Cognition 234 (2023) 105363

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Constructing complex social categories under uncertainty

Alice Xia [b,1,*], Sarah H. Solomon [a,1], Sharon L. Thompson-Schill [a], Adrianna C. Jenkins [a]

[a] Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA, 19104, United States
[b] Department of Cognitive, Linguistics, and Psychological Sciences, Brown University, 190 Thayer Street, Providence, RI 02912, United States

ARTICLE INFO

Keywords:
Social concepts
Conceptual combination
Impression formation
Attribute perceptions
Uncertainty

ABSTRACT

Conceptual combination is the act of building complex concepts from simpler ones. Although research has examined how inferences about compound objects (e.g., *fuzzy chair*) are produced from their constituent concepts, little is known about the combinatorial processes that produce inferences about compound social categories (e.g., *Irish musician*). Using a computational approach, we investigated the relationship between ratings of 25 nationality-occupation combinations and ratings of their constituent concepts along the attribute dimensions of warmth and competence. We found that people incorporate uncertainty into their perceptions of compound social categories. Further, people are more likely to use a linear combination strategy when they are more certain about the attributes of the constituents and less familiar with the combination. Conversely, when social combinations are more familiar, their judged attributes deviate further from the predictions of a combinatorial model and are shared across participants, suggesting that stereotype-based knowledge plays a central role in the representation of complex social groups. Twenty-five non-human animal combinations (e.g., *circus snake*) serve as a comparison and were rated on size and ferocity. We found evidence that familiarity has different effects on the strategies used to combine person concepts and animal concepts, pointing to the possible existence of both common and distinct mechanisms for constructing social and non-social categories.

1. Introduction

You know the old joke that begins, "An Irish musician, a Mexican lawyer, and a Japanese cheerleader walk into a bar…"? Whether you do or not, you are probably able to generate expectations about this situation and the people in it by drawing on prior knowledge (about lawyers, for example, or people from Japan) and by using the impressive feat of human cognition by which people make predictions, draw inferences, and derive meaning from novel events, words, or ideas. Cognitive scientists have long been fascinated, and at times perplexed, by the processes that govern generative thought, in part because even very simple examples highlight how complicated these processes must be. Perhaps the most well-known example was first offered by Jerry Fodor, who asked us to observe that the concept *pet fish* does not in any obvious way inherit the seemingly prototypical features of its constituent parts: the goldfish that one imagines as a good pet fish is quite dissimilar from a prototypical pet (e.g., golden retriever), as well as a prototypical fish (e.g., trout) (Fodor & Lepore, 1996). Yet, as we encounter, interpret, and construct novel concepts, we must make use of, and build upon, the scaffolding of prior knowledge.

Although progress has been made toward understanding how the mind combines object concepts (e.g., "cactus" and "rug") to form new ones ("cactus rug") (Boylan, Trueswell, & Thompson-Schill, 2015; Boylan, Trueswell, & Thompson-Schill, 2017; Estes, 2003; Estes & Jones, 2008; Gagné, 2001; Kenett & Thompson-Schill, 2020; Medin & Shoben, 1988; Murphy, 1988; Smith, Osherson, Rips, & Keane, 1988), less is known about how the mind combines social concepts, including concepts about people. Investigating how people combine social concepts is relevant for at least three main reasons. First, each person can be described as the amalgamation of a multitude of constituent concepts, ranging from those related to demographics (e.g., "woman"; "Pennsylvanian") to those related to social roles ("parent"; "boss") to those related to activities and affiliations ("surgeon"; "libertarian"). Second, judgments of what other people are like have important consequences for social decision-making. For example, members of social groups that are generally perceived to be more warm are offered more during resource allocation decisions in the laboratory and are more likely to receive callbacks for job interviews in field studies (Jenkins,

* Corresponding author at: 190 Thayer Street, Providence, RI 02912, United States.
E-mail address: alice_xia@brown.edu (A. Xia).
[1] Denotes equal contribution.

https://doi.org/10.1016/j.cognition.2022.105363
Received 23 February 2022; Received in revised form 8 December 2022; Accepted 19 December 2022

Karashchuk, Zhu, & Hsu, 2018; Kobayashi, Kable, Hsu, & Jenkins, 2022). Third, there are reasons to not automatically assume that the principles that govern the combination of object concepts also apply to the combination of social concepts. For example, neuroimaging evidence suggests that the brain represents social knowledge differently, or at least separately, from object knowledge (Mitchell, Banaji, & Macrae, 2005), and social concepts may be more abstract or more uncertain than object concepts, on average (Berkay & Jenkins, 2022).

Let us return to the case of the social category “Mexican lawyer”, which we will refer to as a *combined concept*, and its constituent parts, “Mexican” and “lawyer”, which we will refer to as *simple concepts*. There are attributes of a Mexican lawyer that we know with approximate certainty (such as the person’s occupation and likely level of education, as well as the country of origin of the person’s family). However, there are other characteristics of a Mexican lawyer about which we might merely guess, with less certainty, such as the person’s level of intelligence or how likely they would be to have a pet turtle. To make these guesses, we might rely on any of a variety of possible sources of information, including what we believe about members of the separate social categories of Mexicans and lawyers in general (i.e., stereotypes about those groups) and perhaps the attributes of any Mexican lawyers we happen to know. The balance between these two distinct yet complementary strategies for interpreting a combination—extracting information from the simple concepts alone or relying on existing knowledge and beliefs about the combined concept itself—may depend on one’s certainty about the attributes of the simple concepts and on one’s preexisting familiarity with the combination. In particular, it may be under conditions of uncertainty about the attributes of a complex concept (e.g., if one has never considered the idea of a Mexican ballet dancer) that people construct a representation of that concept from its constituents; when they do, uncertainty about the attributes of those constituents ("Mexican", "ballet dancer") may affect the combinatorial process. In this way, (un)familiarity with the combination may affect *when* people build novel complex concepts from simple ones, and (un) certainty about the attributes of each simple concept may affect *how* they do. Examining both sources of information will allow us to better understand the mechanistic role of uncertainty in social conceptual combination.

The study we describe here investigates how people make inferences about the traits of members of multiple social categories under conditions of uncertainty. To do this, we capitalize on the observation that perceptions of other people’s traits can be organized along core attribute dimensions, including their warmth (how good or bad their intentions are toward others) and competence (how capable they are of acting on those intentions) (Fiske, Cuddy, & Glick, 2007). We specifically investigate perceptions of people with various combinations of occupations and nationalities on these dimensions. For comparison, we investigate people’s perceptions of combinations of animal types and the habitats in which they live (e.g., “desert cat”) on the dimensions of size and ferocity, which have been shown to organize the semantic space of animals in a similar fashion (Henley, 1969). Despite the multitude of attributes that may be idiosyncratically associated with particular concepts in a binary fashion (e.g., gavel-wielding or not; long-tailed or not), these dimensional frameworks of person and animal perception make it possible to quantify outputs of the conceptual combination process across a variety of concepts in a common attribute space. We will argue that social categories provide a window through which to observe how the mind constructs complex and potentially novel ideas under uncertainty, in ways that inform scientific understanding of both social cognition and conceptual processing. Additionally, because perceptions of others’ traits are known to play a key role in social decision-making, uncovering the principles that govern how people generate perceptions of members of multiple social categories has implications for our ability to predict how people will treat members of these complex social groups.

1.1. Combining social concepts

In the social world, people frequently make decisions about how to treat others (e.g., which individuals to befriend and which to keep at an arm’s length). In an ideal context, these decisions would be made with perfect information about what another person is like, but such information is rarely directly accessible in the absence of a long history of personal interactions. Instead, people routinely construct inferences about what others might be like based on indirect cues. For example, people might expect a nurse to be both warm and competent, whereas they might expect a surgeon to be highly competent but less warm. These social inference strategies are more complex when multiple pieces of information need to be considered simultaneously.

Across the history of psychology, various ideas have been proposed to describe how different pieces of information about a person might be combined in the mind of a perceiver, ranging from single category dominance (Ho, Sidanius, Levin, & Banaji, 2011; Macrae, Bodenhausen, & Milne, 1995) to equal contribution of both constituents to the combination, which may produce emergent attributes if the combination is surprising (Hutter & Crisp, 2008; Kunda, Miller, & Claire, 1990). Notably, however, despite a documented role for uncertainty in conceptual combination outside the social domain (e.g., Solomon & Thompson-Schill, 2020), the role of uncertainty in the combination of social concepts has not been directly examined. This is particularly surprising given that inferences about other people are generally characterized by high uncertainty (Berkay & Jenkins, 2022) and uncertainty has been treated as a critical variable in adjacent areas of research, e.g., social decision-making (FeldmanHall & Shenhav, 2019). We provide a brief overview of existing models of social combinatorial processing below and demonstrate how our approach, which explicitly examines uncertainty as a variable in the construction of social combined concepts, addresses gaps in previous work.

1.1.1. Category dominance models

Social categories are considered richer and more complex compared to categories of physical objects, which are mostly defined by form and function (Cantor & Mischel, 1979; Dahlgren, 1985; Gelman & Spelke, 1981). When a person falls into multiple social categories, one possible strategy for addressing this complexity may be to simply overweight one category by inhibiting activation of the others, thus preserving the economical function of using category knowledge to form an overall evaluation without relying on integration (Macrae et al., 1995). For example, hypodescent refers to the phenomenon by which mixed-race individuals tend to be perceived as members of the lower status race (e.g., Ho et al., 2011). This form of “all-or-none” categorization may maximize both the ability to differentiate between categories and to judge similarity within them (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976).

It has been proposed that all-or-none categorization can be automatically triggered by salient social dimensions such as age, race, and sex (Allport, 1979). However, others later showed that these dimensions facilitate the categorization of unfamiliar faces but not familiar (i.e., famous) faces (Quinn, Mason, & Macrae, 2009). Category-level knowledge therefore might be most useful when individuating information is absent. While these findings imply an interaction between category-level knowledge and other factors such as prior experience or familiarity, which could in turn increase uncertainty, the exact nature of that relationship remains unclear.

1.1.2. Integration models

Dating back at least as far as Asch (1946), questions have arisen concerning whether an impression of another person is (or is not) more than the sum of its parts. Two pieces of information can be integrated in a variety of ways, but perhaps the most basic integrative model is an additive one, in which individual cues are linearly combined to predict the meaning of those cues in a combination (e.g., Mitchell & Lapata, 2010; Smith, Patalano, & Jonides, 1998). In the social domain, additive models define a perceiver's integrated impression of a target as the sum of each social cue multiplied by a weighting factor (Rokeach & Rothman, 1965). Although initially influential, substantial evidence now converges on the idea that perceptions of others are not additive. For example, one early study showed that evaluations of a combined concept like "immoral priest" do not reduce to a simple aggregation of the features evoked by the individual concepts "immoral" and "priest" (Rokeach & Rothman, 1965). Similarly, additive models may be less successful when the simple concepts to be combined are more incongruent. For example, these models might be able to predict the perceptions of congruent combinations such as "intelligent lawyer" but might fail at predicting people's perceptions of incongruent combinations such as "blind lawyer" or "Harvard-educated carpenter", which typically include attributes not associated with any of the constituent concepts alone (Kunda et al., 1990).

Further, simple additive models may be insufficient because the influences of specific social attributes on perceptions of other people are known to be context-dependent. When presented with pairs of types of social cues (e.g., age and race, Kang & Chasteen, 2009; race and sexual orientation, Remedios, Chasteen, Rule, & Plaks, 2011), the effect of those cues on people's perceptions is often interactive rather than additive. For example, "old" may have a negative influence on perceptions of white people, but a positive influence on perceptions of Black people (Kang & Chasteen, 2009). Despite general agreement that overall impressions of others are likely more than the sum of individual pieces of information, attempts to quantify the mechanisms underlying this construction process have largely been absent.

Many kinds of integrative models have been proposed outside of the social domain, but whether these models translate to perceptions of complex social concepts is unclear. In contrast to additive models, multiplicative models combine cues in a nonlinear fashion and are thus integrative in a representational sense. These combinatorial approaches have been used in distributional semantics (Baroni & Zamparelli, 2010; Chang, Cherkassky, Mitchell, & Just, 2009; Mitchell & Lapata, 2010) as well as in neuroimaging approaches to conceptual combination (Baron & Osherson, 2011; Solomon & Thompson-Schill, 2020). For example, in a study of adjective-noun concept combination, Solomon and Thompson-Schill (2020) found that representations of noun concepts (e.g., "rabbit", "molasses") were modulated more by the adjective concepts (e.g., "light", "dark") when uncertainty about the relevant attributes of the noun concept was higher (a rabbit could be either quite light or quite dark) than when it was lower (all molasses is dark). Many cognitive theories of conceptual combination can also be considered integrative since information from two constituent concepts is integrated in some fashion (whether linearly or otherwise) to form the resulting combination. Integration often occurs on the level of specific conceptual features. In property mapping models, for example, a highly salient property of one simple concept is mapped directly onto another simple concept (Wisniewski, 1996) the phrase "tiger hound" might be interpreted as a hound with stripes since stripes are salient features of tigers. In the selective-modification model (Smith et al., 1998), the representation of a combined concept is generated by adjusting the strength of specific features (e.g., a "green apple" is an apple with its *green* feature strengthened). Similarly, the attribute inheritance model integrates features of constituent concepts to create a combined concept (Hampton, 1997, 1998). These theories all share the main assumption that the conjunction of two concepts entails the conjunction of the *attributes* of those concepts, not their category memberships. However, the extent to which attributes of constituent concepts are used to generate interpretations of combined concepts is disputed (Fodor & Lepore, 1998; Connolly, Fodor, Gleitman, & Gleitman, 2007; Gagné & Shoben, 1997).

How do these computational and cognitive models translate to the social domain? In our case, it could be that people understand complex social concepts such as *Mexican lawyer* by integrating the assumed warmth and competence of Mexican people with the assumed warmth and competence of lawyers. When integrating information across social concepts, uncertainty may or may not play a role. Applying existing models of conceptual combination to the social domain will reveal how social concepts are integrated when people consider complex social groups and the conditions under which this integration occurs. It also has the potential to reveal how different sources of uncertainty influence the interpretation of complex social concepts.

1.2. Study aims and predictions

Here we translate existing models of conceptual combination to the social domain to specifically evaluate the extent to which uncertainty influences judgments of combined social concepts. We focus on two sources of uncertainty: (i) the estimated range of possible values of each constituent concept on a given attribute (e.g., the range of possible warmth values of a lawyer and of a Mexican person) and (ii) familiarity with the combined concept (how much experience one has had thinking about and/or interacting with Mexican lawyers). We also directly compare conjunctions of social and non-social concepts to address the generalizability of the cognitive processes that subserve our capacity for constructing and interpreting compound, and potentially novel, concepts.

We use computational models to assess how judgments of complex social concepts (e.g., *Mexican lawyer*) are derived from information about their constituent, simple concepts (e.g., *Mexican* and *lawyer*). Specifically of interest is the extent to which uncertainty associated with each simple concept influences the interpretation of the resulting combination. Warmth and competence are core attributes used to evaluate others and influence aspects of others' social and economic opportunities (Fiske et al., 2007; Jenkins et al. 2018). We thus analyze simple and complex social concepts within a 2D warmth-competence space. As a nonsocial comparison, we explore the same questions for animal habitat-animal species combinations (e.g., *cave pig*) within 2D ferocity-size space since the semantic structure of animals is defined largely by these two attributes (Henley, 1969).

In our additive model, evaluations of complex concepts are predicted by a weighted average of the simple concepts. This reflects a case in which people's perceptions of, for example, the warmth and competence of *Mexican lawyer* is the weighted average of their perceptions of warmth and competence of Mexican people in general and lawyers in general. Predictions borne out of this model assume that the individual components of complex social concepts can be summed in a weighted linear fashion. This would align with previous evidence indicating that social knowledge structures are more amenable to linear separability, since such separability permits greater within-category flexibility (Wattenmaker, 1995).

Our Bayesian models, on the other hand, predict evaluations of complex concepts based on the degree of uncertainty about the individual attributes of the two simple concepts. We capture this uncertainty by quantifying the range of warmth and competence values that participants assign to each social group within the constituent concepts (e.g., *Mexicans*, *lawyers*). Different trait-based probability distributions are then constructed for each simple social concept before they are combined to generate a probabilistic estimate of the warmth and competence of *Mexican lawyer.* The Bayesian models' prediction of a combined concept will be drawn toward whichever constituent concept is associated with greater certainty within each dimension. Our computational goal is to test whether the inclusion of uncertainty in our Bayesian models results in increased performance relative to the additive model. This will reveal whether uncertainty influences how social and nonsocial concepts are combined.

While these combinatorial models illuminate how concept uncertainty contributes to combined concept evaluations, they are not informative about other factors that may influence combinatorial processing, such as prior knowledge about the combinations themselves. In particular, familiarity effects in person perception can be induced through simple repeated exposure to the pairing of a certain behavior with a certain trait, resulting in, e.g., greater favorability for behaviors previously repeatedly associated with the trait “intelligent” (Smith, 1989). Additionally, interpretations of less familiar combinations are not necessarily derived from the most typical properties of their constituents (Connolly, Fodor, Gleitman, & Gleitman, 2007). In particular, exemplar-based models of social judgment claim that such judgments are based on interactions between past experiences with specific individuals and knowledge or assumptions about the social categories to which those individuals belong, which linear attribute-based models of person perception cannot fully explain (Smith & Zarate, 1992). This effect, which has also been demonstrated by nonsocial category exemplars (Holyoak & Koh, 1987; Novick, 1988), raises the possibility that familiarity influences how we flexibly apply social conceptual knowledge. However, the exact role of familiarity in conceptual combination is unclear. We therefore examine how participants’ previous exposure to and direct experience with our social and nonsocial combinations relate to the traits they ascribe to those combinations.

2. Methods

2.1. Participants

A total of 591 participants (66% female; mean age ± SD in years, 44.5 ± 12.8) completed online surveys via Amazon’s Mechanical Turk (AMT), either in the social or nonsocial domain, and were compensated according to standard rates. Participants were eligible if they were located in the United States and had at least 10,000 HITs approved on AMT with at least a 99% HIT approval rate. The research was approved by the University of Pennsylvania Institutional Review Board. Informed consent was obtained for all participants prior to participation. Data and analysis code are available at https://osf.io/s73uk/.

2.2. Stimuli

Within the social domain, we chose five nationality concepts (*Irish, Japanese, Mexican, Moroccan, Russian*) and five occupation concepts (*cheerleader, lawyer, musician, nurse, police officer*), resulting in 10 simple person concepts and 25 combined person concepts (e.g., *Mexican lawyer*). To serve as a nonsocial comparison, we additionally chose five animal habitat concepts (*cave, circus, desert, farm, savanna*) and five animal species concepts (*cat, owl, pig, rat, snake*) resulting in 10 simple animal concepts and 25 combined animal concepts (e.g., *desert cat*). Concepts in each domain were analyzed on two relevant semantic dimensions: social concepts were analyzed in terms of warmth and competence (e.g., Cuddy et al., 2009) and nonsocial concepts were analyzed in terms of size and ferocity (e.g., Henley, 1969).

The final set of simple concepts was selected based on results from a pilot study conducted on AMT in which participants rated 20 simple concepts along the two dimensions of interest either in the social (N = 100) or nonsocial (N = 104) domain. Our goal was to select a set of 10 concepts in each domain that were as decorrelated as possible along our dimensions of interest, while simultaneously covering a sufficient range along those same dimensions in order to capture a representative sample of concepts that spans the relevant attribute space. In our final set of 10 simple social concepts, warmth and competence were not reliably associated ($r = 0.08$, $p > .8$). In our final set of 10 nonsocial concepts, size and ferocity were positively but not reliably associated ($r = 0.61$, $p = .06$).

2.3. Semantic attribute rating tasks for simple and combined concepts

2.3.1. Task

One group of participants ($N = 258$) provided ratings on simple and combined concepts within the social domain. Ratings for simple concepts were collected before ratings for the combined concepts. For each simple concept, we asked participants to provide a range of warmth values and a range of competence values that corresponded to their perception of people within that social category. Warmth was defined as “friendly, caring, and well-intentioned”, and competence was defined as “intelligent, effective, and capable”. Participants were asked to consider the traits of a person based either on nationality or occupation information alone. For each concept, participants indicated the likely range of warmth or competence that they believe captures most of the members within a given social group. Participants selected a minimum and maximum value on a 100-point scale ranging from 1 (“extremely low”) to 100 (“extremely high”), separately for each dimension. The survey questions were organized in four different blocks: nationality-warmth, nationality-competence, occupation-warmth, and occupation-competence. Nationality and occupation concepts were randomized within each block, and the order of the blocks was also randomized.

Combined concept questions were presented after all the simple concept blocks were completed. Each participant rated a subset of 10 combined concepts (five each on warmth and competence separately) out of a total of 25 possible combinations. The subset was randomly selected, but a simple concept would only be shown once per participant per dimension. For example, if a participant were asked to rate the warmth of *Irish Musician*, then they would not be shown *Irish Nurse* or *Japanese Musician*. On each combined concept trial, participants evaluated a person or animal based on two pieces of information: nationality and occupation in the social domain; animal habitat and animal species in the nonsocial domain. Participants provided a point estimate for each combined concept on a single 1–100 numeric scale anchored by the same endpoints as the scales used for the simple concept task. We also collected familiarity ratings for each combined concept (described

Table 1
Summary of participant-generated ratings and measures derived from those ratings.

Measure	Description	Participant generated?	Section
	Simple Concepts		
Attribute Ratings	Min and max values on a 1–100 numeric scale for each attribute dimension	Yes	2.3.1
Mean Attribute Values	Points in 2D attribute space, each one an average of min and max values from Attribute Ratings	No	2.3.2
Attribute Ranges (R)	Absolute difference between min and max values from Attribute Ratings	No	2.3.3
Concept Uncertainty (U)	Product of two simple concepts’ R values in the two attribute dimensions	No	2.5
	Combined Concepts		
Attribute Ratings	Point estimate on a 1–100 numeric scale for each attribute dimension	Yes	2.3.1
Familiarity	Average of three values, each one a point estimate from a 1–100 numeric scale corresponding to one of three different questions	Yes	2.4
Agreement (A)	Sum of the inverse standard deviations of Attribute Ratings of each combined concept, across all participants	No	2.7
Differentiation (D)	[Euclidean distance between combined concept Attribute Rating and center of 2D attribute space] − [Euclidean distance between midpoint of two simple concepts in 2D space and center of 2D attribute space]	No	2.8

below in Section 4). A summary of all the ratings collected from participants and the measures we derived from those ratings is provided in Table 1.

A different group of participants ($N = 242$) provided ratings on simple and combined concepts in the nonsocial domain, in an otherwise identical design to the social task above. Specifically, participants were asked to evaluate an animal given either information about its habitat or its biological classification (at the species level). Ferocity was defined as "dangerous and deadly", and participants responded on a 100-point scale ranging from 1 ("extremely low") to 100 ("extremely high"). Size questions were framed as "how big and massive" an animal was likely to be and the scale ranged from 1 ("extremely small") to 100 ("extremely large").

Additional participants were excluded ($N = 91$) if they did not complete the survey in full, or if their data reflected that they had not followed task instructions (e.g., they entered the same response across all trials).

2.3.2. Attribute values

We used data from the tasks above to represent the simple concepts (e.g., *Mexican*) and combined concepts (e.g., *Mexican lawyer*) in a two-dimensional (2D) space defined by the social or nonsocial attributes of interest (see Figs. 1 and 2). Social concepts and combinations were represented as points in a 2D space in which the dimensions corresponded to warmth and competence; nonsocial concepts and combinations were represented in a 2D space defined by ferocity and size. This approach to conceptual representation is consistent with previous work in which concepts are defined as points in a high-dimensional semantic space (e.g., Baroni & Zamparelli, 2010; Lund & Burgess, 1996; Mitchell & Lapata, 2008, 2010); however, here we focus on two semantic attribute dimensions to make an analysis of conceptual change tractable. Analyzing concepts in 2D space enabled us to examine the computations and strategies used to interpret combined concepts factoring in the two relevant dimensions simultaneously.

For each simple concept (e.g., *Irish*), for each attribute (e.g., warmth), we averaged the minimum and maximum values within each participant to generate a subject-specific scalar estimate, and then averaged these values across participants to determine the concept's value in that attribute dimension. Thus, each social concept was defined by a single warmth value and a single competence value, and each nonsocial concept was defined by a single ferocity value and a single size value (see Fig. 2A).

For each combined concept, each participant provided ratings for the two attributes of interest; we averaged these values across participants in order to define the social and nonsocial combinations in the same 2D spaces that were used to define the individual concepts.

2.3.3. Attribute ranges

For each simple concept and each attribute dimension, we calculated the absolute difference between the minimum and maximum ratings given by each participant and averaged these differences across participants. This resulted in a range magnitude R for each concept and attribute. Each concept's R values were used to generate a measure of constituent uncertainty and were also used in the Bayesian predictive models, described below.

2.4. Combination familiarity

To determine how uncertainty about a combined concept influences its interpretation above and beyond uncertainty about its constituent concepts, we collected familiarity ratings for each combined concept after the attribute ratings were completed. Combination uncertainty was operationalized as the inverse of familiarity (less familiar concepts are likely to be more uncertain). The familiarity measure also accounted for potential novelty effects—while our simple concepts were chosen to be highly familiar across participants, our combined concepts were constructed from all possible pairs of the simple concepts and participants' familiarity with these combinations were unknown. We asked three questions that probed experience with and exposure to the combined concepts and/or their referents. Specifically, we probed personal familiarity (Q1: "How much have you personally observed or interacted with this kind of person/animal (relative to other kinds of people/animals)?"), prior beliefs (Q2: "Did you have an idea of what this kind of person/animal is like, prior to this survey? If so, how strong was your idea of this kind of person/animal (relative to other kinds of people/animals)?"), and frequency of contemplation (Q3: "Did you ever think about this kind of person/animal, prior to this survey? If so, how frequently did you think about this kind of person/animal (relative to other kinds of people/animals)?"). These three measures of familiarity were highly associated with one another (r's > 0.9). We therefore

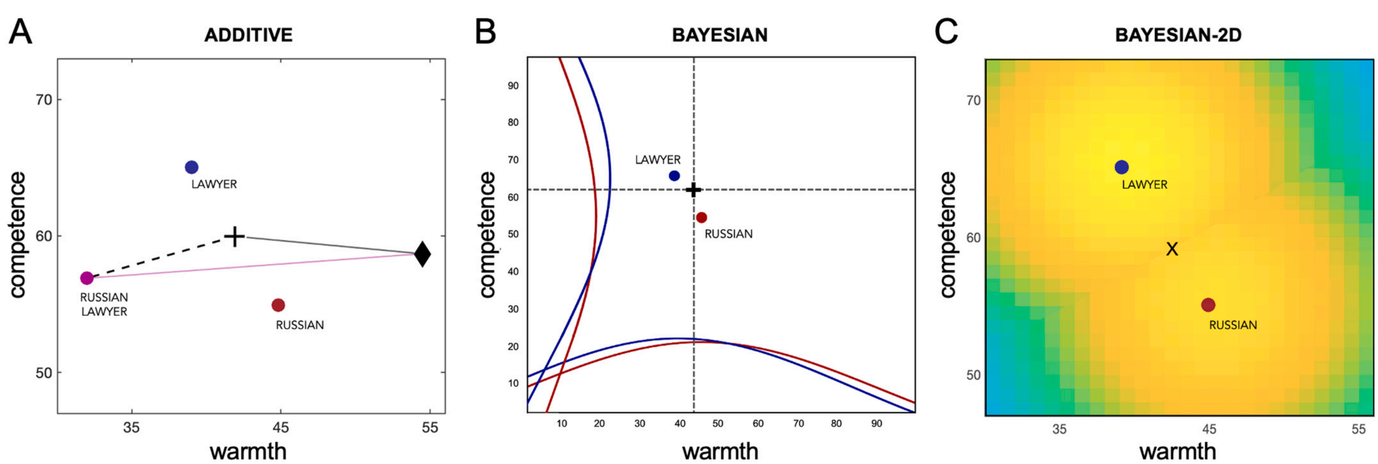

Fig. 1. Additive and Bayesian modeling methods. (A) In the unweighted additive model, the prediction for the combination (+) is midway between C1 (e.g., Russian) and C2 (e.g., lawyer). Error (dashed line) was calculated as the Euclidean distance between the model's estimate and the true combination value (purple dot). The Differentiation measure was calculated by subtracting the distance between the estimate of the unweighted additive model (+) and the center of attribute space (black diamond) from the distance between the true combination and the center of attribute space. (B) In the standard Bayesian model, the predicted value on each dimension was calculated as the maximum a posteriori (MAP) estimate (dashed lines) of the product of the C1 and C2 distributions for that dimension (posterior distribution not shown). Uncertainty in each dimension was captured in the variance of the simple concept distributions. (C) In the 2D-Bayesian model, C1 and C2 were represented as bivariate probability distributions. The model's prediction was derived from the product of the simple concepts' multivariate distributions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

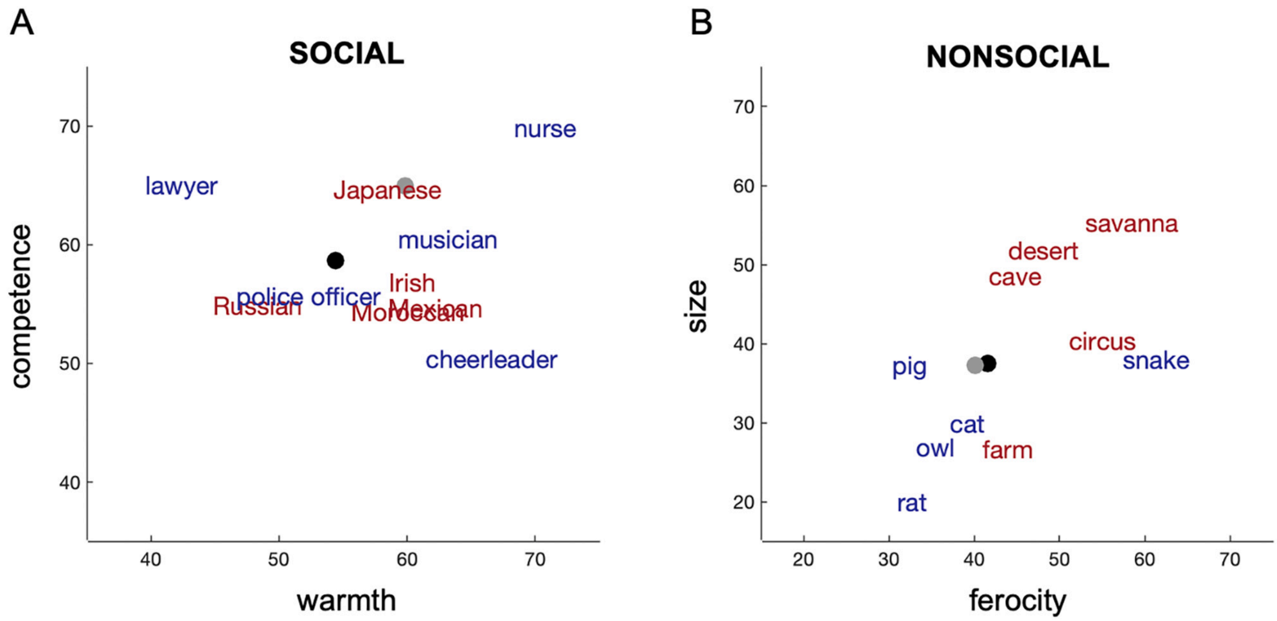

Fig. 2. Social and nonsocial concepts in 2D attribute space. (A) Social concepts. (B) Nonsocial concepts. C1 in red; C2 in blue. Black dot represents the center of the attribute space defined by the simple concepts. Grey dot represents the center of the attribute space defined by the combinations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

averaged the three ratings within each combined concept in order to form a composite familiarity score used in all subsequent analyses.

We also collected similar familiarity ratings for the nonsocial combinations, using the same question format. The three measures were again highly associated with one another (r's > 0.8); we therefore averaged the ratings within each combination to create a composite familiarity score, as we did within the social domain.

2.5. Concept uncertainty

We generated estimates of uncertainty for each concept, for each semantic attribute, using the range magnitude, R, in each relevant attribute dimension. In our 2D conceptual spaces, a concept's uncertainty (U) was represented as an area in this 2D space determined by the product of the concept's R values in the two relevant attribute dimensions. For example, the uncertainty U corresponding to *Irish* was calculated as the product of the range magnitudes in the warmth and competence dimensions:

$$U_{IRISH} = R(warmth)_{IRISH} \bullet R(competence)_{IRISH}$$

Thus, this measure of uncertainty (U) captures the magnitude of the area of 2D space in which each concept could hypothetically be located. Note that this measure does not incorporate the concept's actual location in 2D space, only the degree of uncertainty surrounding that location.

2.6. Predictive models

o understand how combined concepts are interpreted, and whether the relevant processes differ in the social and nonsocial domains, we created a set of predictive models that made different assumptions about how the interpretations of combined concepts are influenced by the consistuent concepts. This approach has been used in distributional semantics, where high-dimensional semantic vectors that capture word co-occurrence statistics are combined in order to generate predicted semantic representations of combined concepts (e.g., Baroni & Zamparelli, 2010; Mitchell & Lapata, 2008, 2010), and can be scaled down to the level of individual semantic dimensions or features (Solomon & Thompson-Schill, 2020). Here, we used predictive models to make predictions about combined concepts in a 2D conceptual space. We analyzed two noncombinatorial models (C1 and C2 models) and two kinds of combinatorial models (Additive, Bayesian).

2.6.1. C1 and C2 models

The noncombinatorial models, analogous to category dominance models mentioned earlier, predict that the representation of the combined concept is identical to either the first (C1) or second (C2) constituent concept. For example, the C1 model predicts that the representation of *Irish musician* will be identical to the representation for *Irish*. D1 and D2 represent values in the first (e.g., warmth) and second (e.g., competence) dimensions:

$$D1, D2_{IRISH\ MUSICIAN} = D1, D2_{IRISH}$$

Conversely, the C2 model predicts that the representation of *Irish Musician* will be identical to the representation for *Musician*:

$$D1, D2_{IRISH\ MUSICIAN} = D1, D2_{MUSICIAN}$$

Neither the C1 nor C2 model is combinatorial, or integrative, because only one of the concepts factors into the prediction for the combination.

Error for each combination was calculated as the Euclidean distance between the model's estimated $D1$, $D2$ coordinates and the true $D1$, $D2$ coordinates for each combination. The errors for the 25 social combinations were squared and then averaged, resulting in a mean squared error (MSE) for each model in the social domain. Error was similarly calculated in the nonsocial domain.

2.6.2. Additive model

The combinatorial Additive model predicts that the representation of the combined concept is a (weighted) sum of the values of the constituent concepts. The unweighted version of this model is as follows:

$$D1, D2_{IRISH\ MUSICIAN} = (5 \bullet D1_{IRISH}) + (5 \bullet D1_{MUSICIAN}), (5 \bullet D2_{IRISH}) + (5 \bullet D2_{MUSICIAN})$$

Here, C1 (Irish) and C2 (Musician) are weighted equally, and the dimension values for each concept are, in effect, averaged together. Weighted versions of this model allow us to examine whether the Additive model performs better when either C1 or C2 contributes more to the predicted combination. The scalar weight of 0.5 in the unweighted example was replaced with parameters W_{C1} and W_{C2}which always sum to 1 and indicate the weights given to C1 and C2, respectively:

$$D1, D2_{IRISH\ MUSICIAN} = (W_{C1} \bullet D1_{IRISH}) + (W_{C2} \bullet D1_{MUSICIAN}), (W_{C1} \bullet D2_{IRISH}) + (W_{C2} \bullet D2_{MUSICIAN})$$

For example, weights of $W_{C1} = .8$ and $W_{C2} = .2$ would result in an estimate for the combined concept that relies more heavily on the representation of C1 relative to C2. We tested the parameter space from $0 \leq W_{C1} \leq 1$, at increments of 0.01. Note that the weighted Additive model at parameter values of $W_{C1} = 0$ and $W_{C1} = 1$ is identical to the C2 and C1 model, respectively. The Additive model's MSE was calculated using Euclidean distance between the true and predicted location in 2D space, separately for social and nonsocial combinations (dashed line in Fig. 1A).

2.6.3. Bayesian models

Combinatorial Bayesian models incorporate the uncertainty of the attributes of the constituent concepts when generating a prediction of the combination (Solomon & Thompson-Schill, 2020). In standard Bayesian approaches, relevant representations are captured in probability distributions within a dimension of interest. Thus, in our standard Bayesian model, warmth and competence were treated as independent variables (Fig. 1B). For each variable, within each simple concept, probability distributions were defined based on the mean value (μ) along with a variance (σ), which we estimated using the R measure described above. The predicted value for each dimension was the maximum a posteriori (MAP) estimate of the product of the C1 and C2 distributions:

$$D1_{IRISH\ MUSICIAN} = arg\ max\, f\ \{P_{\mathrm{D1-IRISH}}(\mu, \sigma) \bullet P_{\mathrm{D1-MUSICIAN}}\ (\mu, \sigma)\ \}$$

In our two-dimensional Bayesian model (Bayes-2D), we translated these standard approaches into 2D space by representing concepts and combinations in terms of multivariate probability distributions (Fig. 1C). In the Bayes-2D model, the two dimensions are no longer treated independently but have a conjoint effect on the model predictions. In this case, each simple concept distribution is defined by a value in each dimension along with a 2D variance estimated using the R measure. For example, the concept *Irish* was represented as:

$$N_{IRISH} = (\mu, \Sigma)$$

$$\mu = [D1_{IRISH}, D2_{IRISH}]$$

$$\Sigma = \begin{bmatrix} R1_{IRISH} & 0 \\ 0 & R2_{IRISH} \end{bmatrix}$$

The Bayes-2D model calculates the product of N_{C1} and N_{C2} and finds the peak ($D1$, $D2$) of the posterior distribution.

In both Bayesian models, constituent uncertainty is reflected in the variance of the respective distributions, and thus is able to influence the resulting posterior distribution. The difference between the standard Bayesian model and the Bayes-2D model is that the former captures D1 and D2 in separate distributions, whereas in Bayes-2D they are jointly captured in a single distribution. The standard Bayesian model can be considered to be a special case of a linear model in which the dimension weights in C1 and C2 are determined by concept-specific uncertainty. Importantly, the Bayesian models differ from the Additive model in that the Bayesian models incorporate uncertainty within each simple concept whereas the Additive model does not. The Bayesian models' MSE was calculated using Euclidean distance between the true and predicted location in 2D space, separately for social and nonsocial combinations.

2.7. Combination Agreement

It is possible that interpretation of combined concepts is driven by knowledge or stereotypes that are shared across people. In order to examine this possibility, we generated a measure of combination Agreement, A, which captures the extent to which ratings of the combined concepts were consistent across people. More specifically, the value of A for a combined concept was calculated as the sum of the inverse standard deviations of ratings in each dimension:

$$A_{IRISH\ MUSICIAN} = \frac{1}{S(warmth)_{IRISH\ MUSICIAN}} + \frac{1}{S(competence)_{IRISH\ MUSICIAN}}$$

where S indicates the standard deviation of ratings across participants. In other words, the value of A increases as the variance across dimension ratings decreases, within a specific combination.

2.8. Combination Differentiation

When predictive models fail at predicting the representations of combined concepts, the reasons for these failures are not provided by the models themselves. Additional analyses are required to understand how the representations of the combinations differ from those of their constituent concepts. One approach, and the one we chose here, is to quantify movement of the combination away from the center of the relevant 2D attribute space. The center of 2D space was calculated as the mean across the 10 simple concepts in both attribute dimensions separately within each domain (black dot in Fig. 1A and 2A). We chose to use the simple concepts to define the center of semantic attribute space because the simple concepts are more basic and familiar relative to the potentially novel combinations. However, we separately calculated the center of combined concept space by taking the mean across the 25 combined concepts in both dimensions within each domain (grey dot in Fig. 2A).

One possibility regarding the interpretations of combined concepts is that, instead of relying on the constituents or a linear combination thereof, people revert to the center of semantic attribute space. In other words, the combination will be interpreted to have the characteristics most typical of the domain if nothing else about the referent is known. Conversely, it could be that the characteristics of the combined concept are differentiated from the prototypical attributes—that is, the characteristics of the combination might become more extreme, or exaggerated, relative to what would be predicted from the constituents alone.

We quantified this kind of movement through 2D space in our measure of combination Differentiation (D). For each combination, we first calculated the Euclidean distance between the center of (simple concept) conceptual space and the midpoint between the C1 and C2 points (i.e., the prediction of the unweighted Additive model). This distance measure captures the degree to which a linear combination of the constituent concepts diverges from the central tendency in the relevant attribute space (grey line in Fig. 1A). We then calculated the Euclidean distance between the center of semantic attribute space and the true (participant-rated) location of the combined concept (purple line in Fig. 1A). Subtracting the former from the latter results in a measure that reflects the degree to which people's judgments of combined concepts are more differentiated (i.e., farther from the center of 2D space) than would be expected given its constituents alone. Note that this Differentiation measure does not reflect the angle of movement through space, merely the magnitude of movement toward or away from the central point.

3. Results

3.1. Attribute ratings in simple and combined social concepts

Mean attribute values for the social and nonsocial concepts are shown in Fig. 2. Consistent with pilot data, attribute values were not correlated across simple concepts in either the social domain (warmth and competence: $r = 0.078$, $p > .8$) or the nonsocial domain (ferocity and size: $r = 0.61$, $p = .063$).

3.1.1. Social combinations are best predicted by the Bayesian model

We compared the ability of the Nationality (C1), Occupation (C2), Additive, and both Bayesian models to predict the attributes of the social combinations. Mean squared error (MSE) across the 25 combinations for each model is shown in Fig. 3A. A one-way ANOVA revealed an overall difference in model performance ($F(124)=4.87$, $p=.001$). Paired dependent *t*-tests indicated that the Bayesian model outperformed the Nationality (C1) model ($t(24)=3.98$, $p < .001$), the Additive model ($t(24)=4.17$, $p < .001$), and the Bayesian-2D model ($t(24)=3.33$, $p=.003$). The difference between the Bayesian model and the Occupation (C2) model was not significant ($t(24)=1.63$, $p=.12$). There are two results to highlight here. First, the fact that the Bayesian model outperformed the Additive model suggests that uncertainty plays a role in how people combine concepts—incorporating uncertainty into the model reduces prediction error (Fig. 3B). Second, the finding that the standard Bayesian model outperformed Bayes-2D suggests that while uncertainty factors into the perception of the combined concept, uncertainties within different dimensions are treated independently. The Additive model outperformed the Nationality (C1) ($t(24)=3.62$, $p=.001$) and Bayes-2D ($t(24)=2.75$, $p=.01$) models; performance of the Additive and Occupation (C2) models did not differ from each other ($p > .7$). Based on these results, we suggest that when interpreting nationality-occupation social combinations, people may perform a linear combination of the perceived nationality and occupation concepts, factoring in uncertainty, or they may rely on the perceived attributes of the occupation concept alone. In comparison, for the nonsocial combinations, a one-way ANOVA revealed that the five models (i.e., Habitat, Species, Additive, Bayesian, Bayes-2D) reliably differed in their performance ($F(124)=2.5$, $p=.046$; Fig. 3A); however, the difference between the Bayesian model and the Additive model was not significant ($t(24)=1.40$, $p= .17$).

In the Additive models reported above, C1 and C2 were weighted equally. Given the relative accuracy of the Occupation (C2) and Additive models in the social domain, we sought to determine whether a weighted combination of the Nationality (C1) and Occupation (C2) models would increase performance of the Additive model. We examined a range of W values and calculated the MSE across the 25

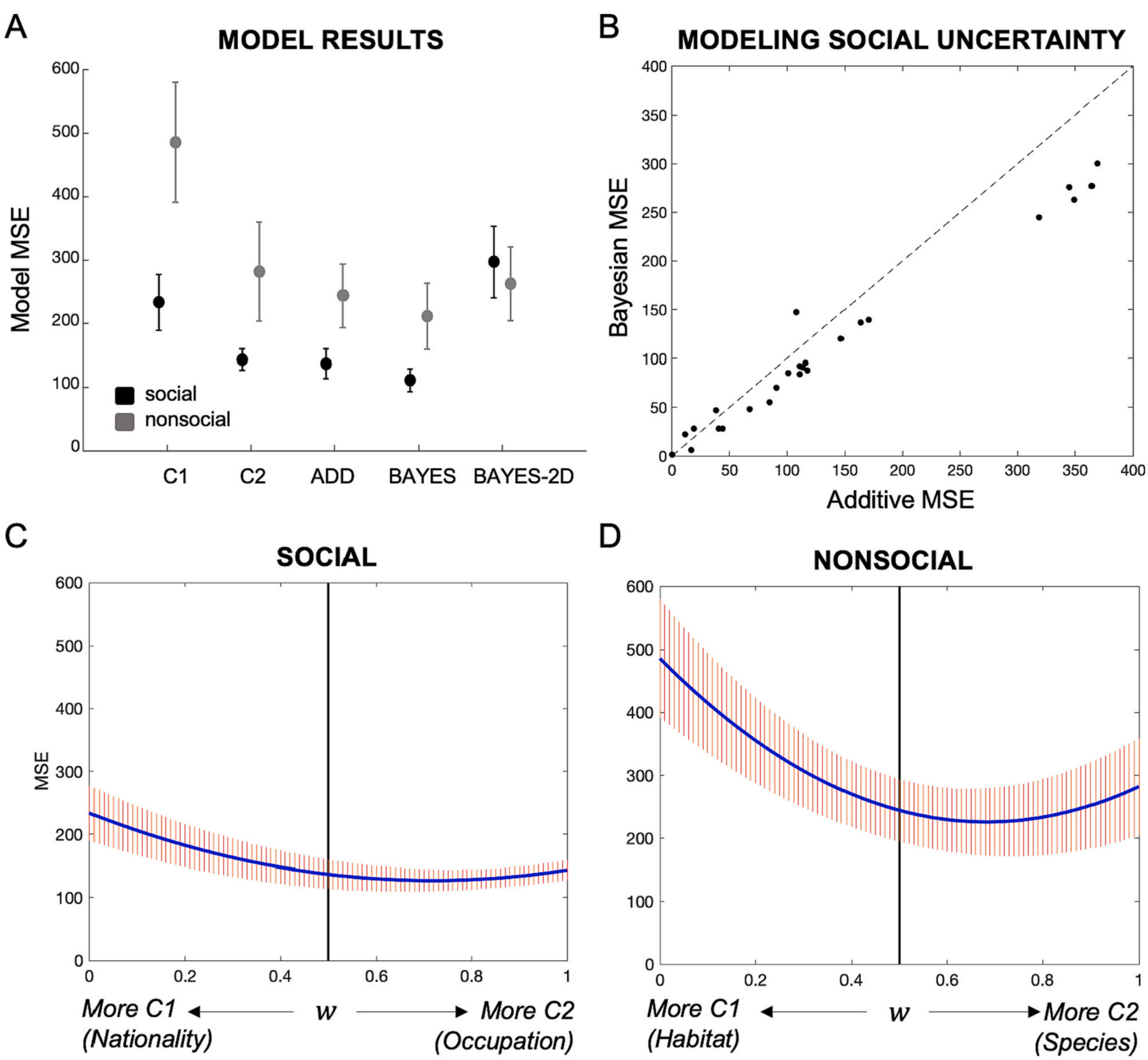

Fig. 3. Predicting combination attributes. (A) Mean squared error (MSE) for the C1, C2, Additive, Bayes, and Bayes-2D models' predictions for the social (black) and nonsocial (gray) combinations. (B) The incorporation of uncertainty in the Bayesian model reduces error relative to the Additive model across the 25 social combinations. (C) Weighted additive model for social combinations. (D) Weighted additive model for the nonsocial combinations.

combinations for each value of W (Fig. 3 CD). Note that when $W_{C1} = 0$ in the weighted Additive model it is identical to the Occupation (C2) model alone. We did not find an increase in Additive model performance at any values of W relative to the unweighted model in either the social or nonsocial domain.

3.1.2. Factors that influence whether attributes of social concepts are linearly combined

When do people use the attributes of simple concepts (e.g., *Mexican, lawyer*) to generate predictions of the attributes of combined social concepts (e.g., *Mexican lawyer*)? The performance of our predictive models reveals that people may rely on a linear combination of nationality and occupation concepts, or the occupation alone, when judging the attributes of a nationality-occupation combination. Under which circumstances are people more likely to rely on these strategies? The success of the Bayesian model at predicting combination attributes suggests that uncertainty may factor into the interpretation of combinations in multiple ways. For example, people may be more likely to combine representations of the simple concepts when the attributes of those concepts are more certain, and thus would act as more reliable inputs in a linear combination process. In other words, people may be less likely to linearly combine concepts when their attributes are more uncertain. Our data support this: for social concepts, we observed a positive relationship between constituent concept uncertainty U and Bayesian model errors for the combinations (r=0.43, p=.032; Fig. 4A). The relationship between concept uncertainty and Occupation (C2) model errors was positive but not statistically significant (r=0.39, p=.056).

Another possibility is that people are more likely to rely on a linear combination of conceptual attributes when the combination itself is

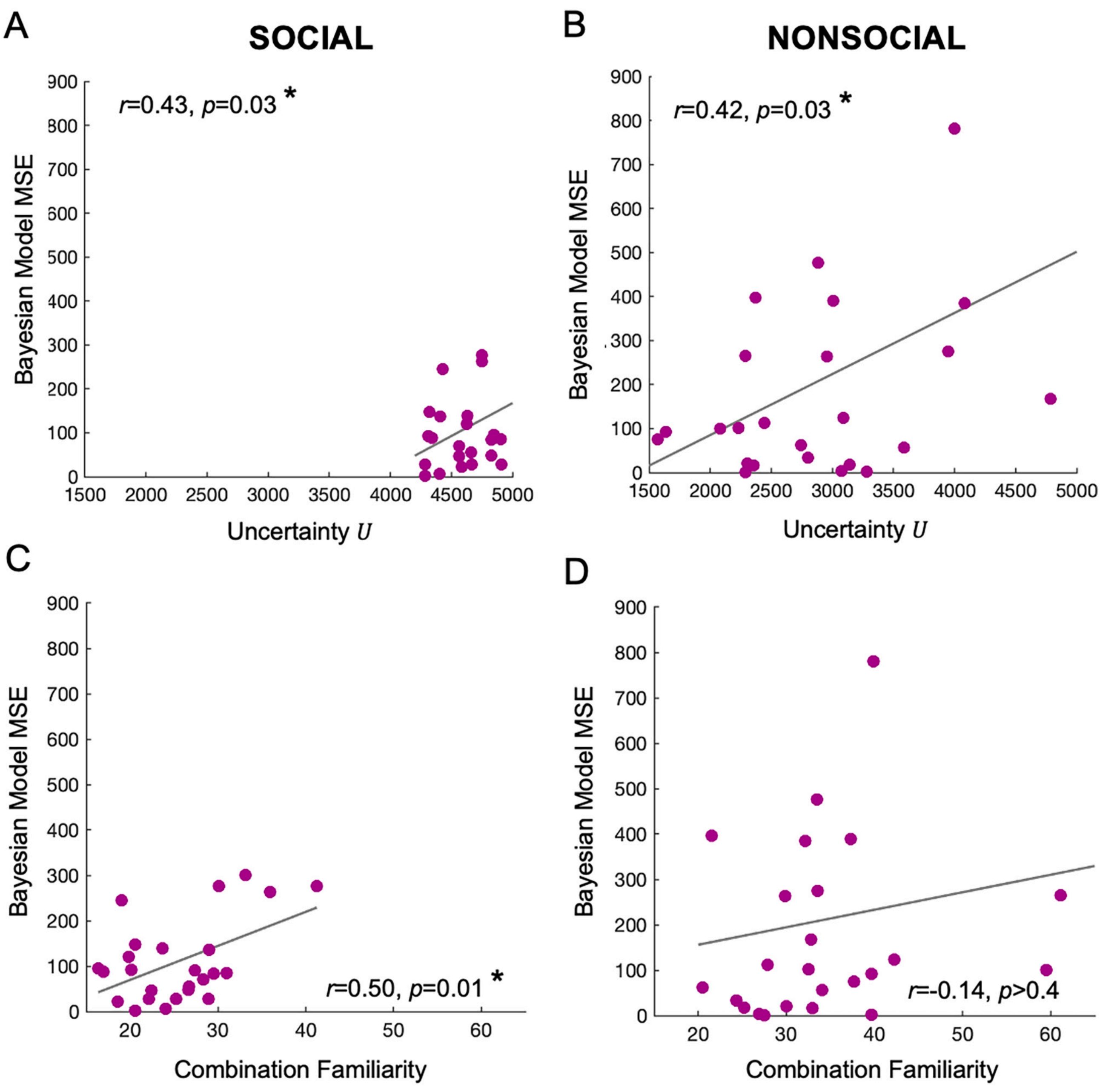

Fig. 4. Simple concept uncertainty and combination familiarity. (A) Increased uncertainty around the attributes of social concepts predicts worse performance of the Bayesian model. (B) This relationship was also significant for nonsocial combinations. (C) Increased familiarity of social combinations predicts worse performance of the Bayesian model. (D) This relationship was not observed for nonsocial combinations.

unfamiliar, thus requiring people to generate a judgment of the combination on-the-fly. Our findings also support this hypothesis: familiarity positively predicted Bayesian model errors across the 25 social combinations (r=0.50, p=.011; Fig. 4C), meaning that more linear combinations occurred when the combined social concept was less familiar. A similar relationship was observed between familiarity and Occupation (C2) model errors (r=0.47, p=.017), which is perhaps unsurprising given that the C2 model is a special case of a linear model in which C1 is given zero weight. Taken together, these results show that when a social combination itself is more familiar, people are more likely to rely on prior knowledge about the attributes of that combination than to rely on knowledge about the attributes of the relevant simple concepts that make up the combination.

In the nonsocial domain, the relationship between constituent uncertainty U and Bayesian model errors was also significantly positive (r=0.42, p=.035; Fig. 4B). In contrast, no relationship between combination familiarity and Bayesian model errors was observed for nonsocial concepts ($p > .4$; Fig. 4D).

Together, these patterns of results are consistent with the possibilities that (i) the state of one's knowledge of the constituent concepts may influence the conceptual combination process for both social and nonsocial concepts, with more certainty producing more linear combinations, but that (ii) the state of one's knowledge about the combinations themselves may play a particular role in social domain, whereby more familiarity is associated with less linear combinations, perhaps because conceptual combination is precluded by the application of pre-existing stereotypes about particular combinations of person attributes.

3.2. Non-linear combinatorial processing in the social domain

So far, it appears that people are less likely to linearly combine social concepts when the attributes of the constituent concepts are uncertain and also when the social combination is familiar. This does not, however, tell us what strategy people are using when they are not using a linear combination. In other words, what explains the failures of the predictive models? We used the Differentiation measure to test two possibilities. The first possibility is that when people are not combining concepts linearly, they rely on the central tendency within the relevant attribute space to make their judgments. The second possibility is that they might instead differentiate the combination from the central tendency by exaggerating its attributes away from the center of attribute space. We use the direction of the correlation between Bayesian model errors and Differentiation to distinguish between these two hypotheses, which predicts either a negative or a positive relationship, respectively. Our results support the second hypothesis: we observed a positive relationship between Bayesian model errors and Differentiation (r=0.92, $p < .001$). That is, when people are not combining concepts linearly, they judge the members of the combined concept group to have more extreme attributes than would be expected given the simple concepts alone.

Furthermore, we find that people agree upon the exaggerated attributes of social combinations: we observed a strong, positive relationship between Differentiation and Agreement (r=0.59, p=.002; Fig. 5A). We also observe a positive relationship between Agreement and Bayesian model errors (r=0.73, $p < .001$). One possible explanation for these findings is simply that people agree on the attributes of a combination because they are all similarly relying on the central tendency of nationality-occupation combinations in attribute space—not because people agree on unique, exaggerated locations in attribute space for specific combinations. If this were true, we would expect a positive relationship between Agreement and the proximity of the combination to the center of combination attribute space (i.e., the gray circle in Fig. 2A). However, this relationship was not found (r=0.03, $p > .9$). Furthermore, a linear regression model revealed that combination familiarity (p = .007), Agreement ($p < .001$), and Differentiation ($p < .001$) all explained unique variance in Bayesian model errors. It thus appears that familiarity with social combinations decreases the likelihood of using a linear combination process; in turn, when people are not combining social concepts linearly, attributes of the combination are exaggerated away from the center of attribute space in unique, combination-specific ways, and people agree on what these attributes are.

Why do people agree on the attributes of combinations with which they are more familiar? One possibility is that people are familiar with a combined social concept because representations of that group's members (whether true or false) are common in popular culture. If large numbers of individuals have access to these representations, then they may be both agreed upon within the population and also familiar to

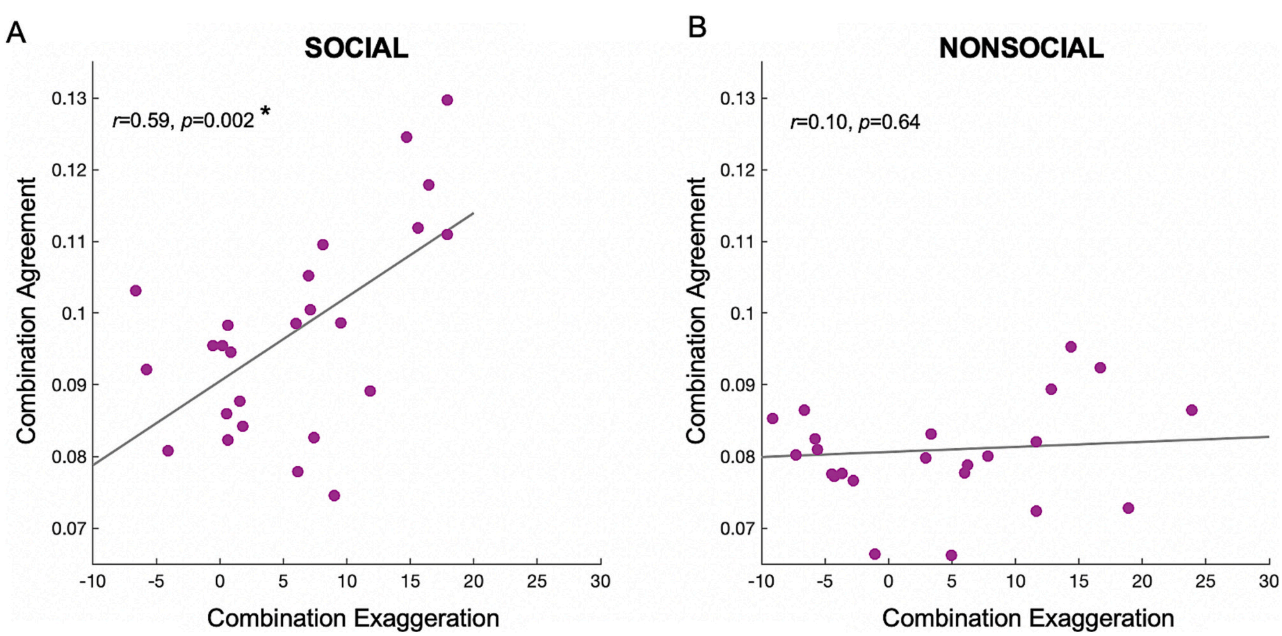

Fig. 5. Combination Exaggeration predicts Combination Agreement in the social domain. (A) When the attributes of the social combinations are exaggerated away from the center of attribute space, there is greater agreement across people on what those attributes are. (B) This relationship did not exist in the nonsocial domain.

individuals. Thus, the relationship between familiarity and Differentiation could be explained by the relationship between familiarity and Agreement. A linear regression model provided support for this hypothesis: when familiarity and Agreement were simultaneously used to predict Differentiation, Agreement predicted unique variance (p=.006) whereas familiarity did not (p=.11). Thus, the relationship between familiarity and Differentiation may be best understood as reflecting available representations of certain combined social groups in popular culture, such that shared exposure to these representations results in both personal familiarity with these groups and also population-level agreement regarding their (exaggerated) attributes.

In the nonsocial animal combinations, we also observed a positive relationship between Differentiation and Bayesian model errors (r=0.78 $p < .001$), suggesting that the failures of the Bayesian model can partially be explained by the exaggeration of a combination's attributes away from the center of semantic space, rather than a reliance on the central attribute tendencies. However, there was no relationship found between Agreement and Bayesian errors (r=0.04, $p > .8$). There was also no relationship between Differentiation and Agreement in the nonsocial domain (r=0.10, p>.6; Fig. 5B).

3.3. Comparisons between the social and nonsocial domains

Social combinations (M=25.5; SD=6.1) were overall rated as less familiar than nonsocial combinations (M=34.3; SD=9.7; t(48)=3.85, $p < .001$). Composite attribute uncertainty for the social combinations was higher (M=4622, SD=263) than the composite attribute uncertainty for the nonsocial combinations (M=2909, SD=797; t(48)=10.2, $p < .001$). At the same time, however, there was greater agreement across people on the attributes of the social combinations (M=0.10; SD=0.014) than of the nonsocial combinations (M=0.08, SD=0.007; t(48)=5.1, $p < .001$). There was no difference in combination Differentiation across social (M=5.8, SD=7.2) and nonsocial domains (M=2.8, SD=9.9; t(48)= 1.2, p=.23).

Regarding the performance of our predictive models, there is some evidence that the Bayesian model may perform better in the social domain (M=111.0, SD=91.7) than the nonsocial domain (M=211.5; SD=260.8), but this difference was not reliable (t(48)=1.8, p=.08). The C1 (Nationality) model in the social domain performed significantly better than the C1 (Habitat) model in the nonsocial domain (t(48)=2.4, p=.02). The improved performance of the C2 and Additive models in the social domain relative to the nonsocial domain was not reliable (p = .09, p = .06).

4. Discussion

Our ability to construct rich and complex social combinations from their constituent parts has been characterized in a variety of ways, including as a simple additive process (Asch, 1946) or as the result of causal reasoning (Kunda et al., 1990). However, one previously overlooked aspect is the explicit role of uncertainty in the construction and evaluation of combined concepts such as *Mexican lawyer*. We addressed this gap by examining how judgments of combined concepts are influenced both by uncertainty about the attributes of the simple concepts they comprise (e.g., the warmth and competence of *Mexican* and *lawyer* separately), and by prior knowledge about those combined concepts (e.g., familiarity with *Mexican lawyers*). We examined the performance of predictive models to compare different combinatorial strategies. For both social and nonsocial combinations, we found that people did not weight the attributes of two simple concepts equally. In the social domain, we found that a Bayesian model outperformed an Additive model, indicating that people are more likely to judge social combinations in a probabilistic manner rather than on a simple weighted average of the simple concepts' attributes. However, the Bayesian model performed worse when the attributes of the simple concepts were more uncertain and also when the combinations themselves were more familiar, revealing that linear combination is used more when individuals have relatively certain ideas about the constituent concepts of a combination and are relatively unfamiliar with the combination itself. Further, we found that people tend to judge social combinations to have more extreme attribute values than predicted by linear integration, and that people tend to agree on what these attributes are. We address each of these results in turn below, then note points of divergence from the nonsocial domain.

What might explain the differential weighting between the first concept and second concept in both social and nonsocial combinations? One possibility is offered by the competition among relations in nominals (CARIN) model of conceptual combination, which assumes that noun-noun compounds are interpreted by selecting a thematic relation between the constituent concepts (Gagné & Shoben, 1997). On this view, the first concept acts as a modifier for the second concept after an appropriate thematic relation is found. For example, *Mexican lawyer* would be interpreted as a lawyer who is from Mexico, and *cave cat* would be interpreted as a cat that lives in a cave. In essence, the modifier constrains the space of possible judgments of the head noun, thus making the head noun a more reliable input to the overall evaluation of the combination. A second possibility is that people simply believe (implicitly or explicitly) that certain types of information are more diagnostic than others and should accordingly carry more weight in the combination. For example, there may be differences in participants' relative willingness to use nationality information when other, potentially more diagnostic information (e.g., occupation) is present versus when it is not. Another related but distinct possibility is that the relative weights are influenced by the smaller range of social perception of the nationality concepts. With this in mind, it would be premature to conclude any special role for head versus modifier nouns per se in social conceptual combination on the basis of the current findings. More generally, the extent to which these combinations are best understood as reflecting a merging of two whole concepts or an integration of specific features of those concepts remains open to future research.

On its own, the observation that constituent concepts are weighted unequally in a combination does not tell us how a final interpretation of the combination is achieved. To gain mechanistic insight into this process, we compare the performance of predictive models that make different assumptions about how information from the constituent concepts is used. The goal of this modeling approach was not to find the best possible model for predicting social combinations, but rather to test different assumptions embedded in relatively simple models. Our Additive model relies on a weighted combination of the constituent concept attribute judgments, while the Bayesian model relies on the probability distribution of those same judgments, and therefore captures participants' degrees of uncertainty. The Bayesian model outperformed the Additive model in the social domain, indicating that attribute uncertainty factored into participants' judgments of the combined concepts. While this influence of uncertainty is consistent with our previous findings in the domain of object concepts (Solomon & Thompson-Schill, 2020), we did not find that the Bayesian model outperformed the Additive model for non-social animal combinations.

Additional analyses revealed that uncertainty can also influence the *strategy* used to interpret the attributes of social combinations. This influence of uncertainty manifested in two ways. First, we found that the Bayesian model performed worse under conditions of high constituent concept uncertainty. That is, a combinatorial model was best able to predict people's judgments when the simple concepts' attributes were more certain, but produced more errors when the single concepts' attributes were more uncertain. Based on these findings, we suggest that people are more likely to use a linear combination strategy when the attributes of the constituent concepts are known, but rely on another strategy when the attributes are more uncertain. Second, the Bayesian model performed worse when combinations in the social domain were more familiar, suggesting that uncertainty surrounding the combinations themselves influences the interpretive strategy used. Our combination familiarity measure captured the degree to which participants have interacted with or hold beliefs about people from our combined nationality and occupation groups. We therefore suggest that previous experience with a complex social group reduces the use of a linear combination process. Together, we argue that there are at least two sources of uncertainty that influence judgments of complex social concepts: people are more likely to linearly combine the attributes of two social groups to form a judgment of the complex group when the attributes of the individual concepts are evaluated with greater certainty and when the attributes of the combined concept are unfamiliar.

While uncertainty appears to influence the strategies used to combine social concepts—specifically, if simple concepts are linearly integrated—these relationships alone cannot reveal the nature of any alternative strategies to linear integration. We therefore took a separate approach to explore what other strategies people might be using. One possibility is that instead of integrating the simple concepts' attributes, people predict the attributes of a complex group by reverting to the prototypical attributes of the population. Another possibility is that people might predict the attributes of the complex group to be substantially different from the prototypical attributes, thereby differentiating the complex group from the population as a whole. We considered the center of 2D warmth-competence space to reflect prototypical attribute values (i.e., the average warmth and competence estimates across all simple social groups; Fig. 1A). This enabled us to construct a Combination Differentiation measure that reflected whether each complex group was more or less similar to the prototype than would be predicted by linear integration of its constituents. We found that the Bayesian model failed when combinations were highly differentiated from the prototype, indicating that when people do not linearly combine concepts, they are likely to differentiate the social group from the prototype, rather than revert to it. In other words, people may predict that members of complex social groups have more extreme attributes than would be expected given their constituent concepts alone.

So far, our results show that people are less likely to linearly combine social concepts when the simple concepts are unknown and the combination is more familiar, and that when social concepts are not linearly combined, the attributes of the combination are exaggerated. These observations are consistent with the possibility that people arrive at nonlinear judgments of combinations by relying on their own personal experience with the relevant social groups. However, another possibility is that people (also) rely on societally-shared stereotypes about those social groups. Our Combination Agreement measure—reflecting the inverse variance associated with each combination across all participants—was positively related to Combination Differentiation, indicating that people tend to agree on the exaggerated attributes of social combinations and that the consistency of combination judgments might emerge from social perceptions that are shared across the population. The idea that shared knowledge across the population is leveraged to interpret complex social concepts aligns well with the possibility that stereotypes are used as heuristics for making judgments in the absence of other information. Further, while people tended to agree on the attributes of highly familiar combinations, Agreement predicted degree of Differentiation above and beyond familiarity, strengthening the claim that people rely on shared stereotypes to judge the attributes of complex social groups. This suggests that at least one possible source of familiarity could be impressions of highly stereotyped groups in popular culture. It is also possible that people's own experiences with individuals from a given social group could lead them to have "exaggerated" perceptions of that group, though we cannot formally examine this here. Future work could assess the degree to which specific social combinations are associated with stereotypes in part by indexing individuation (i.e., isolating personal experience).

There were several points of divergence between our results in the social compared to nonsocial domain. First, while the Bayesian model reliably outperformed the Additive model for social combinations, no such difference was found for the nonsocial combinations. We also did not observe a relationship between linear integration and combination familiarity in the nonsocial domain, nor did we observe a relationship between Differentiation and Agreement for the nonsocial combinations. Therefore, while people rely on common knowledge or beliefs about social groups when judging the attributes of social combinations (e.g., *Mexican lawyer*), people do not appear to rely on shared knowledge or beliefs when judging the attributes of nonsocial animal combinations (e.g., *circus pig*). One possible source of distinction between our findings in the social and nonsocial domains may be self-other similarity. For example, the degree to which warmth and competence are attributed to people from a certain social group is influenced by a variety of social cues (e.g., nationality, occupation, gender, etc.), all of which may interact with the identity of the perceiver, which was not considered in the present study. Given that in-group/out-group status can affect mentalizing and empathy (Cikara, Bruneau, & Saxe, 2011; Jenkins, Macrae & Mitchell, 2008), attribute ratings of different social groups in our study may be partly driven by participants' judgments of their own warmth and competence, as well as their degree of affiliation with a given nationality and occupation. Our nonsocial concepts, on the other hand, likely do not elicit the same kind of self-other comparison.

An advantage of dimensional frameworks of person perception is that they are general. In principle, any collection of attributes of a person can be translated into their effects on core perceptions of that person's warmth and competence. To the extent that we take these frameworks seriously, combinations of other types of attributes of people besides occupation and nationality should be expected to be characterized by similar effects of uncertainty as in the present investigation (Jenkins et al., 2018; Kobayashi et al., 2022). At the same time, the current findings suggest that an important difference between the social and nonsocial concepts may be that some of the social combinations, but not the nonsocial combinations, are associated with culturally shared assumptions that deviate from the intersection of the representations of the simple concepts, thus precluding novel conceptual combination. Accordingly, a promising route for future research will be to test if nonsocial combinations that have become common in popular culture behave like highly familiar social combinations.

5. Conclusion

Membership in any given social group (nationality, occupation, gender, etc.) is not an isolated feature of our perceptions of others, but is rather constructed alongside other information, including membership in other groups. The idea that people build on different sources of information to approximate features of others, such as their warmth and competence, may not be controversial; however, we demonstrate here for the first time that this approximation process is influenced by uncertainty in two main ways: uncertainty about the attributes of individual concepts affects *how* the attributes of those concepts are combined, and unfamiliarity with the combination itself affects *when* the attributes of the individual concepts are combined. We suggest that uncertainty is a critical aspect of the mechanisms underlying conceptual combination, and that it influences our interpretations of complex concepts through balancing bottom-up generative processes with top-down effects of prior knowledge.

CRediT authorship contribution statement

Alice Xia: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Data curation, Formal analysis. **Sarah H. Solomon:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Data curation, Formal analysis. **Sharon L. Thompson-Schill:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Adrianna C. Jenkins:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

Link to data/code provided in revised manuscript

Acknowledgements

We thank members of the Thompson-Schill Lab and the Jenkins Lab for feedback on the conceptualization of this work. This work was supported by funding from the National Institutes of Health [R01 DC015359] awarded to S.T.S and institutional funding to A.C.J.

References

Allport, G. W. (1979). *The nature of prejudice*. Addison-Wesley.
Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology, 41*(3), 258.
Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *NeuroImage, 55*(4), 1847–1852.
Baroni, M., & Zamparelli, R. (2010, October). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing* (pp. 1183–1193).
Boylan, C., Trueswell, J. C., & Thompson-Schill, S. L. (2015). Compositionality and the angular gyrus: A multi-voxel similarity analysis of the semantic composition of nouns and verbs. *Neuropsychologia, 78*, 130–141.
Boylan, C., Trueswell, J. C., & Thompson-Schill, S. L. (2017). Relational vs. attributive interpretation of nominal compounds differentially engages angular gyrus and anterior temporal lobe. *Brain and Language, 169*, 8–21.
Cantor, N., & Mischel, W. (1979). Prototypes in person perception. In *, Vol. 12. Advances in experimental social psychology* (pp. 3–52). Academic Press.
Chang, K. M. K., Cherkassky, V. L., Mitchell, T., & Just, M. A. (2009, August). Quantitative modeling of the neural representation of adjective-noun phrases to account for fMRI activation. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP* (pp. 638–646).
Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and them: Intergroup failures of empathy. *Current Directions in Psychological Science, 20*(3), 149–153.
Connolly, A. C., Fodor, J. A., Gleitman, L. R., & Gleitman, H. (2007). Why stereotypes don't even make good defaults. *Cognition, 103*(1), 1–22.
Connolly, A. C., Fodor, J. A., Gleitman, L. R., & Gleitman, H. (2007). Why stereotypes don't even make good defaults. *Cognition, 103*(1), 1–22.
Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S., Leyens, J. P., & Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology, 48*(1), 1–33.
Dahlgren, K. (1985). The cognitive structure of social categories. *Cognitive Science, 9*(3), 379–398.
Estes, Z. (2003). A tale of two similarities: Comparison and integration in conceptual combination. *Cognitive Science, 27*(6), 911–921.
Estes, Z., & Jones, L. L. (2008). Relational processing in conceptual combination and analogy. *Behavioral and Brain Sciences, 31*(4), 385–386.
FeldmanHall, O., & Shenhav, A. (2019). Resolving uncertainty in a social world. *Nature Human Behaviour, 3*(5), 426–435.
Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77–83.
Fodor, J., & Lepore, E. (1996). The red herring and the pet fish: Why concepts still can't be prototypes. *Cognition, 58*(2), 253–270.
Fodor, J. A., & Lepore, E. (1998). The emptiness of the lexicon: Reflections on James Pustejovsky's the generative lexicon. *Linguistic Inquiry, 29*(2), 269–288.
Gagné, C. L. (2001). Relation and lexical priming during the interpretation of noun–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 236.
Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(1), 71.
Gelman, R., & Spelke, E. (1981). The development of thoughts about animate and inanimate objects: Implications for research on social cognition. In J. H. Flavell, & L. Ross (Eds.), *Social cognitive development: Frontiers and possible futures* (pp. 43–65). Cambridge, MA: Cambridge University Press.
Hampton, J. A. (1997). Conceptual combination. In K. Lamberts, & D. R. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 133–159). Cambridge, MA: The MIT Press.
Hampton, J. A. (1998). Similarity-based categorization and fuzziness of natural categories. *Cognition, 65*(2–3), 137–165.
Henley, N. M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8*(2), 176–184.
Ho, A. K., Sidanius, J., Levin, D. T., & Banaji, M. R. (2011). Evidence for hypodescent and racial hierarchy in the categorization and perception of biracial individuals. *Journal of Personality and Social Psychology, 100*(3), 492.
Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15*(4), 332–340.
Hutter, R. R., & Crisp, R. J. (2008). Reduced constituent category application in surprising combinations. *The Journal of Social Psychology, 148*(2), 247–252.
Jenkins, A. C., Karashchuk, P., Zhu, L., & Hsu, M. (2018). Predicting human behavior toward members of different social groups. *Proceedings of the National Academy of Sciences, 115*(39), 9696–9701.
Jenkins, A. C., Macrae, C. N., & Mitchell, J. P. (2008). Repetition suppression of ventromedial prefrontal activity during judgments of self and others. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 4507–4512. 42.
Kang, S. K., & Chasteen, A. L. (2009). The moderating role of age-group identification and perceived threat on stereotype threat among older adults. *The International Journal of Aging and Human Development, 69*(3), 201–220.
Kenett, Y., & Thompson-Schill, S. L. (2020). *Novel conceptual combination can dynamically reconfigure semantic memory networks. PsyArXiv*.
Kobayashi, K., Kable, J. W., Hsu, M., & Jenkins, A. C. (2022). Neural representations of others' traits predict social decisions. *Proceedings of the National Academy of Sciences, 119*(22), Article e2116944119.
Kunda, Z., Miller, D. T., & Claire, T. (1990). Combining social concepts: The role of causal reasoning. *Cognitive Science, 14*(4), 551–577.
Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*(2), 203–208.
Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology, 69*(3), 397.
Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology, 20*(2), 158–190.
Mitchell, J., & Lapata, M. (2008, June). Vector-based models of semantic composition. In *Proceedings of ACL-08: HLT* (pp. 236–244).
Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science, 34*(8), 1388–1429.
Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *Neuroimage, 28*(4), 757–762.
Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12*(4), 529–562.
Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 510.
Quinn, K. A., Mason, M. F., & Macrae, C. N. (2009). Familiarity and person construal: Individuating knowledge moderates the automaticity of category activation. *European Journal of Social Psychology, 39*(5), 852–861.
Remedios, J. D., Chasteen, A. L., Rule, N. O., & Plaks, J. E. (2011). Impressions at the intersection of ambiguous and obvious social categories: Does gay + black = likable? *Journal of Experimental Social Psychology, 47*(6), 1312–1315.
Rokeach, M., & Rothman, G. (1965). The principle of belief congruence and the congruity principle as models of cognitive interaction. *Psychological Review, 72*(2), 128.
Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382–439.
Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science, 12*(4), 485–527.
Smith, E. E., Patalano, A. L., & Jonides, J. (1998). Alternative strategies of categorization. *Cognition, 65*(2–3), 167–196.
Smith, E. R. (1989). Procedural efficiency: General and specific components and effects on social judgment. *Journal of Experimental Social Psychology, 25*(6), 500–523.
Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*(1), 3.
Solomon, S. H., & Thompson-Schill, S. L. (2020). Feature uncertainty predicts behavioral and neural responses to combined concepts. *Journal of Neuroscience, 40*(25), 4900–4912.
Wattenmaker, W. D. (1995). Knowledge structures and linear separability: Integrating information in object and social categorization. *Cognitive Psychology, 28*(3), 274–328.
Wisniewski, E. J. (1996). Construal and similarity in conceptual combination. *Journal of Memory and Language, 35*(3), 434–453.