanding of audience members’ tastes, similarity perceptions, and sampling behaviors.

Using labels to infer category membership could be more tenuous than we have assumed. For instance, multiple labels do not differentiate between <food-court> and <fusion> situations (Baron, 2004). That is, a <Mexican>/<French> restaurant might list Mexican dishes on one side of its menu and French dishes on the other side; or it might serve only dishes that fuse elements of the two cuisines. These are qualitatively distinct forms of spanning, and these restaurants would attract different audiences. We cannot, however, tell these cases apart in our data. Future work is needed to address this distinction both theoretically and empirically, for

example, by analyzing restaurant menus—see Kovács and Johnson (2014). Relatedly, we argued that spanning restaurants and books get devalued because they confuse audience members. It is possible that in certain cases the fact that the restaurant or the book is classified into multiple genres actually clarifies the expectations of audience members. For example, it could be the case that a restaurant that is classified as <vegan> and <sushi> only serves vegetarian sushi. To explore this possibility, we collected the menus of multiple spanning restaurants. Although we did not conduct a formal analysis of the menus, it seems to us that spanners are more likely to offer a wide range of dishes than a specific set of dishes that lie in the overlap between the listed categories.

The measures proposed for estimating distances in conceptual spaces could aid researchers in studying the evolution of genres over time. Genres that are distant in one time period might move closer in subsequent ones. Adjusting for the dissimilarity of genres seems particularly useful when the conceptual structure is in flux. Indeed the finding that distance among concepts and the contrast of the associated categories influence objects' evaluation might have interesting dynamic consequences. If objects in low-contrast categories are more likely to get categorized in multiple genres, then the contrasts of these categories further decrease, and the contrasts of crisper categories would further increase or at least remain stable. Pontikes and Hannan (2014) find evidence of such a pattern in the software industry. These processes would imply a tendency toward the macro-level polarization of categories' contrasts.

Another possibly interesting dynamic links genre similarity and the distances among them. On one hand, distance affects the prevalence of spanning. Genre combination, however, influences how audiences perceive the distance among genres: research in cognitive psychology and linguistics show that categories that tend to occur together are perceived to be similar (Church and Hanks 1990). This feedback loop between spanning and genre distances implies a polarization of distances between genres: initially similar pairs will get combined more often, thereby increasing their similarity; and dissimilar pairs will rarely be combined, keeping their similarity low.

The arguments we developed and tested in this article relate to the general reactions to genre spanning. This is not to say that we think more specific argumentation cannot be developed to predict cases in which genre spanning could be beneficial. A potentially useful line of argument builds on the middle-status conformity argument (Phillips and Zuckerman, 2001). Producers with a strong organizational identity (net of the categorical identities that apply) likely face weaker pressures to conform to audience expectations. If so, then spanners with high status and high visibility might face little penalty for spanning distant categories so long as the spanning is consistent with the individual identity. Using this kind of argument in empirical research requires *a priori* identification of the contours of individual identities, so a strategic move in this direction requires greater knowledge about audience schemas for individual objects than has been available in any research we know.

Another potential extension could be to use the proposed approach to study the diversity of organizations entering a given category (McKendrick and Carroll 2001;

Carnabuci et al. forthcoming). The process of legitimation would likely change if *de alio* entrants come from a set of industries that are close versus distant from one another. Future research might look at how, when novel forms first emerge, their distance relative to others affects legitimation (Ruef 2000).

Future research could also explore alternative measurement approaches. For example, one could put more emphasis on the analysis of audience structures and on the taste of audience members regarding genre spanning by using the novel approach of relational class analysis (Goldberg, 2011). Future research that uses review data should also consider the selection problems that arise in such data: while this article analyzed how genre spanning influences whether the restaurants get high or low ratings, we did not model the chance that these restaurants are visited or reviewed. We suspect that genre spanning does influence selection, but our data do not allow us to investigate this question further.

## Notes

- 1 As explained above, *concepts*, *categories*, and *genres* do not exactly have the same meaning (see Goldberg et al. (2014a)). However, both in the current literature and in everyday life these terms are used interchangeably because these terms *usually* go together. As being too pedantic about differentiating the usage of *concepts*, *categories*, and *genres* would result in awkward prose, throughout the text we use these terms interchangeably unless we explicitly want to evoke a specific meaning, in which cases we note this.
- 2 Two studies find that investors are much less sensitive to genre combination than members of general audiences (Pontikes, 2012b; Smith, 2011).
- 3 We restrict attention to the concepts and categories in the same domain, e.g., cuisine or banking. In general, if sets of concepts are seen as unrelated, then establishing membership in more than one does not seem to cause problems.
- 4 We build on the recent body of research that considers the interface between two roles: producer and audience. Incumbents of the producer role create offers; incumbents of the audience role inspect, evaluate, and consume the offers.
- 5 Obviously we would prefer to have access to data that tell what schemas audience members associate with the relevant labels. Then categories could be represented as sets in a space of the values of conceptually relevant features and relations. Questions about combining genres could then be addressed in terms of positions in the feature space.
- 6 In an important study on which we build, Hsu (2006) analyzed data from multiple intermediaries to specify a generalized label function. Her research focused on consensus among the intermediaries as indicative of typicality.
- 7 One of the datasets that we analyze actually contains a restaurant with this pair of labels.
- 8 For a detailed discussion of alternative similarity and dissimilarity measures, see Batagelj and Bren (1995). Some preliminary results show that the main findings of this paper apply to other measures as well, but we leave this direction of investigation for further research.
- 9 In the usual language of logic and linguistics, the extension of a label refers to the set of objects that bear the label.
- 10 We refer to *categorical* niche width (not genre niche width) because what we can observe are the categorical assignments.

- 11 A formal version of the argument can be found in Kovács and Hannan (2013).
- 12 The ratings on both sites represent the overall satisfaction of users with the book or restaurant. Admittedly, the rating might relate to multiple dimensions of value, such as food quality or service for restaurants (Kovács and Johnson (2014) found that the most important dimension is food quality), and the quality of the writing or the reputation of the author in the case of books. We do not attempt to disentangle these dimensions here. Rather, we take a constructivist stance and leave it for the user to decide what dimensions are important for her.
- 13 We do not estimate models with fixed effects because ML estimators for ordered logit models with fixed effects are not consistent (Greene, 2004).
- 14 We intentionally excluded January because the system uses January 1 as a default if no exact day is specified for publishing date.
- 15 Kovács and Hannan (2010) reported estimates of effects of niche width and contrast on appeal—but without adjusting for distances—for a subset of the establishments that we analyze here.
- 16 Yelp.com curates the categorization of restaurants. The owner or the patron who first adds the restaurant can choose from a drop-down list of all possible genres. This label assignment, before it goes online, has to be approved by the maintainers of yelp.com. Later on any user or the owner can submit a modification request. For example, if a restaurant is labeled as <Italian> but a patron finds that the restaurant also serves sushi, then she can suggest that yelp.com change the categorization. Nonetheless, the editors at yelp.com need to approve all changes in categorization.
- 17 We also explored various multidimensional scaling representations of the similarity structure of labels (Kruskal and Wish, 1978), and we found that multidimensional scaling does not represent the structure of the data well. We obtained high stress measures (37.9 for the two-dimensional solution, and 28.4 for the three-dimensional solution), which indicates that the underlying similarity structure is not metric. Indeed, the hierarchical clustering figure supports this conclusion as well.
- 18 We used the hclus procedure in R with the average distance option. We measured distance using Equation (10) with the distance between pairs with similarity of zero (nonintersecting extensions) set to 0.001.
- 19 Note that the log-likelihood increases by 72 in the metric approach, although the mean squared prediction errors are essentially the same.

## References

- Baron, James N. 2004. "Employing Identities in Organizational Ecology." *Industrial and Corporate Change* 13:3–32. <http://dx.doi.org/10.1093/icc/13.1.3>.
- Batagelj, Vladimir and Matevz Bren. 1995. "Comparing Resemblance Measures." *Journal of Classification* 12:73–90. <http://dx.doi.org/10.1007/BF01202268>.
- Carnabuci, Gianluca, Elisa Operti, and Balázs Kovács. forthcoming. "Categorical imperative and structural reproduction: Insights from the global semiconductor industry." *Organization Science*.
- Carroll, Glenn R., Mi Feng, Gaël Le Mens, and David G. McKendrick. 2010. "Organizational Evolution with Fuzzy Technological Formats: Tape Drive Producers in the World Market, 1951–1998." *Research in the Sociology of Organizations* 31:203–234. [http://dx.doi.org/10.1108/S0733-558X\(2010\)0000031009](http://dx.doi.org/10.1108/S0733-558X(2010)0000031009).

- Carroll, Glenn R. and Dennis R. Wheaton. 2009. "The Organizational Construction of Authenticity: An Examination of Contemporary Food and Dining in the U.S." *Research in Organizational Behavior* 29:255–282. <http://dx.doi.org/10.1016/j.riob.2009.06.003>.
- Cerulo, Karen A. 2006. *Never Saw It Coming: Cultural Challenges to Envisioning the Worst*. Chicago, IL: University of Chicago Press. <http://dx.doi.org/10.7208/chicago/9780226100296.001.0001>.
- Chater, Nick and Paul M. B. Vitányi. 2003. "The Generalized Universal Law of Generalization." *Journal of Mathematical Psychology* 47:346–369. [http://dx.doi.org/10.1016/S0022-2496\(03\)00013-0](http://dx.doi.org/10.1016/S0022-2496(03)00013-0).
- Church, Kenneth W. and Patrick Hanks. 1990. "Word Association Norms, Mutual Information, and Lexicography." *Computational Linguistics* 16:22–29.
- DiMaggio, Paul J. 1997. "Culture and Cognition." *Annual Review of Sociology* 23:263–287.
- Durkheim, Emile and Marcel Mauss. 1969 [1903]. *Primitive Classification*. London: Routledge.
- Gärdenfors, Peter. 2004. *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press.
- Garfinkel, Harold. 1968. *Studies in Ethnomethodology*. Cambridge, UK: Polity Press.
- Goldberg, Amir. 2011. "Mapping Shared Understandings Using Relational Class Analysis: The Case of the Cultural Omnivore Reconsidered." *American Journal of Sociology* 116:1397–1436. <http://dx.doi.org/10.1086/657976>.
- Goldberg, Amir, Michael Hannan, Greta Hsu, Balázs Kovács, Gaël Le Mens, Giacomo Negro, László Pólos, Elizabeth Pontikes, and Amanda Sharkey. 2014a. "Conceptual Spaces and Label Spaces: A Geometric Model." Unpublished Manuscript.
- Goldberg, Amir, Michael T. Hannan, and Balázs Kovács. 2014b. "What Does It Mean to Span Cultural Boundaries." Presented at the Analytical Sociology Conference, Mannheim.
- Greene, William. 2004. "The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects." *Econometrics Journal* 7:98–119. <http://dx.doi.org/10.1111/j.1368-423X.2004.00123.x>.
- Griswold, Wendy. 1987. "The Fabrication of Meaning: Literary Interpretation in the United States, Great Britain, and the West Indies." *American Journal of Sociology* 92:1077–1117. <http://dx.doi.org/10.1086/228628>.
- Griswold, Wendy, Terry McDonnell, and Nathan Wright. 2005. "Reading and the Reading Class in the Twenty-First Century." *Annual Review of Sociology* 31:127–141. <http://dx.doi.org/10.1146/annurev.soc.31.041304.122312>.
- Hampton, James A. 2007. "Typicality, Graded Membership, and Vagueness." *Cognitive Science* 31:355–383. <http://dx.doi.org/10.1080/15326900701326402>.
- Hannan, Michael T. 2010. "Partiality of Memberships in Categories and Audiences." *Annual Review of Sociology* 36:159–181. <http://dx.doi.org/10.1146/annurev-soc-021610-092336>.
- Hannan, Michael T., László Pólos, and Glenn R. Carroll. 2007. *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton, NJ: Princeton University Press.



- Hsu, Greta. 2006. "Jacks of All Trades and Masters of None: Audiences' Reactions to Spanning Genres in Feature Film Production." *Administrative Science Quarterly* 51:420–450.
- Hsu, Greta, Michael T. Hannan, and Özgeçan Koçak. 2009. "Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests." *American Sociological Review* 74:150–169. <http://dx.doi.org/10.1177/000312240907400108>.
- Hsu, Greta, Michael T. Hannan, and László Pólos. 2011. "Typecasting, Legitimation, and Form Emergence: A Formal Theory." *Sociological Theory* 29:97–123. <http://dx.doi.org/10.1111/j.1467-9558.2011.01389.x>.
- Jaccard, Paul. 1901. "Étude Comparative de la Distribution Florale dans une Portion des Alpes et des Jura." *Bulletin de la Société Vaudoise des Sciences Naturelles* 37:547–579.
- Johnston, Josée and Shyon Baumann. 2007. "Democracy Versus Distinction: A Study of Omnivorousness in Gourmet Food Writing." *American Journal of Sociology* 113:165–204. <http://dx.doi.org/10.1086/518923>.
- Koçak, Özgeçan, Michael T. Hannan, and Greta Hsu. 2014. "Emergence of Market Orders: Audience Interaction and Vanguard Influence." *Organization Studies* 35:765–790. <http://dx.doi.org/10.1086/518923>.
- Kovács, Balázs and Michael T. Hannan. 2010. "The Consequences of Category Spanning Depend on Contrast." *Research in the Sociology of Organizations* 31:175–201. [http://dx.doi.org/10.1108/S0733-558X\(2010\)0000031008](http://dx.doi.org/10.1108/S0733-558X(2010)0000031008).
- Kovács, Balázs and Michael T. Hannan. 2013. "Category Spanning, Distance, and Appeal." Research Report 2081R, Stanford Graduate School of Business.
- Kovács, Balázs and Rebeka Johnson. 2014. "Contrasting Alternative Explanations for the Consequences of Category Spanning: A Study of Restaurant Reviews and Menus in San Francisco." *Strategic Organization* 12:7–37. <http://dx.doi.org/10.1177/1476127013502465>.
- Kruskal, Joseph B. and Myron Wish. 1978. *Multidimensional Scaling*. Beverly Hills, CA: Sage.
- Leahey, Erin. 2007. "Not by Productivity Alone: How Visibility and Specialization Contribute to Academic Earnings." *American Sociological Review* 72:533–561. <http://dx.doi.org/10.1177/000312240707200403>.
- Leung, Ming and Amanda Sharkey. 2014. "Out of Sight, Out of Mind: Evidence of Perceptual Factors in the Multiple-Category Discount." *Organization Science* 25:171–184. <http://dx.doi.org/10.1287/orsc.2013.0828>.
- Leung, Ming D. 2011. "Apples to Oranges: How Category Overlap Facilitates Commensuration in an Online Market for Services." Presented at the Meetings of the Academy of Management, San Antonio, TX.
- Leung, Ming D. 2014. "Dilettante or Renaissance Man? How the Order of Job Experiences Affects Perceptions of Ability in an External Labor Market." *American Sociological Review* 79:136–158. <http://dx.doi.org/10.1287/orsc.2013.0828>.
- McKendrick, David G. and Glenn R. Carroll. 2001. "On the Genesis of Organizational Forms: Evidence from the Market for Disk Drive Arrays." *Organization Science* 12:661–683. <http://dx.doi.org/10.1287/orsc.12.6.661.10083>.



- Negro, Giacomo, Michael T. Hannan, and Hayagreeva Rao. 2010a. "Categorical Contrast and Audience Appeal: Niche Width and Critical Success in Winemaking." *Industrial and Corporate Change* 19:1397–1425. <http://dx.doi.org/10.1093/icc/dtq003>.
- Negro, Giacomo, Michael T. Hannan, and Hayagreeva Rao. 2011. "Category Reinterpretation and Defection: Modernism and Tradition in Italian Winemaking." *Organization Science* 22:1449–1463. <http://dx.doi.org/10.1287/orsc.1100.0619>.
- Negro, Giacomo, Özgeçan Koçak, and Greta Hsu. 2010b. "Research on Categories in the Sociology of Organizations." *Research in the Sociology of Organizations* 31:1–35. [http://dx.doi.org/10.1108/S0733-558X\(2010\)0000031003](http://dx.doi.org/10.1108/S0733-558X(2010)0000031003).
- Negro, Giacomo and Ming D. Leung. 2013. "'Actual' and Perceptual Effects of Category Spanning." *Organization Science* 24:684–696. <http://dx.doi.org/10.1287/orsc.1120.0764>.
- Olzak, Susan and Nandini Roy. 2013. "Does Spanning Ideological Categories Affect Terrorist Organizations?" Presented at the Organizational Ecology Conference, Budapest.
- Peterson, Richard A. 1992. "Understanding Audience Segmentation: From Elite and Mass to Omnivore and Univore." *Poetics* 21:243–258. [http://dx.doi.org/10.1016/0304-422X\(92\)90008-Q](http://dx.doi.org/10.1016/0304-422X(92)90008-Q).
- Phillips, Damon J. and Ezra W. Zuckerman. 2001. "Middle-status Conformity: Theoretical Refinement and Empirical Demonstration in Two Markets." *American Journal of Sociology* 107:379–429. <http://dx.doi.org/10.1086/324072>.
- Pontikes, Elizabeth G. 2008. *Fitting In or Starting New? An Analysis of Invention, Constraint, and the Emergence of New Categories in the Software Industry*. Ph.D. thesis, Stanford University, Graduate School of Business.
- Pontikes, Elizabeth G. 2012a. "Fitting In or Starting New? An Analysis of Invention, Constraint, and the Emergence of New Categories in the Software Industry: Invention, leniency, and category emergence in the software industry, 1990–2002." Technical report, University of Chicago Booth School of Business.
- Pontikes, Elizabeth G. 2012b. "Two Sides of the Same Coin: How Ambiguous Classification Affects Multiple Audience Evaluations." *Administrative Science Quarterly* 57:81–118. <http://dx.doi.org/10.1177/0001839212446689>.
- Pontikes, Elizabeth G. and Michael T. Hannan. 2014. "An Ecology of Social Categories." *Sociological Science* 1:311–343. <http://dx.doi.org/10.15195/v1.a20>.
- Rao, Hayagreeva, Philippe Monin, and Rodolphe Durand. 2005. "Border Crossing: Bricolage and the Erosion of Categorical Boundaries in French Gastronomy." *American Sociological Review* 70:968–991. <http://dx.doi.org/10.1177/000312240507000605>.
- Rosch, Eleanor and Carolyn B. Mervis. 1975. "Family Resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 7:573–605. [http://dx.doi.org/10.1016/0010-0285\(75\)90024-9](http://dx.doi.org/10.1016/0010-0285(75)90024-9).
- Rosch, Eleanor H. 1975. "Cognitive Representations of Semantic Categories." *Journal of Experimental Psychology: General* 104:192–233. <http://dx.doi.org/10.1037/0096-3445.104.3.192>.

- Ruef, Martin. 2000. "The Emergence of Organizational Forms: A Community Ecology Approach." *American Journal of Sociology* 106:658–714. <http://dx.doi.org/10.1086/318963>.
- Shepard, Roger N. 1987. "Toward a Universal Law of Generalization for Psychological Science." *Science* 237:1317–1323. <http://dx.doi.org/10.1126/science.3629243>.
- Simmel, Georg. 1978 [1907]. *The Philosophy of Money*. Boston: Routledge.
- Simpson, Edward H. 1949. "Measurement of Diversity." *Nature* 163:688.
- Smith, Edward Bishop. 2011. "Identities as Lenses: How Organizational Identity Affects Audiences' Evaluation of Organizational Performance." *Administrative Science Quarterly* 56:61–94. <http://dx.doi.org/10.2189/asqu.2011.56.1.061>.
- Tenenbaum, Joshua B. and Thomas L. Griffiths. 2002. "Generalization, Similarity, and Bayesian Inference." *Behavioral and Brain Sciences* 24:629–640. <http://dx.doi.org/10.1017/S0140525X01000061>.
- Vaisey, Stephen. 2009. "Motivation and Justification: A Dual-Process Model of Culture in Action." *American Journal of Sociology* 114:1675–1715. <http://dx.doi.org/10.1086/597179>.
- Verheyen, Steven, James A. Hampton, and Gert Storms. 2010. "A Probabilistic Threshold Model: Analyzing Semantic Categorization Data With the Rasch Model." *Acta Psychologica* 135:216–225. <http://dx.doi.org/10.1016/j.actpsy.2010.07.002>.
- Wezel, Filippo, Balázs Kovács, and Gianluca Carnabuci. 2014. "Being There: Patent Class Contrast and the Impact of Technological Innovations." Unpublished manuscript, University of Lugano.
- Widdows, Dominic. 2004. *Geometry and Meaning*. Stanford, CA: CSLI Publications.
- Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Legitimacy Discount." *American Journal of Sociology* 104:1398–1438. <http://dx.doi.org/10.1086/210178>.
- Zuckerman, Ezra W., Tai-Young Kim, Kalinda Ukanwa, and James von Rittmann. 2003. "Robust Identities or Non-entities? Typecasting in the Feature Film Labor Market." *American Journal of Sociology* 108:1018–1075. <http://dx.doi.org/10.1086/377518>.

**Acknowledgements:** We appreciate the helpful comments of Glenn Carroll, Amir Goldberg, Greta Hsu, Gaël Le Mens, Giacomo Negro, László Pólos, and Elizabeth Pontikes. We thank the Università della Svizzera italiana, Stanford Graduate School of Business, the Stanford GSB Faculty Trust, and the John Osterweis and Barbara Ravizzi Faculty Fellowship at the Stanford GSB for generous financial support for this project.

**Balázs Kovács:** Università della Svizzera italiana. E-mail: kovacs@usi.ch

**Michael T. Hannan:** Graduate School of Business, Stanford University. E-mail: hannan@stanford.edu.