A Dynamic Interactive Theory of Person Construal

Jonathan B. Freeman and Nalini Ambady

Tufts University

A dynamic interactive theory of person construal is proposed. It assumes that the perception of other people is accomplished by a dynamical system involving continuous interaction between social categories, stereotypes, high-level cognitive states, and the low-level processing of facial, vocal, and bodily cues. This system permits lower-level sensory perception and higher-order social cognition to dynamically coordinate across multiple interactive levels of processing to give rise to stable person construals. A recurrent connectionist model of this system is described, which accounts for major findings on a) partial parallel activation and dynamic competition in categorization and stereotyping; b) top-down influences of high-level cognitive states and stereotype activations on categorization; c) bottom-up category interactions due to shared perceptual features; and d) contextual and cross-modal effects on categorization. The system's probabilistic and continuously evolving activation states permit multiple construals to be flexibly active in parallel. These activation states are also able to be tightly yoked to ongoing changes in external perceptual cues and to ongoing changes in high-level cognitive states. The implications of a rapidly adaptive, dynamic, and interactive person construal system are discussed.

Keywords: person perception, social cognition, dynamical systems, connectionist, face perception

If humans treated every new stimulus as a unique experience, we would quickly drown in a bewildering amount of redundant information. To improve the situation, the cognitive system groups stimuli that share similar characteristics into meaningful categories (Murphy, 2002; Rosch, 1978). In general, such categorization has numerous benefits. The brown, four-legged, wooden flat surface appearing before your eyes is instantly rendered "table," something to put things on, and not "chair," something to sit on. This saves time in making sense of the object and leads to adaptive behavior (e.g., that you do not sit on a table). This is but one example of how, in sorting through an impossibly complex world, categorization provides an efficient cognitive strategy to make the perceiver's job far easier.

The benefit of streamlining mental resources by categorizing other people rather than furniture, however, is not so straightforward. Indeed, social psychologists recognized the tremendous implications of person categorization early on. Seminal writers, such as Allport (1954), Sherif (1967), and Tajfel (1969), converged on the argument that categorizing other people was an inevitable economizing strategy used to simplify the cognitive demands of dealing with others. Their work had wide-sweeping influences on person perception research, and for nearly half a century their arguments set the stage for work on social categorization. Mere exposure to another person, it was thought, automatically triggered a relevant social category (e.g., sex, race, age), and along with that category, its corresponding knowledge structure. Activating category knowledge, it was shown, spontaneously triggered a variety of cognitive, affective, and behavioral outcomes. Countless studies documented how a White person's exposure to a Black man, for instance, unleashes a specific cascade of events. Encountering a Black man automatically activates the category, Black, which molds subsequent judgments and impressions (e.g., "he's aggressive"), triggers evaluations (e.g., "I don't like him"), and elicits patterns of behavior

(e.g., increases in aggression; Bargh, 1994; Bargh, 1999; Brewer, 1988; Devine, 1989; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Fiske & Neuberg, 1990; Gilbert & Hixon, 1991; Sinclair & Kunda, 1999). It became clear that social categorization influenced stereotyping and prejudice and had a powerful role in shaping interpersonal interaction. Given the implications, social psychological research placed a great deal of focus on the downstream dynamics of categorization, on the ways that categorical thinking shapes interpersonal outcomes.

Until quite recently, person perception research by and large investigated how perceivers make judgments and evaluations from written behavioral descriptions (but see McArthur & Baron, 1983). Real-world social targets, however, are not generally encountered through behavioral descriptions. Rather, in real life perceivers encounter other people first through sensory cues of the face, voice, and body. The theoretical and empirical work examining the links between lower-level perceptual processing and higher-order social cognition began only recently (see Bodenhausen & Macrae, 2006; Zebrowitz, 2006). Although it was long understood that perceivers frequently categorize other people along a variety of dimensions (e.g., sex, race, age) from mere exposure to their face (Brewer, 1988; Fiske & Neuberg, 1990; Strangor, Lynch, Duan, & Glas, 1992), the mechanisms and perceptual determinants underlying these categorizations received considerably less attention.

While social psychologists were documenting the downstream implications of perceiving others, cognitive psychologists and neuroscientists were examining person perception from a different perspective. They were concentrating their efforts on investigating the perceptual mechanisms of face processing (Bruce & Young, 1986; Burton, Bruce, & Johnston, 1990; Calder & Young, 2005; Farah, Wilson, Drain, & Tanaka, 1998; Haxby, Hoffman, & Gobbini, 2000). Recently, by integrating the social cognitive framework of person perception with insights from the cognitive literature on face processing, a growing body of research has begun to link lower-level perceptual processing with higher-order social cognition. This emerging body of work has come to be referred to as "person construal" research. Traditional social cognition research focused on the relatively high-level cognitive processes involved in person categorization and individuation, especially how these shape downstream phenomena (e.g., stereotyping, behavior). Person construal research, on the other hand, seeks to understand the lower-level perceptual mechanisms that produce these social cognitive phenomena in the first place.

Purpose of the Article

Social cognition researchers have developed a number of models of person perception, including models that explain how we reason about other people and infer their personality traits, how we categorize and individuate, and how explicit knowledge and memory of other people is learned, stored, and accessed (Bodenhausen & Macrae, 1998; Brewer, 1988; Chaiken & Trope, 1999; Fiske, Cuddy, Glick, & Xu, 2002; Fiske & Neuberg, 1990; Higgins, 1996; Kunda & Thagard, 1996; Read & Miller, 1998b; Smith & DeCoster, 1998; Srull & Wyer, 1989; van Overwalle & Labiouse, 2004). These models tend to place categorization as a starting point, after which subsequent interpersonal phenomena are richly explained (e.g., impressions, memory, behavior). Thus, the focus of these models is not to explain the categorization process; it is to explain the higher-order social cognitive processing that comes after.

Person construal research seeks to examine the lower-level perceptual mechanisms and determinants of categorization, including how categories and stereotypes are activated from cues of the face, voice, and body. To our knowledge, there has yet to be a comprehensive framework that details how such lower-level perceptual processing contributes to higher-order social cognitive phenomena. Here, we introduce such a framework, which utilizes increasingly popular approaches to cognition, namely connectionism and dynamical systems theory (Kelso, 1995; Port & van Gelder, 1995; Rogers & McClelland, 2004; Rumelhart, Hinton, & McClelland, 1986; Smolensky, 1989; Spivey, 2007). Recently, researchers have applied connectionist models to understand social cognitive phenomena as well (e.g., Kunda & Thagard, 1996; Read & Miller, 1993, 1998a; Read, Vanman, & Miller, 1997; Smith & DeCoster, 1998, 1999; van Overwalle, 2007; van Overwalle & Labiouse, 2004; Zebrowitz, Fellous, Mignault, & Andreoletti, 2003). In the present article, we apply connectionism and dynamical systems theory to comprehensively explain the process of person construal. Thus, we aim to provide a framework that explains social categorization processes at a perceptual level and links these processes to the higher-order social cognitive phenomena emphasized in prior models of person perception.

In the past decade, person construal research has documented a number of fascinating effects that have yet to be comprehensively accounted for by theoretical models. These range from findings of partial parallel activation, dynamic competition and continuous temporal dynamics, contextual and cross-modal effects on categorization, bottom-up category interactions due to shared perceptual features, and top-down effects on categorization, such as influences of motivational state, stereotype activation, and prejudice, among others (e.g., Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Eberhardt, Dasgupta, & Banaszynski, 2003; Freeman & Ambady, 2009, in press; Freeman, Ambady, Rule, & Johnson, 2008; Freeman, Pauker, Apfelbaum, & Ambady, 2010b; Hugenberg & Bodenhausen, 2003, 2004; Pauker et al., 2009). Together, these emerging findings suggest that person construal is a dynamic and highly interactive process. Here, we will offer

a framework for person construal that can explain these recent advances.

First, we will describe our dynamic interactive theory of person construal. Then, we will introduce a computational model that captures our theoretical claims. We will then explain how a number of recent findings are consistent with the model, and we will conduct several simulations to demonstrate this. Lastly, we will discuss how the theory and model compare with extant accounts, and we will discuss several important implications for present understandings of person construal.

A Dynamic Interactive Theory of Person Construal

We view the task of perceiving others as a dynamic interactive process, and we expound on this view below.

Top-Down and Bottom-Up Interactivity

In perceiving the world, we are continually extracting sensory information to guide our attempts in discerning what it is that lies before us. Even with the most mundane kinds of construal, such as perceiving objects or environments, we bring a great deal of knowledge to the perceptual process. This is only truer in the case of perceiving other people. Our rich set of prior experiences with another person or the regularities in our experience with whole groups of people (e.g., sex, race, age) undoubtedly provide a lens through which we construe others. Beyond the prior knowledge that might contextualize perception, our everyday encounters with others are also replete with complex affective and motivational states. Though there is much prior knowledge about the objects or environments we might encounter, this only pales in comparison to what is brought to the table when perceiving other people. We may have stereotypic beliefs about people of a certain sex, we may feel disdain for someone who has made us cry, or we may be motivated to make a good impression in order to land the job. In short, there is an enormity of prior knowledge and high-level states that may be brought to bear on the perception of our social world. Although traditionally it was long assumed that perception is primarily a bottom-up phenomenon and insulated from any top-down influence of higher-order processes (e.g., Fodor, 1983; Marr, 1982), it is becoming increasingly clear that perception arises instead from both bottom-up and top-down influences, likely mediated by large-scale neural oscillations (e.g., Engel, Fries, & Singer, 2001; Gilbert & Sigman, 2007). Even the earliest of responses in primary visual cortex, for example, are altered by top-down factors (Li, Piëch, & Gilbert, 2004). We argue, therefore, that our prior knowledge and expectations about people, our stereotypes, and our affective and motivational states may all dynamically interact with incoming sensory information in the perceptual process to shape person construal.

The person construal process invites another form of interactivity as well, one that is driven directly by the incoming sensory information itself. Whereas the perception of an object, for example, generally affords only one focal type of construal (e.g., "that's a table"), multiple construals are simultaneously availed to person perceivers, including sex, race, age, emotion, or inferences of personality characteristics, to name a few. Given how many construals are available, sometimes the perceptual cues supporting certain construals will, by chance, overlap. For instance, the cues specifying another person's sex and emotional state can overlap (Becker et al., 2007). An adult's facial features might by chance happen to overlap with the facial features more common in babies or with the facial features of another person we know, in turn shaping our inferences of his or her personality characteristics (Zebrowitz & Montepare, 2008). Thus, certain person construals may be thrown into interaction with one another because they are directly confounded in the bottom-up sensory information itself.

Time Dependence and Continuous Temporal Dynamics

We argue that the process of person construal is dynamic, in the sense that it takes time and fluctuates over time, and that representations triggered during this process are inherently time-dependent. For instance, recent evidence shows that, after catching sight of another person, representations of social categories and stereotypes dynamically evolve across hundreds of milliseconds until stabilizing over time (Freeman & Ambady, 2009; Freeman et al., 2008; Freeman et al., 2010b; also see Kunda, Davies, Adams, & Spencer, 2002). Thus, at each moment during the categorization process, representations are varying as a function of time, making time-dependent transitions between, for instance, ~0% activation and ~100% activation (Dale, Kehoe, & Spivey, 2007; Freeman et al., 2008). This is not particularly surprising when considering how a social categorization would be implemented in an actual human brain.

For instance, there is now a great deal of evidence suggesting that mental representations, as realized in the brain, are neuronal populations that convey information (e.g., "he's a man!") through patterns of activity distributed across many neurons (Rogers & McClelland, 2004; Spivey, 2007; Spivey & Dale, 2004). This was confirmed with regard to representations of the face by studies that recorded populations of temporal cortex neurons in nonhuman primates (Rolls & Tovee, 1995; Sugase, Yamane, Ueno, & Kawano, 1999). Thus, most modern-day accounts assume that mental representations, such as a representation of a social category, involve continuous changes in a pattern of neuronal activity (e.g., Rogers & McClelland, 2004; Smith & Ratcliff, 2004; Spivey, 2007; Spivey & Dale, 2004; Usher & McClelland, 2003). For instance, about 50% of a face's identity is transiently represented in macaque temporal cortex as early as only 80 ms after a face's presentation, but the remaining 50% of its representation gradually accumulates over the following hundreds of milliseconds

(Rolls & Tovee, 1995). Thus, in early moments of processing representations of a face's category memberships would reflect a rough "gist," because the initial rough sketch of the face is partially consistent with multiple interpretations (e.g., both male and female). As the ongoing accrual of more and more information continues, however, the pattern of neuronal activity gradually sharpens into an increasingly clear interpretation (e.g., male) while other competing, partially-active representations (e.g., female) are pushed out (Freeman, Ambady, Midgley, & Holcomb, in press; Freeman et al., 2008; Smith & Ratcliff, 2004; Spivey & Dale, 2004; Usher & McClelland, 2003).

Indeed, by tracking the categorization process as it unfolds in real-time (through measuring the trajectory of hand movements en route to category responses on a screen), such a dynamic competition between multiple partially-active representations has been observed (Dale et al., 2007; Freeman & Ambady, 2009; Freeman et al., 2008; Freeman et al., 2010b). These findings suggest that a single category representation (e.g., male) does not discretely activate at an instantaneous moment after a target's presentation, nor does a single category representation transition from zero activation to full activation across time. Instead, such findings suggest that person construal involves alternative, competing categories that are simultaneously and partially active, and these evolve over time until stabilizing onto ultimate construals. Given such continuous dynamics, we argue that person construal is a temporally dynamic process and that person construal phenomena (e.g., a social categorization; activation of a stereotype) are best understood as gradual time-dependent transitions between mental states (e.g., from state A, the initial sight of another person, transitioning to state B, the ~100% confident recognition that the person is a White man). Further, we argue that during this time-dependent process, representations of a person's category memberships (e.g., male, White) as well as other candidate category memberships (e.g., female, Black, Asian) are rapidly fluctuating over time until achieving a stable, steady state.

Complex Integration

Person construal routinely involves complex integration. Even the simplest of construals, such as categorizing a person's sex, requires simultaneous integration of an enormous amount of information. For instance, all the various cues of the internal face in addition to peripheral cues such as hair must be integrated into a coherent interpretation of a target's sex. In many person construal tasks of the laboratory, this may be the only information available to perceivers—and even these simple tasks require already a substantial integration among cues. In everyday person construal and more complicated laboratory tasks, however, the integration is even more complex. For instance, perceivers receive information from multiple sensory modalities at the same time. Thus, to perceive the sex of real-world social targets, sex-specifying cues of the face and body arriving in the visual system must be integrated together with the vocal cues arriving in the auditory system. Moreover, not only does bottom-up sensory information need to be integrated, so too do top-down informational sources, as described earlier. For instance, high-level motivational states influence the perception of a face's race (Pauker et al., 2009). Moreover, priming context, expectations, stereotypes, cultural knowledge, among many other top-down factors, shape basic perceptions (e.g., Balcetis & Dunning, 2006; Eberhardt et al., 2003; Hugenberg & Bodenhausen, 2004; Johnson, Pollick, & McKay, 2010; MacLin & Malpass, 2001; Pauker, Rule, & Ambady, 2010). Thus, there is a complexity of information involving many sources—some bottom-up, some top-down—that must be integrated together in a very short amount of time to perceive others.

Theoretical Claims

In consideration of the above, we propose that perceptions of other people are accomplished by a dynamical system in which they gradually emerge through ongoing cycles of interaction between categories, stereotypes, high-level cognitive states, and the low-level processing of facial, vocal, and bodily cues. As such, this system permits lower-level sensory perception and higher-order social cognition to continuously coordinate across multiple interactive levels of processing to give rise to stable person construals. We capture this dynamic interactive theory of person construal with a computational model, which is introduced below.

The Integrative Power of a Recurrent Connectionist Network

Dynamical systems, such as a recurrent connectionist network or the human brain, are powerful in their ability to integrate multiple simultaneous sources of information. In a recurrent connectionist network, there are a number of nodes with connections that can be positive (excitatory) or negative (inhibitory). These nodes are not intended to represent individual neurons, but the overall structure of a network is often intended to be approximately neurally plausible. Arguably, nodes may operate like large populations of neurons and the connections between them operate like synapses (Smolensky, 1989). The critical feature of a recurrent connectionist network, which distinguishes it from other (feedforward) connectionist networks, is that many of these connections are bidirectional. Thus, as one node's activation tends to excite the nodes connected to it, the excitation of the nodes connected to it send feedback to the original node. Thus, many nodes in a recurrent network experience feedback, where they both influence and are influenced by the other nodes connected to them. Indeed, it is now clear that many neuronal projections in the human brain are bidirectional, producing recurrent feedback loops across both local and large-scale neural networks (e.g., Brefczynski & DeYoe,

1999; Dragoi, Sharma, & Sur, 2000; Lamme & Roelfsema, 2000, also see Spivey, 2007). Thus, recurrent connectionist networks have relatively high neural plausibility (Smolensky, 1989). It is this feedback among nodes that leads to the powerfully integrative nature of a recurrent network. Initially, a network is stimulated by external input. This input could come from bottom-up sources (e.g., facial or vocal cues) as well as top-down ones (e.g., motivation, task demands, prejudice). Activation then spreads among all nodes simultaneously (as a function of their connection weights). Because many of the nodes receive feedback, complex feedback loops are produced within the system. This causes the system to dynamically converge on an overall stable pattern of activation that best fits the input. This convergence involves the network's flows of activation gradually settling into a stable, steady state, where the activation of each node reaches an asymptote (Smolensky, 1989).

Dynamic constraint satisfaction. Because a node's activation is a function of all the positive and negative connections to other nodes that are activated in parallel, the final activation of a node (i.e., when the system stabilizes on a steady state) can be thought of as the satisfaction of multiple constraints. Each connection between nodes is a constraint. For instance, a node representing the category MALE might excite and be excited by another node representing the stereotype AGGRESSIVE. When these two nodes are incorporated in a larger recurrent network that is stimulated by, for instance, a male face, this MALE–AGGRESSIVE between-node connection serves as a constraint on the network. That is, for the network to ever achieve a stable state, activation must flow through that connection and incorporate it into an overall stable pattern (in addition to all other connections). Thus, the steady states that a recurrent network eventually stabilizes on are end-solutions that maximally satisfy all the constraints in the network, including between-node connections (e.g., MALE–AGGRESSIVE) and the input (e.g., facial cues, vocal cues, task demands, prejudice). As such, nodes in a recurrent network constrain each other in finding a best overall pattern that fits the input. In person construal, therefore, the stable states that a recurrent network achieves could be thought of as the satisfaction (i.e., integration) of many pieces of potentially conflicting information, including bottom-up sources (e.g., facial cues), in addition to top-down ones (e.g., motivational factors, task demands). This property of recurrent connectionist networks—dynamic constraint satisfaction—makes these networks powerfully integrative, much like the person construal process itself. A thorough explanation of how a dynamical system with feedback leads to the emergent ability to stabilize on steady states that maximally satisfy system constraints is beyond the scope of this article, but extensive discussions may be found elsewhere (e.g., Hopfield, 1982; Rumelhart et al., 1986; Smolensky, 1989).

Attractor dynamics. The stable states that a recurrent network settles into may be described as *attractors* (see Churchland & Sejnowski, 1989), in the sense that the network is attracted to be in that activation pattern (because it maximally satisfies all the constraints). Given different initial conditions (e.g., different faces, different high-level cognitive states), a network has many attractive overall patterns of activation (attractors) that it will gravitate towards. The mathematical properties of the phenomenon of attraction, which is inherent to nonlinear dynamical systems across nature (including cognitive and neural systems), has been studied extensively (with respect to cognitive and neural systems, see Churchland & Sejnowski, 1989; Izhikevich, 2006; Kelso, 1995; Rumelhart et al., 1986; Spivey, 2007).

If we modeled sex categorization in a recurrent connectionist network, for instance, eventual judgments (e.g., "that's a man!") would simply correspond with the person construal system gradually stabilizing on a state of activation that best fits the input (e.g., a male face). Thus, we can conceive person construal as the process by which the person construal system settles into an attractor state—the overall pattern of activation that provides the best global and integrated solution for the various inputs. These inputs would include visual cues of the face, but also potentially many other simultaneous inputs, such as visual cues of the body, vocal cues, motivations, task demands, among many others. Thus, the attractor dynamics of a recurrent connectionist network allow multiple sources of information—both bottom-up cues and top-down factors—to powerfully interact and integrate over time to produce stable person construals.

Structure of the Model

A diagram of the dynamic interactive model of person construal appears in Figure 1. It provides a general description of what specific instantiations of the model would involve, although specific instantiations of the model need not (and often will not) involve all elements appearing in Figure 1. The model has a recurrent connectionist architecture that may be classified as a stochastic interactive activation network (McClelland, 1991; Rumelhart et al., 1986). How the activation of a node changes over time is determined by three factors: the node's prior activation, how quickly this activation decays, and the net input of activation into the node from other nodes. We assume that excitation and inhibition summate algebraically, and that the influence of input on a node is dependent on the node's prior history of activation. We also assume that processing is stochastic rather than deterministic (see McClelland, 1991). On each iteration, therefore, the input to every node is altered by normally distributed random noise. Thus, the system's activation states are inherently probabilistic.

Before the presentation of each stimulus, activations of all nodes in the network are set equal to a resting activation value (zero), and external inputs are presented to certain
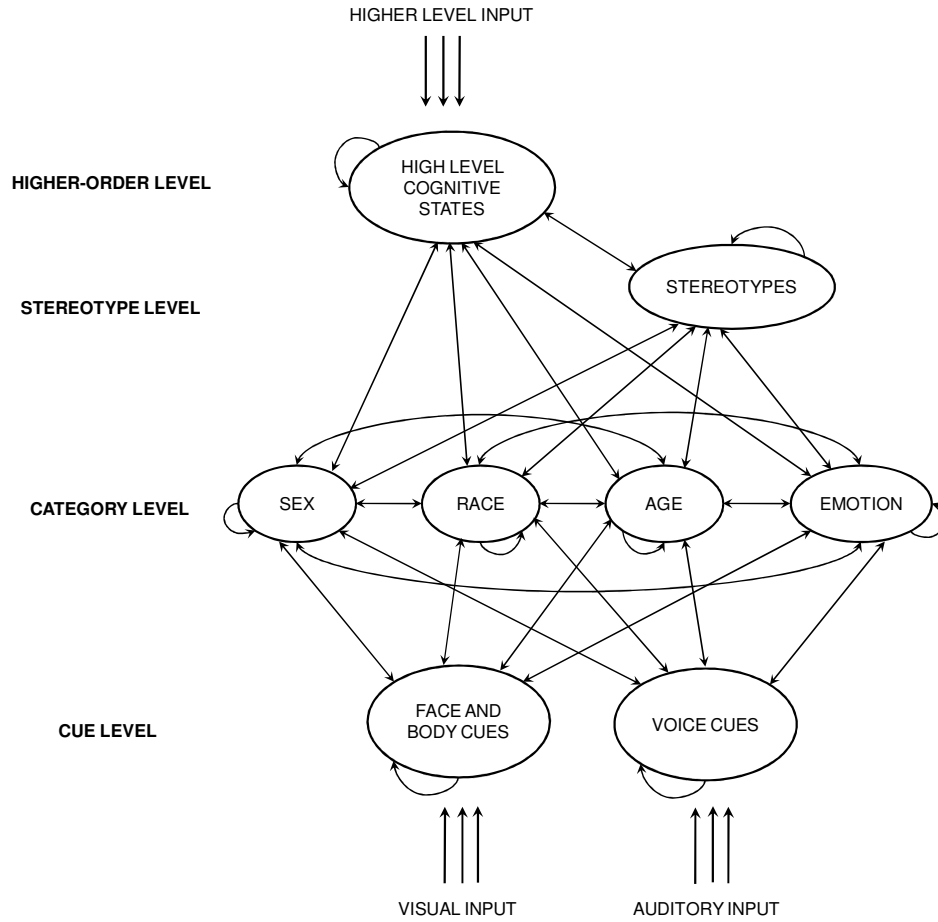
HIGHER LEVEL INPUT

HIGHER-ORDER LEVEL

HIGH LEVEL
COGNITIVE
STATES

STEREOTYPE LEVEL

STEREOTYPES

CATEGORY LEVEL

SEX   RACE   AGE   EMOTION

CUE LEVEL

FACE AND
BODY CUES

VOICE CUES

VISUAL INPUT   AUDITORY INPUT

Figure 1. A general diagram of the dynamic interactive model of person construal.

nodes for processing. When at resting activation level, a node is inactive and therefore assumed to not be represented in the processing landscape. Processing occurs over a number of iterations. On each iteration, each node computes its net input from the nodes connected to it based on their latest activation. Specifically, the net input to node i is:

$$net_i = \sum_j w_{ij} o_j + ext_i + \varepsilon_\sigma$$

where w_{ij} is the connection weight to node i from node j, o_j is the greater of 0 and the activation of node j, ext_i is any external input to node i, and ε_σ is a small amount of normally distributed random noise with mean 0 and standard deviation σ. Once the net input into all nodes has been computed, the activation of node i is updated as:

If $net_i > 0$:
$$\Delta a_i = I(M - a_i)net_i - D(a_i - r)$$

If $net_i \leq 0$:
$$\Delta a_i = I(a_i - m)net_i - D(a_i - r)$$

such that M is the maximum activation, m is the minimum activation, r is the resting activation level, I is a constant that scales the influence of external inputs on a node, and D is a constant that scales a node's tendency to decay back to rest. In all instantiations of our model, the parameters are as follows: $M = 1$, $m = -0.2$, $r = 0$, $I = 0.4$, $D = 0.1$, and $\sigma = 0.01$. These are standard values used in connectionist networks of this type (McClelland, 1991; Rumelhart et al., 1986). Connection weights will be specified for each instantiation of the model later.

We assume that the person construal system is organized into 4 interactive levels of processing: cue level, category level, stereotype level, and a higher-order level. Within each of these levels are one or several pools of nodes (Figure 1). Most nodes represent some feature or micro-hypothesis. For instance, the RACE pool would include a node for WHITE category and another node for BLACK category. Most of these pools are competitive in the sense that all the nodes are mutually exclusive and related by inhibitory connections. However, this is not necessarily the case for all pools. For instance, in the STEREOTYPES pool would be many nodes for different stereotypes. Some of these may inhibit one another (and thus be competitive), such as

AGGRESSIVE and NICE, whereas others might have no relationship with one another, and some others might excite one another, such as AGGRESSIVE and DANGEROUS. Nodes that excite another node have a positively weighted connection, nodes that do not influence another node have no connection (zero weight), and nodes that inhibit another node have a negatively weighted connection.

Each node has a transient level of activation at every moment in time. This level of activation corresponds with the strength of a tentative interpretation or hypothesis that the node is represented in the input (e.g., a face). Thus, in situations where a face is presented, the activation level of the MALE category node could be said to represent, at every moment in time, the strength of the hypothesis that the face is male. A node whose activation level exceeds a threshold excites other nodes with which it has an excitatory connection and inhibits other nodes with which it has an inhibitory connection. Importantly, most of the connections in our model are bidirectional, producing feedback and making the network highly interactive.

In simulations, the network's ultimate response is given by the response alternative associated with the node with the largest activation in a pool after a given amount of iterations (once the network has stabilized). The network's reaction time is given by the number of iterations it takes to for the winning node to reach 90% of its final activation state, which is then scaled by and added with constants to approximate human reaction time data (ms).

Cue level. The cue level contains a set of detectors for visual features (facial and bodily cues) and auditory features (vocal cues), which are directly stimulated by bottom-up sensory information of another person. The cue level contains two pools: a FACE/BODY CUES pool and a VOICE CUES pool. Sensory information of another person arriving in the visual system (facial and bodily cues) directly activates nodes in the FACE/BODY CUES pool. Sensory information arriving in the auditory system (vocal cues) directly activates nodes in the VOICE CUES pool. Depending on specific modeling interests, these pools have the flexibility to contain different arrangements of nodes. For instance, the FACE/BODY CUES pool could contain one node corresponding with all male facial features and another node corresponding with all female facial features. However, different strategies could be used. For instance, one node could describe a specific feature (e.g., LONG HAIR or DARK SKIN). Similarly, the VOICE CUES pool could contain a node corresponding with all male vocal features or it could contain a node corresponding with something specific such as FORMANT RATIO.

Nodes for cues that are along the same dimension (e.g., MALE CUES and FEMALE CUES) are related by mutually inhibitory connections because they compete for the same visual/auditory input. Thus, excitation of the MALE CUES node will inhibit the FEMALE CUES node, and vice-versa. Nodes that have no direct relationship with one another (e.g., LONG HAIR and DARK SKIN) have no connection

between them. Cue nodes excite all category nodes consistent with them and inhibit all of those inconsistent with them. For instance, the cue node for male facial features would activate the MALE category node and inhibit the FEMALE category node. Similarly, the cue node for female facial features would activate the FEMALE category node and inhibit the MALE category node. Note that the connections between cue nodes and category nodes are bidirectional. Thus, cue nodes both influence and are influenced by category nodes. This produces feedback and a recurrent flow of activation, as discussed earlier.

Category level. The category level contains a number of competitive pools that correspond with social category dimensions. For instance, in Figure 1, we have 4 pools: SEX, RACE, AGE, and EMOTION. Any number of different categories could be used, however (e.g., SOCIAL CLASS, SEXUAL ORIENTATION, OCCUPATION, ETHNICITY). These could include categories that are relatively static (e.g., sex) as well as categories that are dynamic (e.g., emotion). [1] Each of these pools contain category nodes. The pool for SEX would include a MALE node and a FEMALE node; the pool for RACE would include, for example, a WHITE node, a BLACK node, and an ASIAN node. Nodes within a pool compete with one another through mutual inhibition. In the broad model depicted in Figure 1, bidirectional connections exist between all 4 of the category pools. This is not required for all instances of the model, but they are depicted because in some instances category nodes may be directly related to one another. For instance, if perceivers have learned in their lifetime that women tend to be happy and men tend to be angry (see Fabes & Martin, 1991), then the node for MALE (in the SEX pool) may have a bidirectional excitatory connection with ANGRY (in the EMOTION pool). Similarly, the node for FEMALE may have a bidirectional excitatory connection with HAPPY.

Category nodes receive input from cue nodes (which directly receive bottom-up sensory information) and they also send feedback to cue nodes. Category nodes activate stereotype nodes (e.g., MALE excites AGGRESSIVE and FEMALE excites DOCILE), and they also receive feedback from these nodes as well. Thus, not only will the category node, MALE, tend to activate the stereotype node, AGGRESSIVE, but activation of AGGRESSIVE will tend to activate the MALE category. This type of feedback is important, and we will discuss it in detail later. Finally, category nodes may activate and be activated by higher-order nodes.

Stereotype level. The stereotype level contains one pool including nodes for all category-related stereotypes (e.g., AGGRESSIVE or DOCILE). Within this, nodes could mutually

---

[1] Although a dynamic characteristic, such as emotion, changes over time, it exhibits categorical perception effects similar to those of static characteristics, and the perception of emotion is a form of perceptual categorization (Calder, Young, Perrett, Etcoff, & Rowland, 1996; Etcoff & Magee, 1992).

inhibit or mutually excite one another. For instance, AGGRESSIVE and DANGEROUS would mutually excite one another, but AGGRESSIVE and DOCILE may mutually inhibit one another. Stereotype nodes receive input from category nodes and send feedback to them. Stereotype nodes also receive input from higher-order nodes and send feedback to them as well.

Higher-order level. Nodes in this level may correspond with any number of high-level cognitive states, depending on what is being modeled. They could include factors such as prejudice, motivations, processing goals, task demands, among others. We assume that these nodes receive direct input from higher levels of mental processing (e.g., motivational systems or top-down attentional systems). Higher-order nodes may influence category nodes or stereotype nodes, or both. Moreover, they may have a bidirectional connection with these nodes or simply a unidirectional top-down connection only.

For instance, higher-order nodes could be used to model high-level task demands in a particular context. One higher-order node could denote SEX TASK DEMAND and another node could denote RACE TASK DEMAND. During a sex categorization task, the higher-order SEX TASK DEMAND node would be directly activated by higher-level input (e.g., top-down attentional systems, driven by memory of task instructions). Activation of this higher-order node would then have top-down excitatory connections with sex-related category nodes (MALE and FEMALE), but have top-down inhibitory connections with race-related category nodes (WHITE, BLACK, ASIAN), since the task demand compels attention to sex and away from race. As such, attentional effects due to task demands (e.g., placing attention on sex and away from race in a sex categorization task) emerge out of the flows of activation between these higher-order task-demand nodes and the category nodes, consistent with other computational models accounting for task demands (e.g., Cohen & Huston, 1994). This is one example of how the higher-order level could be used to model top-down effects from internal cognitive states, such as task demands, memory, affect, motivations, expectations, situational context, among others.

An Example

We now consider a specific instantiation of the model involving the category and stereotype activation of sex and race (Figure 2). We only use a select amount of the pools from the general model (Figure 1). Namely, we only use the FACE/BODY CUES pool, the SEX category pool, RACE category pool, STEREOTYPES pool, and HIGH-LEVEL pool. Solid-line connections with arrows are excitatory (positive weight) and dashed-line connections with dots are inhibitory (negative weight). Arrows and dots indicate the direction of influence (in this instantiation, all influences are bidirectional).

Let us consider the dynamics of the network when it is presented with a face. Visual input of the face directly

activates nodes in the cue level. For simplicity, we use individual cue nodes to represent all facial features associated with a category. Thus, there is a cue node for MALE CUES, FEMALE CUES, BLACK CUES, WHITE CUES, and ASIAN CUES. The cue nodes specifying a target's sex (MALE CUES and FEMALE CUES) nodes mutually inhibit one another, and the cue nodes specifying a target's race (BLACK CUES, WHITE CUES, and ASIAN CUES) also mutually inhibit one another. Cue nodes excite category nodes consistent with them and inhibit category nodes inconsistent with them. They also receive feedback from category nodes. At the same time that cue nodes receive input from visual processing, higher-level input directly activates higher-order nodes. Here, we use one node to denote a task-induced state that compels excitation of the sex-category dimension and another node to denote a task-induced state that compels excitation of the race-category dimension. The higher-level input in this case would originate from top-down attentional systems driven by memory of the task instructions. These higher-order SEX TASK DEMAND and RACE TASK DEMAND nodes mutually inhibit one another. Moreover, they excite category nodes consistent with them, inhibit category nodes inconsistent with them, and are also activated by category nodes as well. Thus, activation of the RACE TASK DEMAND node would facilitate activation of race categories (BLACK, WHITE, ASIAN) and inhibit activation for sex categories (MALE, FEMALE), and vice-versa for the SEX TASK DEMAND node.

As category nodes are activated by cue nodes and higher-order nodes, they also excite stereotype nodes consistent with them and inhibit stereotype nodes inconsistent with them. Here we use two stereotype nodes: AGGRESSIVE and DOCILE. These nodes mutually inhibit one another. Further, as category nodes excite and inhibit stereotype nodes, they are also updated by feedback from the stereotype nodes. Many more stereotype nodes could be included in the model to capture the full gamut of a category's stereotype contents, but for simplicity we use AGGRESSIVE as one example of male-related and Black-related stereotypes and DOCILE as one example of female-related and Asian-related stereotypes. We will describe simulations of this model that can predict several phenomena found in human perceivers. Before this, however, first we describe important properties of the general model (Figure 1).

Properties of the Model

Dynamic, probabilistic, and mutually interactive representations. The representations in our model are interactive, rather than rigid and independent. This is because our model assumes the person construal system experiences large ongoing cycles of interaction between each representation in the system. Thus, the activation of one representation influences all other representations in the system, as one node's activation influences the activation of all other nodes. For example, consider the presentation of a
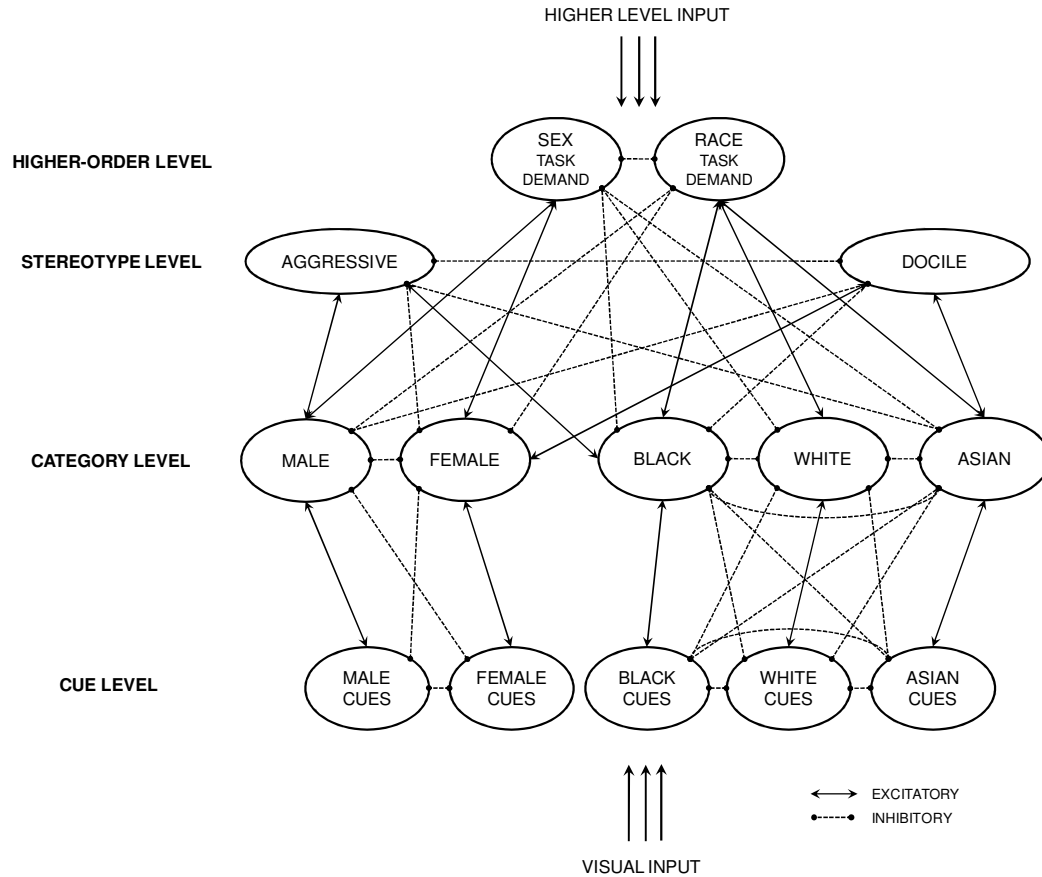
Figure 2. An instantiation of the model, used for simulations to account for Phenomena 1 and 3.

male face. Direct stimulation of the MALE CUES node will facilitate the MALE category, which will inhibit the FEMALE category, which will inhibit the stereotype, DOCILE, in turn inhibiting the category, ASIAN, in turn facilitating the category, BLACK, in turn facilitating the stereotype, AGGRESSIVE, which facilitates the category, MALE, which facilitates the MALE CUES node, and so on and so forth. Thus, the system is highly interactive. These influences gradually taper off as all nodes in the system come to settle into an attractor state. Before the system stabilizes, however, representations are in continuous interaction over time, not encapsulated and independent.

The representations in our model are dynamically and probabilistically reconstructed in every new instance, rather than remaining static and independent. Their real-time development is in continuous interaction with other activations across the system, both dynamically influenced by these activations and a source of influence over them. For example, there is no stand-still, discrete symbol-like representation of the "male" category. Rather, the system will gravitate toward an attractor state that involves stable, strong activation of MALE and weak activation of FEMALE, but this state is not a discrete symbol identically activated

every time the system encounters a male target. Instead, the system's prior history, external inputs, simultaneous activations, internal constraints, and a bit of random noise all work to determine the probabilistic activation of the MALE and FEMALE categories. Thus, the system may frequently visit a similar attractor state involving strong activation of MALE and weak activation of FEMALE every time it encounters a male target. But this is a dynamically reconstructed state of activation that could only approximate an idealized, linguistically identifiable representation of the "male" category (Spivey & Dale, 2006). Even if the system did reach, for example, some idealized attractor state involving 100% MALE and 0% FEMALE, the system does not have much time to dwell there since the ongoing accrual of new sensory information (e.g., facial, vocal, bodily cues) already begins pushing the system into different attractor states to which it must start gravitating. Thus, our model assumes that internal representations of categories and stereotypes as dynamic, probabilistic, and mutually interactive.[2]

Perceiving in a noisy social environment and partial fit. A noteworthy property of a dynamic interactive model is that external stimuli do not need to directly contain all the

perceptual cues required to correctly identify a target's characteristics. Instead, just a few cues can lead the system to generate automatic hypotheses about the stimulus and, under certain conditions, to "run" with those hypotheses. For instance, a fleeting glimpse of a person's face at an obscure angle on a busy street flooded with other people may not directly contain all the cues necessary to discern with confidence whether the person is a man or woman. Nonetheless, the meager amount of cues that are processed by the visual system will activate relevant cue nodes, which will thereafter place excitatory and inhibitory pressures on category nodes. If these pressures are sufficient and if network constraints permit, the MALE and/or FEMALE categories will be pushed above their resting level. If both categories are pushed above their resting level, they will compete with one another to stabilize onto one. Random noise could easily bias the competition toward one category or the other, given the ambiguity in the bottom-up input. If just enough perceptual cues were processed, however, and one category received greater excitation over the other, the system could settle into the correct interpretation. Thus, a mere partial fit in a target's available perceptual cues may be sufficient for the system to generate automatic hypotheses and then potentially commit to them.

Our social environment is rife with these situations involving partial fit. On a busy street, just enough visual information of a target's face might activate relevant cue nodes that then put slight pressure on category nodes, in turn triggering "best guess" partially-active representations of the target's category memberships. As one's eyes rapidly move on to the next passing face, a similar process would occur. Indeed, previous work has shown that minimal, isolated category-specifying perceptual cues (e.g., long hair or short hair) are sufficient to trigger category representations (e.g., female or male, respectively; Macrae

---

[2] A caveat is due regarding our model's use of localist representations (i.e., that each node corresponds with one linguistically identifiable representation, e.g., male category). While many connectionist models use distributed representations, there are some advantages to using localist representations. In contrast to localist models, distributed models map a single identifiable representation (e.g., male category) not to a single local node, but to a particular pattern of activation distributed across several nodes. Indeed, we acknowledge that using localist representations is one step further away from neural plausibility. It is certainly our assumption that representations in the human brain are inherently distributed, with any identifiable representation (e.g., male category) corresponding to a pattern of firing rates—or population code—involving a large number of neurons (Olshausen & Field, 2004; Rumelhart et al., 1986). However, localist representations provide reasonable approximations of distributed representations while also being more intuitive to understand for the purposes of modeling (Grainger & Jacobs, 1998; Smolensky, 1989). Thus, we employ localist representations here for the sake of simplicity, but we assume that these are only useful approximations of actual distributed representations found in neural systems.

& Martin, 2007).

Partial parallel activation and dynamic competition. A central feature of a dynamic interactive model of person construal is that processing involves dynamic competition between partially-active and parallel representations. This is due to the continuous dynamics intrinsic to the person construal system. Perceptual processing triggers partially-active category representations (e.g., "he's [tentatively] male") that continuously compete. Ongoing changes of partially-active, competing category representations, in turn, continuously update stereotype knowledge and higher-order states (while also returning feedback to lower levels of processing). Eventually, these partially-active category representations settle into an attractor state, which in experimental settings (e.g., a sex categorization task) often results in a single, stable categorical outcome (e.g., "that's a man!").

Such continuous dynamics are inherent in our model's structure. For instance, when a face is presented to the network, its visual input begins activating cue nodes, which in turn places excitatory pressures on category nodes consistent with those cues and inhibitory pressures on category nodes inconsistent with them. Importantly, these pressures operate gradually and continuously over time. On each iteration, activated cue nodes update the activation of category nodes by either strengthening or weakening them. At the same time, category nodes are also strengthened or weakened by higher-order nodes and stereotype nodes. They also engage in mutual inhibition with other category nodes (e.g., MALE and FEMALE category nodes compete for activation). Thus, for instance, it might take 60 iterations for the MALE category node to achieve its maximum asymptotic level of activation. But for the vast majority of processing (iterations) prior to that, the MALE category would be partially-active and continuously evolving over time. During this evolution, the MALE category would be continuously incorporating excitatory and inhibitory inputs from a variety of lower nodes (cue nodes), higher nodes (stereotype and higher-order nodes), and other category nodes (e.g., FEMALE category), while also, in turn, feeding activation back to those nodes as well.

Further, these continuously fluctuating and partially-active representations are activated in parallel. For instance, imagine that the face of a feminine-looking man (e.g., .55 male, .45 female) is presented to the network. Visual input will activate the node for MALE CUES and also the node for FEMALE CUES. Activation of cue nodes will begin placing excitatory and inhibitory pressures on category nodes. In this case, both the MALE and FEMALE category nodes will rise above their resting level and become partially-active. These parallel and partially-active category representations will then compete with one another through mutual inhibition. For the system to settle into a stable state (e.g., for sex categorization, either predominantly MALE being active or predominantly FEMALE being active), the parallel and partially-active MALE and FEMALE category nodes must

engage in a dynamic competition, with one gradually gaining activation and the other gradually dying off, as they strangle each other's activation through inhibition. In such a case, the system is simultaneously attracted to be in two different states (i.e., attractor states): one state involving ~100% MALE/~0% FEMALE and another involving ~0% MALE/~100% FEMALE. Such states are highly stable (leading the system to be attracted to them), whereas a state such as ~55% MALE/~45% FEMALE is highly unstable. Although in most cases the system would come to settle into an attractor state involving ~100% MALE/~0% FEMALE activation, thereby achieving a male categorization, along the way it could not help but continually meander near the other attractor (~0% MALE/~100% FEMALE) while the competition is still resolving itself. In short, a dynamic interactive model of person construal assumes many representations are partially active in parallel. Further, conflicts within the same pool (e.g., sex categories; race categories) are resolved through dynamic competition, where all nodes within the pool mutually inhibit one another.

Overview

Across 5 simulations, we show how a dynamic interactive model naturally accounts for a wide range of person construal phenomena. First, we show how it replicates effects of continuous dynamics, partial parallel activation, and dynamic competition in categorization and stereotyping (Phenomenon 1). We then show how it replicates effects of top-down influences of high-level cognitive states and stereotype activation on categorization (Phenomena 2 and 3) as well as bottom-up influences on categorization due to shared perceptual features (Phenomenon 4). Finally, we show how the model replicates contextual and cross-modal effects on categorization (Phenomenon 5). These recently documented phenomena provide converging evidence for our theory that person construal is a dynamic interactive process. The simulations that follow will show how our model captures this process.

Dynamics of Social Categorization and Stereotyping

Faces have been found to trigger simultaneously and partially-active sex categories (Freeman et al., 2008), race categories (Freeman et al., 2010b), and stereotypes (Freeman & Ambady, 2009), which gradually resolve into stable categorical perceptions through dynamic competition. To capture these simultaneously conflicting representations, the person construal process has been examined on-line by recording participants' computer mouse movements en route to responses on the screen. For instance, in one series of studies, participants categorized the sex of male and female faces by moving the computer mouse from the bottom-center of the screen to either the top-left or top-right corners, which were marked "male" and ''female'' (Freeman

et al., 2008). Participants were asked to click on the correct sex category. Meanwhile, their mouse movements were recorded. When categorizing sex-atypical male and female faces (those that contained partial cues of the opposite sex), participants' mouse movements were continuously attracted toward the opposite sex-category (on the opposite side of the screen). For instance, when categorizing a male face that contained some feminine features, participants' mouse movements gravitated a bit closer to the ''female'' response than when categorizing a male face without feminine features. This indicated that sex categorization involved partially-active category representations (male and female), which simultaneously competed over time to gradually stabilize on one categorical outcome.

Such a pattern of results obtained also for race categorization (Freeman et al., 2010b) and stereotype activation (Freeman & Ambady, 2009). The main mouse-tracking results from the stereotype activation study appear in Figure 3. Participants were presented with sex-typical and sex-atypical faces and were instructed to move the mouse and click on the adjective that was stereotypically appropriate for the face (one was always masculine and one always feminine). For sex-atypical faces (which bore a mixture of masculine and feminine cues) the mouse was continuously attracted toward the opposite sex stereotype (e.g., "docile" for a male target) before settling into the correct stereotype (e.g., "aggressive" for a male target), as seen in Figure 3. This finding provided evidence that faces trigger parallel and partially-active stereotypes tied to alternate social categories. These stereotypes then dynamically compete over time to settle onto one (Freeman & Ambady, 2009). Below, we demonstrate how a dynamic interactive model of person construal naturally accounts for these findings. We ran simulations using the instantiation of the model introduced earlier (Figure 2).

Phenomenon 1: The Partial and Parallel Activation of Social Categories and Stereotypes

First, we consider how the model categorizes faces by sex. We consider categorization of two types of targets: a sex-typical White male face and a sex-atypical White male face. Connection weights for this instantiation of the model (Figure 2) are provided in Table S1 of the supplementary materials.[3]

Because this is a sex categorization task and the demands of the task compel attention to sex, higher-level input would directly activate the SEX TASK DEMAND node. We set the higher-level input into the SEX TASK DEMAND node at .9 and higher-level input into the RACE TASK DEMAND node at .1 (see Footnote 3). This simulates the task context of sex categorization, where perceivers would be focusing on targets' sex over their race. This thus facilitates activation of MALE and FEMALE category nodes, and inhibits activation of BLACK, WHITE, and ASIAN nodes. In the cue level, the MALE CUES and FEMALE CUES nodes both receive direct input from visual processing of the face. To
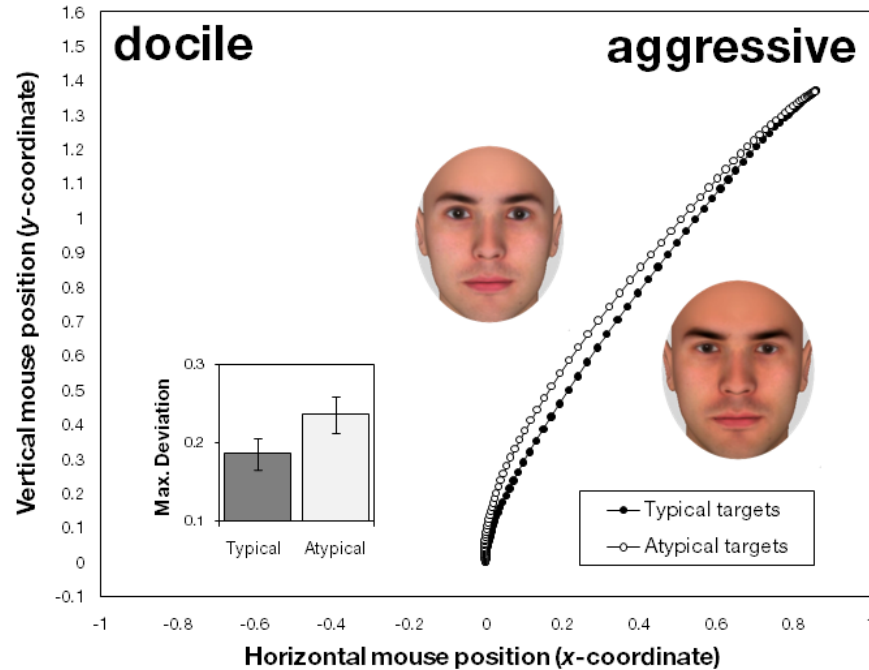
Figure 3. In one study, Freeman and Ambady (2009) found that when stereotyping sex-atypical faces (e.g., a man with feminized cues), participants' computer mouse trajectories were continuously attracted to the stereotype of the opposite sex (e.g., *docile*) before settling into their ultimate response (e.g., *aggressive*). The figure depicts mean mouse trajectories (aggregated across male and female targets). The feminine and masculine stereotype labels appeared in the top -left and top-right corners of the screen (*docile* and *aggressive*, shown here, are examples of the stereotype words that were displayed). In this figure, trajectories for all targets were remapped rightward, with the opposite-sex stereotype on the left and the stereotype label consistent with the target's sex on the right. Sample male face stimuli are also displayed. A typical male face is shown on the right, next to the mean trajectory for typical targets. Its atypical (feminized) counterpart is shown on the left, next to the mean trajectory for atypical targets. During an actual trial, a single face was centered at the bottom of the screen. The bar graph shows trajectories' maximum deviation toward the opposite-sex stereotype from a direct line between trajectories' start and end points, separately for typical and atypical targets (error bars denote standard errors of the mean).

simulate the presentation of a sex-typical White male face, we set visual input into the MALE CUES node at .95 and visual input into the FEMALE CUES node at .05. Thus, this face is inherently 95% masculine and 5% feminine. Because the face is White, we set visual input into the WHITE CUES node at .95 and visual input into the BLACK CUES and ASIAN CUES nodes at .025 each. We ran the simulation 100 times each time for 150 iterations, and plotted the average

activation level of each category node over time, appearing in Figure 4A.

The presentation of a sex-typical White male face sets a process into motion, in which visual processing of the face directly activates cue nodes. Cue nodes inconsistent with one another, such as the MALE CUES and FEMALE CUES nodes, compete for the visual input. The activation of cue nodes, in turn, immediately places excitatory and inhibitory pressures on category nodes (Figure 2). In this case, the highly activated MALE CUES node places strong excitatory pressure on the MALE category node and inhibitory pressure on the FEMALE category node. The highly activated WHITE CUES node places strong excitatory pressure on the WHITE category node and inhibitory pressure on the BLACK and ASIAN category nodes. At the same time, the higher-order SEX TASK DEMAND node places excitatory pressures on the MALE and FEMALE category nodes and inhibitory pressures on the BLACK, WHITE, and ASIAN category nodes. These simultaneous pressures cause the activation levels of some

---

[3] In all simulations, we set connection weights and input values according to our intuitions regarding stimulus and task features. It may be possible in future work to derive these values empirically. However, we are confident given previous studies that our parameters are in accord with participant judgments and task features in these contexts, and we chose parameters that best reflect these intuitions. In this sense, the current simulations serve as existence proofs for the kind of dynamic interactive processing that may take place during construal, though we acknowledge that future work may advance these simulations by deriving network parameters empirically.

category nodes to be pushed above their resting levels, whereas others are inhibited and pushed below their resting levels. Excitatory pressure from both the MALE CUES node and the higher-order SEX TASK DEMAND node leads the MALE category node to rise above its resting level. Positive feedback is then produced between these nodes, which causes the MALE category node to rapidly gain activation until gradually settling into a stable state. Because a small amount of feminine features were presented to the network (the FEMALE CUES node was initialized with .05 visual input), the FEMALE category node also becomes slightly active for a very brief moment early on, and then succumbs to strong inhibition from the MALE CUES node and the MALE category node, resulting in it being pushed below its resting level. Excitatory pressure from the WHITE CUES node leads the WHITE category node to rise above its resting level, but the WHITE category node is also inhibited by the SEX TASK DEMAND node. This leads the WHITE category node to gain a meager amount of activation until eventually settling into a stable state (and thus MALE is more strongly active than WHITE). Finally, inhibitory pressures from the WHITE CUES node and the SEX TASK DEMAND node lead the BLACK and ASIAN category nodes to be rapidly pushed below their resting levels. These dynamics are apparent in Figure 4A.

Note how each category node gradually works over time to settle into a stable attractor state, such that its activation reaches some asymptotic level and tapers off. This stable state would correspond with the fully confident categorization of the target as male. However, before that 100% confident categorization is achieved, bear in mind that partial, tentative evidence for that categorization actually accumulates gradually over time (the dynamics of the MALE

category activation).

Now let us consider how the person construal system settles into a stable state when presented with a sex-atypical White male face in a sex categorization task. To simulate this, we set visual input into the MALE CUES node at .55, input into the FEMALE CUES node at .45, input into the WHITE CUES node at .95, and input into the BLACK CUES and ASIAN CUES nodes at .025 each. As done previously, we set higher-level input into the SEX TASK DEMAND node at .9 and input into the RACE TASK DEMAND node at .1. This simulates attention on sex induced by the task context of sex categorization. The simulation was run 100 times, and the averaged activation level of each category node over 150 iterations appears in Figure 4A.

The activated MALE CUES node begins exciting the MALE category and inhibiting the FEMALE category, while the FEMALE CUES node begins exciting the FEMALE category and inhibiting the MALE category (Figure 2). The MALE CUES and FEMALE CUES nodes also begin inhibiting one another as well. The highly activated WHITE CUES node excites the WHITE category and inhibits the BLACK and ASIAN categories. At the same time, the higher-order SEX TASK DEMAND node excites the MALE and FEMALE categories and inhibits the BLACK, WHITE, and ASIAN categories. The excitatory pressure from both the MALE CUES node and the higher-order SEX TASK DEMAND node leads the MALE category to rise above its resting level. The excitatory pressure from the FEMALE CUES node and the SEX TASK DEMAND node also leads the FEMALE category to rise above its resting level. Pressures from the cue nodes and higher-order nodes cause the WHITE category to gain a meager amount of activation and the BLACK and ASIAN categories to be rapidly pushed below their resting levels.
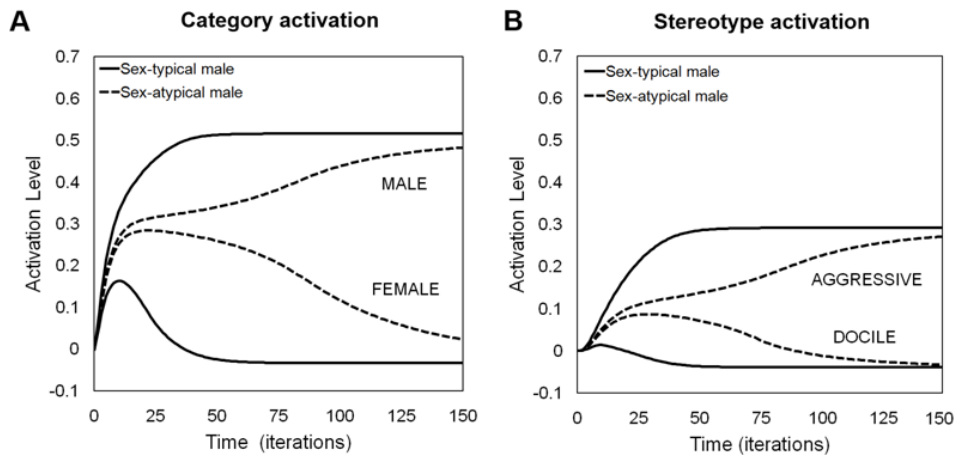
Figure 4. (A) The activation level of the MALE and FEMALE category nodes as a function of time (iterations) following the presentation of a sex-typical male face (solid lines) and sex-atypical male face (dashed lines). (B) The activation level of the AGGRESSIVE (male-related stereotype) and DOCILE (female-related stereotype) nodes are plotted as a function of time (iterations) following the presentation of a sex-typical male face (solid lines) and sex-atypical male face (dashed lines).

With the MALE and FEMALE categories now simultaneously activated, they begin competing with one another through mutual inhibition. Over time, this mutual inhibition between competing MALE and FEMALE categories, in addition to feedback with the cue nodes, leads the FEMALE category to gradually decay while the MALE category gradually rises in activation until a stable state is achieved. This results in the MALE category winning the competition, whereas the FEMALE category dies off and stabilizes on a very weak level of activation. Thus, simultaneously and partially-active sex categories dynamically compete over time to settle onto a single categorical outcome (in this case, a male categorization).

Indeed, such temporally dynamic competition has been shown to underlie sex and race categorization, as described earlier. For instance, when categorizing a sex-atypical male face (bearing partial feminine features), participants' mouse trajectories showed a continuous attraction toward the "female" response before settling into the "male" response. Thus, a female category representation was simultaneously and partially active across construal, which led the mouse to partially curve toward the "female" response on its way to correctly clicking on the "male" response (Freeman et al., 2008). This is precisely what is reflected in Figure 4A, which shows that the FEMALE category node was partially active, simultaneously with the MALE category's activation, until the system settled into a stable state involving strong MALE category activation and the FEMALE category below resting level.

The different stable states that the system settles into in construing a sex-typical versus a sex-atypical male face is noteworthy. When presented with a sex-typical male face, the MALE category stabilized at .52 activation and the FEMALE category at −.03 activation. When presented with a more ambiguous sex-atypical male face, however, the MALE category stabilized at a weaker level of activation (.48) and the FEMALE category at a relatively stronger level of activation (.02). Thus, although the system took an ambiguous mixture of masculine and feminine facial cues and, over time, slotted it into a single categorical outcome (male) through dynamic competition, the outcome is nonetheless graded. A sex-typical male face resulted in stronger activation of the MALE category and weaker activation of the FEMALE category, whereas a sex-atypical male face resulted in less strong activation of the MALE category and less weak activation of the FEMALE category. Indeed, recent findings in person construal research have suggested that steady-state category representations are graded (Blair, Chapleau, & Judd, 2005; Blair, Judd, & Fallman, 2004b; Blair, Judd, Sadler, & Jenkins, 2002; Freeman et al., 2008; Freeman et al., 2010b; Locke, Macrae, & Eaton, 2005; Maddox & Gray, 2002). As the presence of category-specifying facial cues increase (i.e., become more prototypical of a social category), steady-state category representations increase in strength. That representations of social categories are inherently graded has been additionally suggested by neuroimaging work (Freeman, Rule, Adams, & Ambady, 2010c). Further, once triggered these graded representations thereafter influence evaluation and behavior in graded fashion as well (Livingston & Brewer, 2002). For instance, in court trials, individuals with more Black-specifying features are punished more severely and more likely to be sentenced to death (Blair, Judd, & Chapleau, 2004a; Johnson, Eberhardt, Davies, & Purdie-Vaughns, 2006). Thus, a dynamic interactive model naturally accounts for such findings.[4]

What are the implications of the partial and parallel activation of social categories for the activation of stereotypes? Figure 4B shows the level of activation of the masculine stereotype, AGGRESSIVE, and the feminine stereotype, DOCILE (averaged from 100 simulations), across 150 iterations when the network is presented with a sex-typical versus a sex-atypical White male face. As category nodes become activated through excitatory pressures of cue nodes and higher-order nodes, they immediately start placing excitatory and inhibitory pressures on stereotype nodes. For a sex-typical male face, activation of the MALE category in turn activates the AGGRESSIVE stereotype and inhibits the DOCILE stereotype. Positive feedback between the AGGRESSIVE stereotype and MALE category then leads the system to rapidly converge on a stable state involving strong activation of the AGGRESSIVE stereotype and the DOCILE stereotype pushed below resting level. When presented with a sex-atypical male face, however, the activated MALE category node excites the AGGRESSIVE stereotype and inhibits the DOCILE stereotype. However, the simultaneously activated FEMALE category also excites the DOCILE stereotype and inhibits the AGGRESSIVE stereotype. With both masculine and feminine stereotypes activated,

---

[4] Whereas it is sometimes acknowledged in the literature that faces trigger category representations that are graded (Locke et al., 2005), it is often not considered that alternate categories may be partially-active at the same time and that these partially-active representations are graded as well. For instance, as a face becomes less prototypically male (and starts featuring some feminine cues), our model predicts that not only will a representation of the male category become weaker, but also that a simultaneously active representation of the female category will become stronger. This is evident in the results of our simulation (Figure 4A) and also with experimental investigations. For instance, in one study, participants categorized the race (White versus Black) of real faces that varied along a continuum of racial ambiguity (Freeman et al., 2010b). As the amount of Black-specifying features on faces which were ultimately judged as White linearly increased, the hand's simultaneous attraction toward the "Black" response (before settling into the "White" response) increased as well. Similarly, as the amount of White-specifying features on faces which were ultimately judged as Black linearly increased, the hand's simultaneous attraction toward the "White" response (before settling into the "Black" response) increased. Thus, as perceptual cues of an alternate social category increase, the partial activation of that alternate category increases as well, due to the dynamic competition inherent to social categorization.

they then compete with one another through mutual inhibition. Thus, as seen in Figure 4B, a sex-atypical male face leads to the partial and parallel activation of masculine (AGGRESSIVE) and feminine (DOCILE) stereotypes, which dynamically compete until the system achieves a stable state. Indeed, precisely this effect has been empirically demonstrated in human perceivers (Freeman & Ambady, 2009).

In summary, our model accounts for the partial and parallel activation of social categories and stereotypes, as well as the dynamic competition required to resolve these simultaneously conflicting activations into stable person construals.

Top-Down Interactivity in Social Categorization

A dynamic interactive theory of person construal assumes that processing in the category level and stereotype level are interactive, such that, beyond categories obviously feeding forward activation to stereotypes, stereotypes can also feed back activation to categories, thereby exerting top-down pressure on categorization. Recent studies have highlighted several top-down effects on social categorization.

For instance, explicitly labeling a target as "White" or "Black" influences the perception of faces, and this is more strongly the case for perceivers who believe race is fixed rather than malleable (Eberhardt et al., 2003). The perception of a face's race and the subsequent memory of it is also contingent on perceivers' motivation to include ambiguous group members into their in-group or to exclude them from it (Castano, Yzerbyt, Bourguignon, & Seron, 2002; Pauker et al., 2009).

Racial prejudice has numerous top-down effects on race perception. For instance, it moderates the manner by which facial emotion (angry or happy) shapes processing of race (Hugenberg & Bodenhausen, 2004). Moreover, high levels of prejudice lead to less efficient categorizations of racially-ambiguous faces (Blascovich, Wyer, Swart, & Kibler, 1997), and a bias to categorize racially-ambiguous faces as part of the out-group rather than in-group (Pettigrew, Allport, & Bartnett, 1958). How much an individual identifies with his or her in-group also exerts an influence on race categorization. For instance, individuals who strongly identify with their in-group are more likely to exclude racially-ambiguous faces from their in-group and are less efficient in race categorization (Castano et al., 2002).

In categorizing a face's emotion, top-down knowledge (e.g., knowing that a target is watching a horror film vs. comedy show) is readily utilized to resolve ambiguous fearful-happy emotional expressions (Trope & Cohen, 1989). Other top-down knowledge, such as an explicit label, also constrains the activation of emotion categories. In one study, participants encoded emotionally ambiguous faces while given an explicit label such as "angry" or "happy"

(Halberstadt & Niedenthal, 2001). Faces that were paired with an angry label were subsequently remembered as more angry, just as faces paired with a happy label were remembered as more happy. Further data ruled out the possibility that these influences of top-down semantic context were due to post-perceptual processes such as memory reconstruction (Halberstadt, 2005).

More recently, Halberstadt, Winkielman, Niedenthal, and Dalle (2009) used facial electromyography (EMG) to demonstrate that explicit labels of "angry" or "happy" induced spontaneous emotion-specific mimicry during on-line face processing, and that EMG activity predicted subsequent memory bias. This suggests activation of emotion categories are flexibly shaped by top-down cues even early in processing. One's social expectations also exert a top-down influence on the perception of facial emotion. For instance, women who were stigma conscious (i.e., who chronically expect to be rejected by men) reported seeing a greater amount of contempt on male faces than female faces, in contrast to women who were low in stigma consciousness (Inzlicht, Kaiser, & Major, 2008). The perception of facial emotion can also be constrained by individuals' emotional states. For instance, participants who were induced to feel happy reported seeing happiness for a longer period of time when viewing dynamic face morphs transitioning between happiness and sadness (Niedenthal, Halberstadt, Margolin, & Innes-Ker, 2000). In short, a growing body of research finds that high-level factors can exert a variety of top-down influences on the perception of other people.

Phenomenon 2: Top-down effects of racial prejudice and facial emotion on race perception

An excellent illustration of top-down interactivity in person construal is the influence of racial prejudice and emotion category on race perception. As described earlier, Black individuals are stereotyped as hostile (Devine, 1989). Hugenberg and Bodenhausen (2004) reasoned that this "Black is hostile" stereotype would lead perceptions of race to be susceptible to influences of emotion category. Specifically, they argued that faces should be more likely to perceived as Black when displaying a hostile emotional expression (e.g., anger) relative to a non-hostile one (e.g., happy). Further, because individuals with higher levels of racial prejudice more readily activate and apply stereotypes (Lepore & Brown, 1997; Wittenbrink, Judd, & Park, 1997), this stereotype-mediated race-emotion interaction should be stronger in high-prejudice individuals (who would more readily activate the "Black is hostile" stereotype) and should be weaker in low-prejudice individuals. Clearly, emotion category would not dramatically alter the perception of race on faces for which race is quite obvious. However, when racial cues are substantially ambiguous, this bottom-up ambiguity opens up the opportunity for top-down factors (e.g., stereotypes) to exert a strong bias on race-category activation. Indeed, Hugenberg and Bodenhausen (2004)
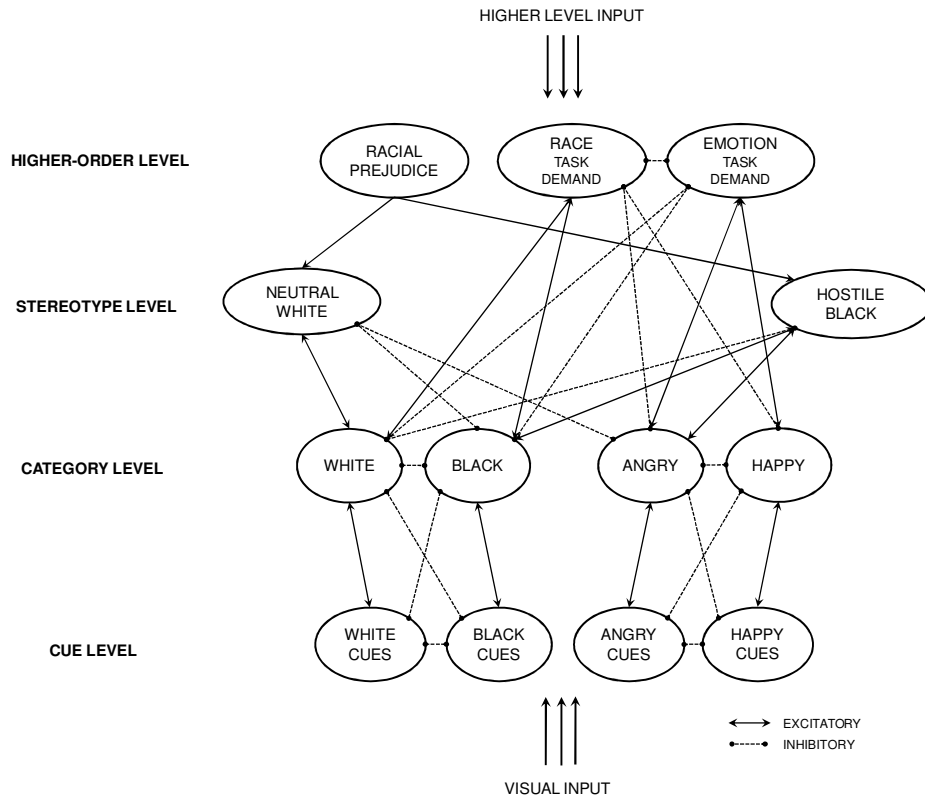
Figure 5. An instantiation of the model, used for simulations to account for Phenomenon 2.

found that, for high-prejudice individuals, racially-ambiguous faces were more likely to be categorized as Black (relative to White) when displaying anger (hostile emotion) than when displaying happiness (non-hostile emotion). For low-prejudice individuals, however, emotion category did not reliably modulate race categorization (presumably because the "Black is hostile" stereotype was not substantially activated for low-prejudice individuals). This finding is a compelling example of the interactive nature of person construal, showing how activation of one category dimension (emotion) fluidly interacts with another category dimension (race), and how these cross-category interactions may be driven by higher-level social cognitive processes, such as stereotype activation and prejudice.

To account for these effects, we developed another instantiation (Figure 5) of our general model (Figure 1). Connection weights are provided in Table S2 of the supplementary materials. As in the previous instantiation of the model (Figure 2), for simplicity, we use individual cue nodes to represent all facial features associated with a category. The cue, category, and higher-order levels are modeled in a similar fashion as the previous instantiation, except here with race and emotion. In this instantiation, however, there is a third higher-order node: RACIAL PREJUDICE. This node is a simplified way of simulating the complex set of memory and affect structures involved in racial prejudice. For high-prejudice individuals, this node will be strongly activated; for low-prejudice individuals, this

will be weakly activated. This RACIAL PREJUDICE node has unidirectional excitatory connections with two stereotype nodes: the HOSTILE BLACK node and the NEUTRAL WHITE node. The HOSTILE BLACK node represents the "Black is hostile" stereotype, and the NEUTRAL WHITE node represents the "White is neutral [non-hostile]" stereotype. Thus, for high-prejudice individuals, the RACIAL PREJUDICE node will be strongly activated, which will excite the HOSTILE BLACK and NEUTRAL WHITE stereotype nodes. For low-prejudice individuals, however, the RACIAL PREJUDICE node will be weakly activated, and in turn, the stereotype nodes will be considerably less active (since low-prejudice individuals are less likely to activate stereotypes; Lepore & Brown, 1997; Wittenbrink et al., 1997). The stereotype nodes are excited by the RACIAL PREJUDICE node (and also excited by category nodes consistent with them and inhibited by category nodes inconsistent with them).

Let us consider how the system would categorize the race of racially-ambiguous angry and happy faces, both for high-prejudice and low-prejudice individuals. To simulate a race categorization task, we set higher-level input into the RACE TASK DEMAND node at .9 and input into the EMOTION TASK DEMAND node at .1. We ran 4 race-categorization simulations: a high-prejudice individual categorizing a racially-ambiguous angry face and happy face, and a low-prejudice individual categorizing a racially-ambiguous angry face and happy face. We ran each simulation 100 times. Each time, we set visual input at .5 for the WHITE

CUES node and at .5 for the BLACK CUES node, thus making the faces perfectly ambiguous with respect to bottom-up racial cues. For angry faces, we set visual input at .9 for the ANGRY CUES node and at .1 for the HAPPY CUES node, and vice-versa for happy faces. For high-prejudice individuals, we set higher-level input at 1 for the RACIAL PREJUDICE node. This is a simplistic simulation of the activation of complex memory and affect structures that mediate an individual's racial prejudice. In contrast, for low-prejudice individuals, we set higher-level input at 0 for this node. After 150 iterations, we selected the race-category node with the highest activation as the network's categorization response.

Figure 6 shows the proportion of times the network categorized the face as Black, separately for a high-prejudice and low-prejudice individual, and separately for angry and happy faces. For happy faces, the network was 50% likely to categorize the face as Black for a low-prejudice individual and 48% for a high-prejudice individual. Thus, happy race-ambiguous faces appeared to be categorized as Black or White due to random noise (50% chance), and uninfluenced by level of prejudice. In contrast, the network was biased toward categorizing angry race-ambiguous faces as Black, with a low-prejudice individual having a greater than chance likelihood of Black categorization (69%), and the likelihood of Black categorization was even stronger for a high-prejudicial individual (88%). Thus, the network appeared to use the emotion category to disambiguate a face's race category, and this was exacerbated with a high level of racial
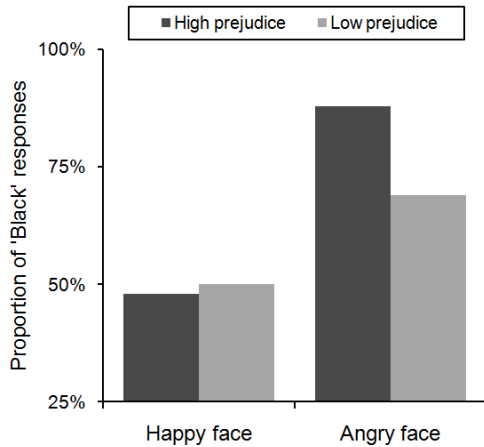
Figure 6. The proportion of times that the network categorized a racially-ambiguous happy face and racially-ambiguous angry face as Black (i.e., when the BLACK category node won the competition), both for a high prejudice individual (darker gray bars) and low prejudice individual (lighter gray bars).

prejudice (thus more readily activating the "Black is hostile" stereotype). How was this prejudice-mediated race-emotion interaction accomplished?

For a high-prejudice individual, presentation of a racially-ambiguous angry face sets a process into motion where ambiguous racial cues push the WHITE and BLACK category nodes above their resting levels, leading them to compete with one another. At the same time, the highly activated ANGRY CUES node strongly excites the ANGRY category node, but this is simultaneously inhibited by the RACE TASK DEMAND node (because this is a race categorization task). Activation of the BLACK category excites the HOSTILE BLACK stereotype node, whereas the WHITE category inhibits it. Similarly, activation of the WHITE category excites the NEUTRAL WHITE stereotype node, whereas the BLACK category inhibits it. Strong activation of the ANGRY category node also excites the HOSTILE BLACK stereotype, leading the HOSTILE BLACK stereotype to become more active than the NEUTRAL WHITE stereotype. Activation of the HOSTILE BLACK stereotype, in turn, feeds back excitation to the BLACK category and inhibition to the WHITE category. Moreover, because of the strongly activated RACIAL PREJUDICE higher-order node (because this is a high-prejudice individual), which has excitatory connections with the stereotype nodes, the stereotype nodes are already primed to be quite active. Stronger activation of the HOSTILE BLACK stereotype then feeds back activation to the WHITE and BLACK category nodes, causing the BLACK category to become more active and the WHITE category to become suppressed. In such a way, stereotypes exerted a top-down effect on race categorization through interactions with the anger category.

For a racially-ambiguous happy face, however, the influence of emotion category on race categorization was not obtained because the happy category is not involved in stereotypic associations with race category. As seen in Figure 6, for a happy face, the proportion of Black categorizations appeared generally the same for low- and high-prejudice individuals (at chance: 50%). Finally, for low-prejudice individuals, the effect of the ANGER category node exciting the HOSTILE BLACK stereotype, which in turn caused the race category nodes to diverge in activation (leading BLACK to win), was not as strong as it was for high-prejudice individuals. This is because the higher-order RACIAL PREJUDICE node primed activation of the stereotype nodes in high-prejudice individuals. Interestingly, Hugenberg and Bodenhausen (2003) also showed that the converse interactive effect holds as well. In categorizing facial emotion, racial prejudice exerts an analogous top-down influence, causing race category to interact with emotion category. High levels of racial prejudice lead White perceivers to activate the anger category more strongly for emotionally-ambiguous Black faces than White faces, whereas this is not as readily seen in perceivers with low levels of prejudice.

Phenomenon 3: Category Interactions due to Overlapping Stereotype Content

The interactive nature of category and stereotype activation suggests that many categorizations may interact due to stereotype contents that happen to overlap. For instance, particular social categories in one dimension (e.g., race) may facilitate and inhibit the activation of categories in another dimension (e.g., sex) due to shared activations in the stereotype level. Stereotypes associated with the sex category, male, include aggressive, dominant, athletic, and competitive, and these are also associated with the race category, Black. Similarly, stereotypes of shy, family-oriented, and soft-spoken apply not only to the sex category, female, but also to the race category, Asian (Bem, 1974; Devine & Elliot, 1995; Ho & Jackson, 2001). Thus, there is some overlap in the stereotypes belonging to the Black and male categories and in the stereotypes belonging to the Asian and female categories.

What would our model predict regarding this overlap? It would predict that category activation along one dimension (e.g., sex) would be constrained by feedback from stereotype activations triggered by the other dimension (e.g., race). Sex categorization, for example, could be potentially constrained by race-triggered stereotype activations. Because the stereotypes of Black and male categories happen to partially overlap, Black men would be categorized more efficiently relative to White and Asian men. This overlap is represented in our previous instantiation of the model (Figure 2), as AGGRESSIVE happens to be positively linked and DOCILE happens to be negatively linked with both BLACK and MALE categories. This overlap would lead the race-triggered excitation of AGGRESSIVE and race-triggered inhibition of DOCILE to feed back excitation to the MALE category and inhibition to the FEMALE category. This would facilitate a male categorization or, in cases of sex-ambiguous targets, bias categorizations toward male (rather than female). A similar effect would occur with the ASIAN and FEMALE categories, where race-triggered excitation of DOCILE and race-triggered inhibition of AGGRESSIVE would come to facilitate a female categorization or bias categorizations toward female. Thus, a dynamic interactive model predicts that incidental overlap in stereotype contents could powerfully shape the perception of another category dimension.

To demonstrate how the feedback from stereotype activation could disambiguate categorization of an alternate dimension, we ran a simulation of sex categorization using our earlier instantiation of the model (Figure 2). As done previously to simulate a sex-categorization task context, we set higher-level input at .9 for the SEX TASK DEMAND node and at .1 activation for the RACE TASK DEMAND node. We ran 3 simulations, one for each race: a sex-ambiguous Black face, a sex-ambiguous White face, and a sex-ambiguous Asian face. For each, we set visual input at .5 for both the MALE CUES and FEMALE CUES nodes (thus making sex-specifying cues completely ambiguous). We set visual input

at .95 for the cue node consistent with the face's race, and at .025 for the cue nodes corresponding with the other two races. Thus, for a sex-ambiguous Black face, we set visual input at .95 for the BLACK CUES node, at .025 for the WHITE CUES and ASIAN CUES nodes, and at .5 for the MALE CUES and FEMALE CUES nodes. We ran each of the 3 simulations 100 times. After 150 iterations we selected the network's sex-category response (male or female) based on whichever node had the highest activation. Figure S1 in the supplementary materials shows the proportion of female responses for each race.

When a sex-ambiguous face was Black, the network was biased towards male categorization, with a 26% likelihood to categorize it as female. When White, random noise seemed to be driving the sex-category competition one way or the other, with a 52% likelihood (random chance: 50%) of female categorization. When Asian, however, the network was biased towards female categorization, with a 75% likelihood of female categorization. Thus, a dynamic interactive model predicts that perceivers would be biased to perceive sex-ambiguous Black faces as men and, conversely, to perceive sex-ambiguous Asian faces as women. This is because the presumably task-irrelevant race category placed excitatory and inhibitory pressures on stereotype nodes which were incidentally shared with sex categories. Thus, the activation of stereotypes from presumably task-irrelevant categories (e.g., race) can powerfully shape the activation of other social categories (e.g., sex). Initial evidence for these sex–race interactive effects, due to incidental stereotype overlap, were recently reported (Goff, Thomas, & Jackson, 2008; Johnson, 2009; Johnson, Freeman, & Pauker, invited revision).

Bottom-up Interactivity in Social Categorization

A dynamic interactive model of person construal permits not just top-down interactions in social categorization, but also bottom-up ones as well. Above, we described how social categories may interact with one another through top-down processes. But such interactions may also be mediated at lower levels in the system as well. For instance, different social categories may interact because they are confounded directly in perceptual cues themselves. Indeed, the face is an extremely complex stimulus that affords many opportunities for bottom-up interactions. The mere fact that so many social categories (e.g., sex, race, age, emotion) are registered through the single percept of a face makes it highly unlikely that each set of category-specifying features is independent. Black individuals have considerably darker skin than White individuals, but also men have darker skin than women. Thus, skin tone, while strongly utilized for discriminating race, is also utilized for discriminating sex (Hill, Bruce, & Akamatsu, 1995). In all likelihood, multiple social categories share a great deal of the face's visual real estate. If correct, the variation in facial features specifying one category (e.g., sex) will partially overlap with the
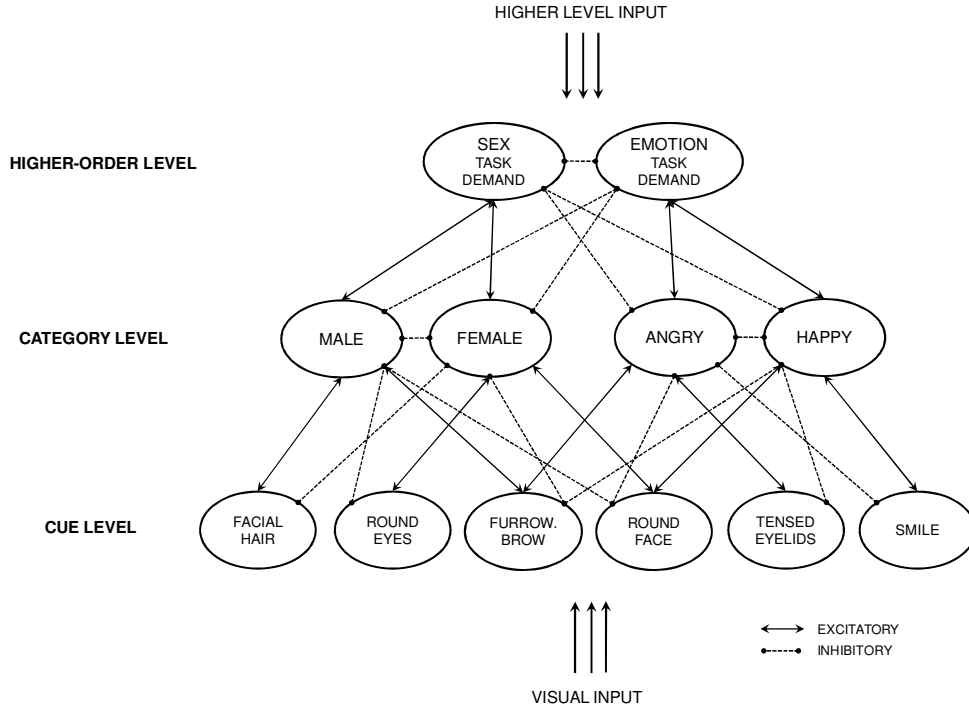
Figure 7. An instantiation of the model, used for simulations to account for Phenomenon 4.

variation in features specifying another category (e.g., emotion). Certain social categories may therefore be directly confounded due to bottom-up featural overlap.

For instance, the facial features specifying anger appear to overlap with the features specifying maturity, whereas the features specifying fear appear to overlap with the features specifying babyishness (Marsh, Adams, & Kleck, 2005). Becker, Kenrick, Neuberg, Blackwell, and Dylan (2007) made a compelling case for the confounded nature of sex and emotion categories from shared bottom-up perceptual cues. In a series of studies, they found that categorizations of sex and emotion were facilitated for faces of happy women and angry men, relative to happy men and angry women. Further, studies using faces displaying neutral emotion provided evidence for direct overlap in male-specifying cues and angry expressions, as well as overlap in female-specifying cues and happy expressions (see also Hess, Adams, Grammer, & Kleck, 2009; Oosterhof & Todorov, 2009). These studies suggest that a portion of the cues that make a face more masculine are the same cues that make a face angrier. Similarly, a portion of the cues that make a face more feminine are the same cues that make a face happier.

For instance, anger displays involve the center of the brow drawn downward, a compression of the mouth, and flared nostrils. These cues also distinguish sex categories. Men have larger brows which may cause them to appear drawn downward. They also have a more defined jaw and thinner lips, which may make the mouth to appear more

compressed, and they have larger noses, which may lead to the appearance of flared nostrils. A similar overlap exists for happy displays and the female face (Becker et al., 2007). For instance, women have rounder faces than men, and the appearance of roundness increases when displaying happiness (i.e., a smile draws out the width of the face). Previous studies suggest that it is this direct, physical overlap in the cues signaling maleness and anger and in the cues signaling femaleness and happiness that leads to more efficient perceptions of angry men and happy women (relative to happy men and angry women).

Phenomenon 4: Facial emotion shapes sex categorization through shared bottom-up cues

To account for these bottom-up interactive effects, we developed another instantiation (Figure 7) of our general model (Figure 1). Connection weights are provided in Table S3 of the supplementary materials. Differing from previous instantiations, here nodes in the cue level represent a single perceptual cue (e.g., defined jaw, smile). We did not use the stereotype level for this instantiation.

The mechanism underlying the bottom-up sex–emotion interaction is modeled in the cue level. Note that one cue node, FACIAL HAIR, has an excitatory connection with MALE and inhibitory connection with FEMALE, whereas another cue node, ROUND EYES, has an excitatory connection with FEMALE and inhibitory connection with MALE. Similarly, one cue node, TENSED EYELIDS, has an excitatory connection with ANGER and inhibitory

connection with HAPPY, and vice-versa for the cue node, SMILE. These four cue nodes represent the perceptual cues that independently relate to sex categories and independently relate to emotion categories. However, also note that one cue, FURROWED BROW, has an excitatory connection both with ANGER and with MALE (since a furrowed brow conveys both categories, described earlier). Similarly, another cue, ROUND FACE, has an excitatory connection both with HAPPY and with FEMALE (since a rounder face conveys both categories, described earlier). Thus, these two cue nodes represent the bottom-up overlap in the perceptual cues conveying sex and emotion. Note that the particular cues used in this simulation were chosen arbitrarily; they are merely intended to simulate the set of non-overlapping and overlapping perceptual cues that convey sex and emotion.

To simulate a sex categorization task, we set higher-level input at .9 for the SEX TASK DEMAND node and at .1 for the EMOTION TASK DEMAND node. We ran 4 simulations: sex categorization of an angry male, angry female, happy male, and happy female. For each simulation, we set visual input at 1 for the cue nodes that would be apparent on a given face stimulus. For instance, to simulate the presentation of an angry male face, visual input was set at 1 for the FACIAL HAIR node (independently cueing MALE category), TENSED EYELIDS node (independently cueing ANGER category), and FURROWED BROW node (cueing both ANGER and MALE categories). Or, to simulate the presentation of an angry female face, visual input was set at 1 for the ROUND EYES node (independently cueing FEMALE category), TENSED EYELIDS node (independently cueing ANGER category), FURROWED BROW node (cueing ANGER category, but also MALE category), and ROUND FACE node (cueing FEMALE category, but also HAPPY category). Each simulation was run 100 times, each time for 75 iterations, and an activation time course was averaged for each of the 4 simulations. Reaction times were calculated as the number of iterations it took for the sex-category node with highest activation (the network's response) to reach 90% of its final activation state. This number was then multiplied by a constant of 12 and added to a constant of 480 to approximate the human reaction time data (ms) of Becker et al. (2007).

Figure S2 in the supplementary materials shows the averaged activation time courses for the 4 simulations. When a male face was angry, the MALE category's activation grew more quickly and stabilized on a stronger state, relative to when a male face was happy. Conversely, however, when a female face was angry, the FEMALE category's activation grew more slowly and stabilized on a weaker state, relative to when a female face was happy. This sex–emotion interaction is reflected in the reaction time data, shown in Figure S3 of the supplementary materials. Categorization of angry men and happy women was facilitated, relative to categorization of angry women and happy men. This is the pattern of results observed in human

perceivers (Becker et al., 2007). Thus, categorizing one dimension (e.g., sex) is shaped by direct bottom-up overlap with the perceptual features supporting another dimension (e.g., emotion). This highlights bottom-up interactivity in social categorization and shows how it is naturally accounted by a dynamic interactive model of person construal.

Contextual and Cross-Modal Interactivity in Social Categorization

One of the most remarkable features of perceiving other people, as compared with everyday objects, is that perceptions of people are frequently grounded in multiple sensory modalities and embedded in a rich set of contexts. The human voice, for example, always contextualizes the human face, continuously over time. The body's motion, for instance, contextualizes the perception of its shape. A growing number of studies have shown that these prevalent contextual and cross-modal cues powerfully constrain the perception of the social percepts under the focus of perceivers' attention.

Hair, for instance, is a cue that may appear stereotypically Black or Hispanic. In a series of studies, racially ambiguous faces were readily disambiguated by their hair, with Black-like hair biasing categorizations toward Black and Hispanic-like hair biasing categorizations toward Hispanic (MacLin & Malpass, 2001). Thus, an identical face was perceived as Black or Hispanic depending on the hair cue that contextualized it. Other cues that contextualize the face, such as cues of the body, also constrain the face's perception. For instance, perceivers' categorization of a face's emotion slows down when the face is coupled with incongruent emotional body cues (Meeren, van Heijnsbergen, & de Gelder, 2005). Aviezer et al. (2008) presented participants with identical faces that were embedded in different body contexts that suggested particular emotions. Perceptions of identical facial expressions were strikingly influenced by contextualizing body cues. Thus, visual contexts surrounding a face, such as emotional body cues, powerfully bias perceptions of facial emotion.

Emotional body cues—whether the body is moving angrily or sadly—heavily bias the perception of the body's sex. Point-light displays depicting angry body motions are more likely to be judged as men and those depicting sad body motions more likely to be judged as women (Johnson et al., 2010). One likely reason for this is that emotion expression is sex-stereotyped, such that men are stereotyped as angry and women stereotyped as sad.

The power of a social percept's context is not limited to visual cues. Cues from other sensory modalities that contextualize the face can also alter its perception. For instance, incongruence between facial and vocal cues (e.g., a slightly feminine male's voice with a male face; a happy voice with a sad face) alters perceptions of the face and

induces longer face-categorization latencies (Campanella & Belin, 2007; Freeman & Ambady, in press). Cross-modal cues originating even in the olfactory system appear to interact with the processing of visual social percepts. The smelling of sex-specific hormones, for instance, biases the perception of a face's sex-category. Perceivers exposed to an androgen (a male-specifying hormone) required less masculine features to perceive a face as male, whereas perceivers exposed to estrogen (a female-specifying hormone) required more masculine features (Kovács et al., 2004). Below we focus on the interactivity between the face and voice in person construal.

**Phenomenon 5: Continuous Face–Voice Interactivity in Social Categorization**

Visual processing of the face and auditory processing of the voice robustly interact to perceive others, specifically in perceiving identity and emotion. For instance, when a face appears sad but is accompanied by a voice that sounds happy, perceivers consistently report seeing the face as more happy than it really is. This remains true even when participants are instructed to disregard the voice (de Gelder & Vroomen, 2000). Furthermore, congruency between vocal and facial features tends to make perceptions of another's emotions more accurate and efficient (for review, Campanella & Belin, 2007). Recently, face–voice interactions have also been explored in the context of sex categorization (Masuda, Tsujii, & Watanabe, 2005; Smith, Grabowecky, & Suzuki, 2007).

We investigated the temporal dynamics through which voice processing interacts with face processing in sex categorization (Freeman & Ambady, in press). Participants categorized slightly ambiguous male and female faces by sex while simultaneously presented with a sex-typical voice (e.g., masculinized male voice for a male face) or sex-atypical voice (e.g., feminized male voice for a male face). We tracked their computer mouse trajectories en route to indicating a "male" or "female" response on the screen.
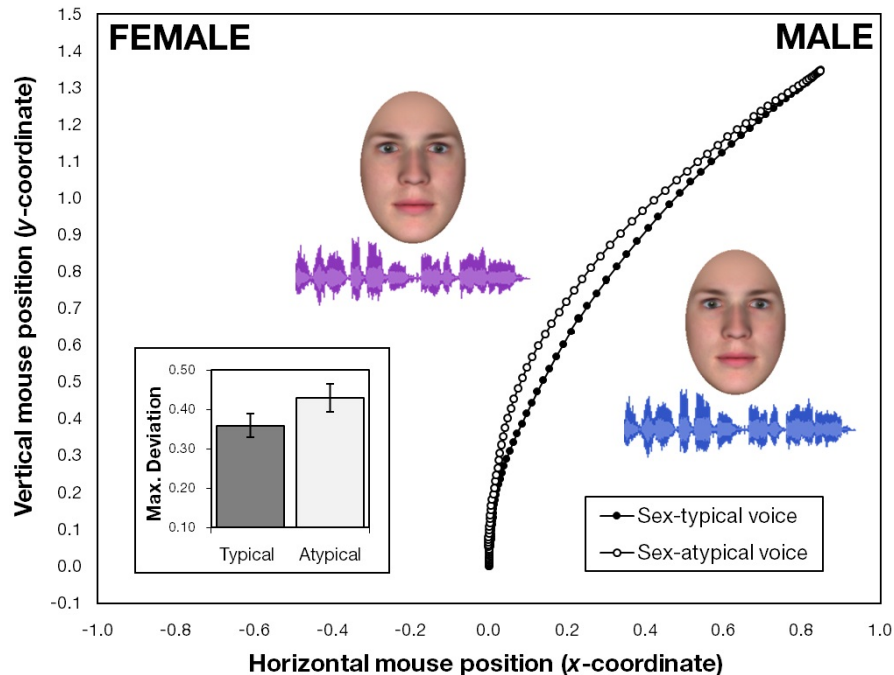
Figure 8. In a series of studies, Freeman and Ambady (in press) found that when categorizing a face's sex, the simultaneous processing of a sex-atypical voice led participants' computer mouse trajectories to be continuously attracted to the opposite sex-category response before settling into the response consistent with the face's correct sex. Mean mouse trajectories from this study are depicted (aggregated across male and female targets). In this figure, trajectories for all targets were remapped rightward, with the opposite sex-category on the left and the sex-category consistent with the face's sex on the right. A sample male face stimulus is displayed (all male and female face stimuli were somewhat sex-ambiguous). A voice stimulus typical for the face's sex (masculine) is shown on the right (audio waveform depicted in blue), next to the mean trajectory for sex-typical trials. Its atypical (feminine) counterpart is shown on the left, next to the mean trajectory for sex-atypical trials (audio waveform depicted in purple). During an actual trial, a single face was centered at the bottom of the screen while the voice stimulus played. The bar graph shows trajectories' maximum deviation toward the opposite sex-category from a direct line between trajectories' start and end points, separately for sex-typical and sex-atypical trials (error bars denote standard error of the mean).
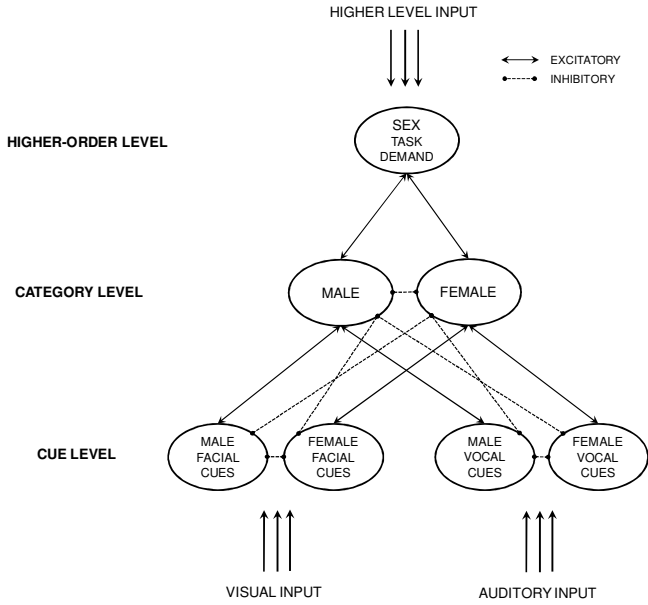
Figure 9. An instantiation of the model, used for simulations to account for Phenomenon 5.

When categorizing a face's sex, the simultaneous processing of a sex-atypical voice led the hand to travel closer to the opposite sex-category continuously across construal (Figure 8). Thus, even when perceivers correctly categorized the face's sex, auditory processing of sex-specifying vocal cues exerted a temporally dynamic influence on the face-based categorization. Specifically, the simultaneous processing of sex-specifying facial and vocal cues triggered partially-active representations of both sex categories (male and female) that simultaneously competed over time to settle into ultimate categorizations.

To account for this temporally dynamic face–voice interactivity in sex categorization, we developed another instantiation (Figure 9) of our general model (Figure 1). Connection weights are provided in Table S4 of the supplementary materials. In this instantiation, the cue level receives input from both visual processing and auditory processing, with nodes for MALE FACIAL CUES, FEMALE FACIAL CUES, as well as MALE VOCAL CUES and FEMALE VOCAL CUES.

To simulate the presentation of a slightly ambiguous male face, we set visual input at .55 for MALE FACIAL CUES and at .45 for FEMALE FACIAL CUES. To simulate the simultaneous presentation of a sex-typical voice, we set auditory input at .95 for MALE VOCAL CUES and at .05 for FEMALE VOCAL CUES. We also set higher-level input at .9 for the SEX TASK DEMAND node to simulate a strong attentional state on targets' sex required by the task. We ran this simulation 100 times, each time over 75 iterations, and plotted the averaged level of activation of the category nodes over time (Figure 10). The slightly ambiguous activation of facial cues nodes fed forward activation onto

the MALE and FEMALE category nodes. Simultaneously, the activation of the vocal cues nodes also fed forward activation onto the category nodes. In doing so, the simultaneous processing of vocal cues placed an immediate constraint on the face-triggered activation of sex categories. This permitted ongoing updates from voice processing to immediately interact with ongoing updates from face processing, continuously over time. The strong activation of MALE VOCAL CUES was therefore immediately brought to bear on resolving the category competition induced by ambiguous facial input. Strong excitation of the MALE category and inhibition of the FEMALE category, due primarily to the unambiguous vocal cues nodes, led the system to rapidly converge on a stable state involving strong activation of MALE category, with FEMALE category pushed below resting level.

When the voice is more atypical, however, the face-triggered category competition does not resolve so quickly. To simulate the presentation of a slightly ambiguous male face coupled with a sex-atypical voice, we kept the input activation the same except, this time, we set input into the MALE VOCAL CUES node at .6 and input into the FEMALE VOCAL CUES at .4. We ran the simulation 100 times, each time over 75 iterations, and plotted the averaged level of activation of the category nodes over time (Figure 10). The slightly ambiguous activation of facial cues nodes and slightly ambiguous activation of vocal cues nodes simultaneously fed forward activation onto the MALE and FEMALE category nodes. This induced a strong competition between the category nodes. Although the system eventually resolved the competition by arriving at a stable state involving strong activation of MALE and weak activation of FEMALE (i.e., a male categorization), the FEMALE category was partially-active in parallel strongly throughout the
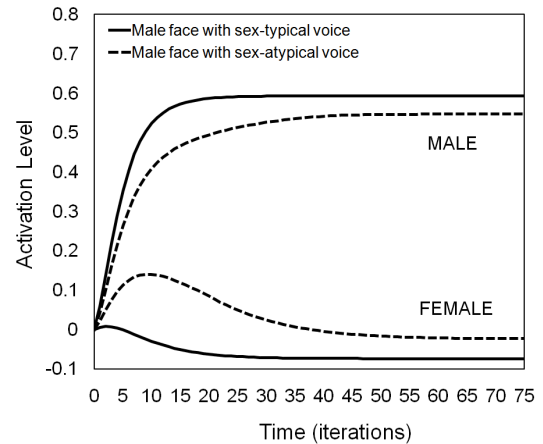
Figure 10. The activation level of the MALE and FEMALE category nodes as a function of time (iterations) following the presentation of a sex-typical male face (solid lines) and sex-atypical male face (dashed lines).

process. This partial activation of the FEMALE category was considerably stronger when the voice was sex-atypical rather than sex-typical (Figure 10).

This pattern is precisely what is observed in the laboratory (Freeman & Ambady, in press). The stronger partial activation of the FEMALE category, which continuously competes with the MALE category, is clearly seen in the human mouse-tracking data of Figure 8. When sex-categorizing a male face, the simultaneous processing of a sex-atypical voice led participants' hands to be continuously attracted toward the "female" response before ultimately arriving at the "male" response. This reflects a stronger partially-active representation of the female category (induced by voice processing) that simultaneously competed over time with the male category during face-based categorization. Thus, in sex categorization, the model predicts (as experimental data show) that voice processing interacts with face processing through temporally dynamic competition. As such, the simultaneous processing of facial and vocal cues place parallel constraints on sex categorization (which are dynamically satisfied over time), permitting the ongoing processing of vocal cues to continuously interact with the ongoing processing of facial cues. In short, our model naturally accounts for continuous cross-modal interactivity in person construal.

Summary

We propose a dynamic interactive theory of person construal. It argues that the perception of other people is accomplished by a dynamical system in which lower-level sensory perception and higher-order social cognition continuously coordinate across multiple interactive levels of processing to give rise to stable person construals. We described a recurrent connectionist model of this system that accounted for a wide range of phenomena, including partial parallel activation and dynamic competition in categorization and stereotyping (Phenomenon 1); top-down influences of high-level cognitive states and stereotype activations on categorization (Phenomena 2 and 3); bottom-up category interactions due to shared perceptual features (Phenomenon 4); and continuous cross-modal interaction in categorization (Phenomenon 5).

In a dynamic interactive model, perceptions of other people gradually emerge through ongoing cycles of interaction between social categories, stereotypes, high-level cognitive states, and the low-level processing of facial, vocal, and bodily cues. Internal representations of categories and stereotypes are dynamically and probabilistically reconstructed, rather than behaving like static, symbol-like structures that wait around inertly until discretely accessed. The real-time evolution of these probabilistic representations is in continuous interaction with other activations across the system, both dynamically influenced by these other activations and a source of influence over them. The entire system's prior history, its visual inputs

(e.g., facial cues), auditory inputs (e.g., vocal cues), and higher-level inputs (e.g., task demands, prejudice, motivation), its internal constraints, and some random noise jointly determine the construal of other people.

Comparison with Extant Models

Extant social psychological models have described how perceivers form high-level impressions of other people, whether they utilize category-based or individual-based information, and how knowledge about individuals and groups is learned, stored, and accessed (Bodenhausen & Macrae, 1998; Brewer, 1988; Chaiken & Trope, 1999; Fiske et al., 2002; Fiske & Neuberg, 1990; Higgins, 1996; Kunda & Thagard, 1996; Read & Miller, 1998b; Smith & DeCoster, 1998; Srull & Wyer, 1989; van Overwalle & Labiouse, 2004). Models in the cognitive face-processing literature, on the other hand, have described the visual and perceptual mechanisms that permit face recognition (Bruce & Young, 1986; Burton et al., 1990; Valentin, Abdi, O'Toole, & Cottrell, 1994). Our dynamic interactive model helps unify these two literatures by describing how the lower-level perceptual processing modeled in the cognitive literature works together with the higher-order social cognitive processes modeled in the social literature to give rise to person construal.

Social psychological models have tended to use categorization as a starting point, with relatively little focus on the perceptual processing that gives rise to it. Thus, in Fiske and Neuberg's (1990) influential model of impression formation it is argued that the utilization of stereotypes, which is derived from a dominant categorization, is prioritized over more individual-based information in forming impressions, unless the perceiver is motivated to move further and individuate the target. This model, like Brewer's (1988) and Kunda and Thagard's (1996) models of impression formation, provide comprehensive accounts of how top-down processes, such as stereotypic expectations, motivation, and attention, interact with the bottom-up process of learning explicit individuated characteristics about a target. In these models, therefore, a target's category memberships are given, and their influence on subsequent interpersonal phenomena are richly described (e.g., impressions, behavior). This is also the case for other models of person perception, such as Bodenhausen and Macrae's (1998) stereotype activation and inhibition model. As such, categorization (and corresponding stereotype activation) is the initial input into these models. The focus of these models is not to explain the categorization process itself; it is to explain the higher-order social cognitive processing that comes after.

Our framework builds on these important models by fleshing out the initial category and stereotype activation process and explaining how this process is dynamically driven by both bottom-up sensory information as well as high-level top-down factors. Notably, this expands on extant

models by explaining how initial category and stereotype activation may be influenced, sometimes considerably, by top-down factors. Although models of person perception have always emphasized the role of top-down factors (e.g., expectations, motivation, and attention), these factors have not been readily acknowledged to seep down into lower levels of processing, into the initial category and stereotype activation process itself. For example, in our model such top-down factors had an important role in Phenomenon 2, where racial stereotypes, more or less activated by prejudice, caused a face's emotional expression to alter the perception of race. Our modeling of the reach of top-down influences into even lower levels of person perception, such as basic category activation, thus builds on extant models that have generally described only the reach of top-down influences into higher levels of processing.

Beyond the importance of accounting for how perceptual processing brings about social cognitive phenomena in general, our modeling of perceptual processing is also important because it can bear a variety of downstream effects. For example, within-category facial or vocal variation affects the dynamic competition inherent to categorization (Freeman & Ambady, in press; Freeman et al., 2008; Freeman et al., 2010b), in turn affecting the eventual stable category representations that perceivers settle into (Locke et al., 2005). Thus, more prototypically masculine facial or vocal features (relative to less), for instance, affects the competition between male and female categories, which results in a stronger stable representation of the male category and weaker stable representation of the female category (see Figure 4A). This can bear a variety of downstream effects, shaping trait attributions (Blair et al., 2005; Blair et al., 2004b; Blair et al., 2002; Ko, Judd, & Blair, 2006; Maddox & Gray, 2002) as well as behavior (Blair et al., 2004a; Johnson et al., 2006). Moreover, as shown in Phenomenon 4, categorization of a focal category membership may be shaped by other memberships because the perceptual cues supporting those memberships are directly confounded (e.g., angry men and happy women; Becker et al., 2007). Thus, our framework builds on extant models by shedding new insights into the relationship between the higher-order processes extant models have described and the lower-level perceptual processing that has received less attention.

Kunda and Thagard's (1996) model of impression formation and Read and Miller's (1998b) Social Dynamics model provide important precedents to the present work. These connectionist models proposed that parallel-constraint-satisfaction principles guide impression formation and social reasoning, and here we proposed that such principles guide person construal. Thus, in Kunda and Thagard's model, categories (e.g., Black) and stereotypes (e.g., aggressive) have an equal priority with individuating information (e.g, pushed someone) in driving high-level impressions, and this information is simultaneously integrated into a coherent impression through a process of constraint satisfaction. Read and Miller's model additionally included detectors for perceptual features and considered these detectors' influence on how a target is identified, with a target's identification thereafter influencing high-level social reasoning (understanding social scenarios and attributing traits). For example, Read and Miller considered how detection of a target's grey hair would lead to the identification that the target is old, which then guides how perceivers reason about a relevant social scenario (e.g., why someone would help carry groceries for the target). At the heart of their model is a constraint-satisfaction process that gives rise to high-level social reasoning, similar to the constraint-satisfaction process underlying Kunda and Thagard's model. Our model shares a kinship with these models and suggests that constraint-satisfaction processes may give rise to a variety of different person-perceptual phenomena, including person construal in our case, as well as impression formation (Kunda & Thagard, 1996) and social reasoning (Read & Miller, 1998b).

These models, however, like other connectionist models in person perception research (Read & Miller, 1993; Read et al., 1997; Smith & DeCoster, 1998; van Overwalle & Labiouse, 2004), did not aim to extensively deal with the perceptual processing that drives the category and stereotype activation process or to examine how specific perceptual interpretations emerge. However, they did make important points about the role of perceptual processing in instigating the higher-order phenomena they were interested in, and the ability for perceptual processing to be potentially influenced by these phenomena (e.g., Read & Miller, 1998b). Our model builds on these prior models by examining how perceptual interpretations emerge and by comprehensively modeling the role of perceptual processing and its dynamic influence by higher-order processes. Our model also extends these models by directly examining the temporal dynamics of construing others. For instance, although extant connectionist models, like our model, assume that processing is continuous and internal representations are dynamic, temporal dynamics and the time-course of person perception were not of primary interest. Instead, these models generally focused on the outcomes of person perception. Therefore, a network's ultimate stable states were used to explain person perception outcomes, with little modeling and simulation of the extended dynamic processing that culminates in those outcomes. Our model builds on these models by explicitly describing the temporal dynamics of person construal and by accounting for these dynamics in several simulations. This is important because a central claim of our theory is that perceptions of other people are the end-result of a time-dependent process of simultaneously and partially active representations continuously interacting over time. Thus, it is important that our model explicitly describe these dynamics and fit them to human data. The fleshing out of temporal dynamics is a novel aspect of our theory and

model, and these add to extant accounts of person perception.

Implications

Our theory and model have several implications for present understandings of person construal, which we discuss here.

Re-thinking the "Multiple Category Problem"

Individuals naturally vary along any number of category dimensions (e.g., sex, race, age). Extant models have often emphasized that one category and the stereotypes tied to that category come to dominate the processing landscape, whereas others are actively suppressed, making the perceiver's job easier and thereby solving the "multiple category problem" (e.g., Bodenhausen & Macrae, 1998; Macrae, Bodenhausen, & Milne, 1995; Sinclair & Kunda, 1999).

In a dynamic interactive model, the selection of one category and winnowing of other categories is accomplished by top-down pressure from higher-order nodes. For instance, the task demands of sex categorization, expressed by higher-order nodes, exerts excitatory pressure on to-be-attended categories (male, female) and inhibitory pressure on task-irrelevant categories (e.g., Black, White, Asian). However, our model introduces several nuances to an understanding of how the person construal system comes to arrive at focal categorizations of others.

Our model assumes that these top-down task-demand pressures exert their differential influence on categories dynamically over time. Thus, although for the purposes of sex categorization an applicable sex category rises in activation (thus becoming focally attended) whereas an applicable race category falls (thus becoming ignored), this pattern of excitation and inhibition is not instantaneous. Rather, higher-order task-demand nodes gradually exert excitatory pressure on certain categories while exerting inhibitory pressure on others. Thus, while these pressures are still at work our model predicts that multiple applicable category memberships (e.g., sex, race) are actually flexibly active in parallel. This places our model in line with neural dynamic models of visual attention (Desimone & Duncan, 1995), which assume a similar parallel activation of multiple representations.

Because multiple applicable category memberships (e.g., Black, angry) may be active in parallel while the system works toward stabilizing on a focal category (e.g., Black), non-focal categories also have an influence over perception. This is because their partial parallel activation can powerfully affect the system's trajectory and the stable states it achieves. A clear demonstration of this, for example, is found in our modeling of the prejudice-mediated race-emotion interactive effects (Phenomenon 2). Due to the context of a race categorization task, higher-order nodes exerted excitatory pressure on race-category nodes and

inhibitory pressure on emotion-category nodes. While these top-down task-demand pressures were at work, for a great deal of processing time emotion categories were still partially active in parallel. This emotion-category activation had a powerful effect on the trajectory of the system and on race categorization in particular. The competition between race categories, which was initiated by a race-ambiguous face (and thus initially equibiased with respect to bottom-up visual input) was powerfully swayed one way (White) or the other (Black) based on the partial parallel activation of presumably task-irrelevant emotion categories. Specifically, when a race-ambiguous target was angry, the partial activation of the ANGRY category biased the race-category competition toward a Black categorization, and this was especially the case for high-prejudice individuals. Thus, non-focal, presumably task-irrelevant categories (e.g., emotion in a race categorization task) can bear powerful influences on focal person construals.

Our model also implies that, in the absence of strong top-down factors that require all but one category to be inhibited (e.g., task demands, goals), the person construal system could settle into stable states that are quite flexible. For instance, without higher-order nodes exerting inhibitory pressures on particular category nodes, the attractor states that the person construal system settles into could easily involve multiple categories (e.g., White, male) that are flexibly active in parallel. Indeed, the quality of having multiple person characteristics (e.g., lazy, friend, lives nearby) partially active in parallel is a central feature of the content-addressable memory modeled in connectionist networks of person memory (Smith, 1996, 2000; Smith & DeCoster, 1998). Just as multiple categories have often been shown to simultaneously constrain high-level impressions and social reasoning (Kunda & Thagard, 1996; Read & Miller, 1998b), a dynamic interactive theory proposes that they also simultaneously constrain lower-level person construals. Thus, the "multiple category problem" might best be characterized not so much as a "problem" that must be eliminated to keep cognitive efficiency (Allport, 1954), but rather as a reflection of the flexibility of interactive, parallel category representations.

The Dynamic Coextension of Category and Stereotype Activation

Recent research has found that variation in facial features may bear effects on stereotype activation that are independent of a target's category membership. For instance, the presence of Black-specifying cues on a person who is not Black (e.g., a White face) increases Black-related stereotypic attributions (Blair et al., 2005; Blair et al., 2002). These effects may thereafter influence behavior as well. For example, in court trials, targets with more Black-specifying features are punished more severely and more likely to be sentenced to death (Blair et al., 2004a; Johnson et al., 2006). Based on such findings, some accounts have argued that these independent feature-based effects on stereotype

activation are accomplished by a special feature-based processing route, where features become associated with stereotypes unmediated by any category representation at all (Blair et al., 2002; Livingston & Brewer, 2002). This direct *features → stereotypes* route is theorized to be separate from a more typical *categories → stereotypes* route.[5]

Our model agrees with these previous accounts that facial features can influence stereotype activation without a discrete categorization. However, because our model permits categorizations to be partially active in parallel, independent feature-based effects on stereotype activation could be mediated by the tentative, partially-active categorization of an alternate category. Specifically, our model suggests that independent feature-based effects on stereotyping are a product of the dynamic processing cascade inherent to the system. Cues of an alternate category (e.g., Black-specifying cues on a White target) trigger partially-active, competing category representations (e.g., "he's [tentatively] White" versus "he's [tentatively] Black"). Both category representations (e.g., White, Black) then immediately pass activation onward to their respective stereotypes before the competition in the category level has resolved and settled into just one alternative. This is reflected in Figure 4B, where feminine cues on a man's face triggered the partial and parallel activation of the FEMALE category, which continuously cascaded into the partial and parallel activation of the female-related stereotype, DOCILE, as was shown with human perceivers (Freeman & Ambady, 2009). Thus, the dynamic coextension of category and

---

[5] It should be noted that this also applies to face overgeneralization effects. Face overgeneralization occurs when adults whose facial characteristics resemble babies, the unfit, a particular emotion, or a familiar person are perceived as possessing the internal characteristics suggested by those cues (e.g., being babyish, being unfit, feeling a particular emotion, or being similar to another familiar person, respectively; Zebrowitz & Montepare, 2008). Conceptually, overgeneralization effects are akin to independent feature-based effects on stereotype activation. For instance, baby-specifying cues on an adult's face appear to partially trigger baby stereotypes independent of the fact that the target is not actually a baby. This is akin to Black-specifying cues on a White target triggering Black-related stereotypes independent of the fact that the target is White (e.g., Blair et al., 2002) or analogous effects with sex categories (Freeman & Ambady, 2009). Thus, face overgeneralization effects are similarly accounted in our model by the dynamically cascading activation inherent to the person construal system. Specifically, activation of baby-specifying cues on an adult's face, for example, would continuously cascade activation into the category level, triggering partially-active representations of both adult category and baby category. These, in turn, would continuously cascade activation into the stereotype level, triggering the partial parallel activation of baby stereotypes. This would thereafter bias high-level judgments and evaluations, driving the perceiver to infer the adult target is somewhat babyish (e.g., innocent, inexperienced). Thus, a dynamic interactive model can account for face overgeneralization effects, complementing connectionist networks that model them explicitly (Zebrowitz et al., 2003; Zebrowitz, Kikuchi, & Fellous, 2010).

stereotype activation permits independent feature-based effects on stereotype activation. As such, our model parsimoniously accounts for independent feature-based effects on stereotyping by one single route involving a dynamic processing cascade.[6]

A Rapidly Adaptive, Ecologically Valid Person Construal System

Like the present model, the ecological approach to social perception (McArthur & Baron, 1983) emphasized the need to study directly the stimulus information that avails perceivers with functionally significant characteristics about other people. It also emphasized the inherently dynamic and multimodal nature of social stimuli. Our dynamic interactive framework is in the best spirit of this approach and builds on it in several ways.

Our framework brings new and helpful ways of thinking about ecologically-valid person construal. Specifically, it assumes that the person construal system's processing is fully continuous and highly interactive, and that its representations are probabilistic, active in parallel, and changing over time. This is exactly the sort of system required for the ecologically-valid person perceiver—the kind of perceiver that must make sense of others in real-time, on-the-fly, and in a rapidly changing social environment. In real-world social encounters the sensory stimulation of another person is almost always in continuous flux (Gibson, 1979). The most obvious example might be the perception of a face's emotion, which continuously fluctuates over time. Rarely do perceivers encounter a static emotional expression. Rather, for just a few fleeting moments, another's face displays slight anger, which then rapidly transitions into some other expression. By the time perceivers are finished processing that subtle anger, however, there are already hundreds of milliseconds of new visual information that needs to be accrued and dealt with. In real-world person construal, therefore, another's face tends not to fit squarely into any one expression (e.g., angry), but is usually in some in-between state amidst one interpretable expression and the next, and rarely standing still.

For simplicity, in our instantiations of the model we supplied external input to the network discretely (at iteration 1). However, the model is flexible to support the more ecologically-valid situation in which external stimulation to the network dynamically changes across time based on changing cues in the social environment. As a face's

---

[6] Although our model can account for independent feature-based effects on stereotyping with a single route involving dynamically cascading activation, future work could attempt to empirically disentangle a one-route account (cues → categories → stereotypes) versus two-route account (categories → stereotypes and cues → stereotypes) using the model. By implementing direct connections between the cue nodes and stereotype nodes, researchers could determine whether experimental data better fit a one-route or two-route account in the future.

emotion, a body's subtle nonverbal behavior, or the ongoing stream of vocal cues fluctuate over time, the visual and auditory inputs into cue nodes would continually change across iterations accordingly. This would thereby continually change, iteration to iteration, the amount of excitatory and inhibitory pressures on category and stereotype nodes. As such, at any given moment while the system is trying to settle into one stable attractor state, new sensory information bombarding the system would already start changing the various attractor states to which the system will start gravitating (Spivey, 2007). This leaves little time for the system to actually rest in any given stable state, since by the time it starts to stabilize it is already being pushed out of its stability by new constraints (e.g., changes of a face's emotion, of the body's behavior, of the voice stream). Thus, the network we have outlined is a rapidly adaptive and dynamic person construal system. Its continually evolving states are able to be tightly yoked to the ongoing sensory stimulation of the social environment.

This adaptive, dynamic person construal system is potentially stimulated by continuous top-down input as well. For instance, ecologically-valid, moment-to-moment changes in one's goals or attentional states, among other top-down factors, would continually stimulate higher-order nodes, which thereafter continually change the amount of excitatory and inhibitory pressures on category and stereotype nodes. Thus, although for the sake of simplicity we modeled external inputs into the network as discrete occurrences, the system is inherently capable of supporting stimulation by a dynamically changing social environment as well as dynamically changing internal cognitive states.

An ecologically-valid person construal system also needs to permit ongoing perceptions of other people to guide action continuously over time. In social interaction, something apparent on individual A's face and gesturing elicits a reaction on individual B's face and gesturing, which then elicits a reaction on individual A's face and gesturing, and so on and so forth. Thus, there is no staccato series of static images and sounds that elicit particular reactions. Instead, ecologically-valid person construal would likely need to involve continuous millisecond-by-millisecond updates of facial, vocal, and bodily information, and these updates need to make their way onto the motor system immediately, not once the system has 100% finalized the processing of each transient image or sound in a social interaction. Indeed, recent neurophysiological evidence suggests that this dynamic person processing is a likely possibility. In a series of event-related potential studies, we demonstrated that the process of social categorization immediately shares its ongoing results with the motor cortex to guide action continuously over time (Freeman et al., in press). This is consistent with multi-cell recordings in nonhuman primates (Cisek & Kalaska, 2005, 2010). Thus, person construal is characterized by continuous perceptual–cognitive–motor dynamics, such that perceptual, cognitive, and motor processing are coextensive. Cognitive representations of a face's category memberships develop over hundreds of milliseconds while perceptual processing is ongoing, and these representations evolve alongside accruing perceptual evidence for category alternatives. Further, ongoing results of this social category processing are immediately cascaded into the motor cortex to guide relevant actions continuously over time. Thus, person construal is continuously coextensive with action. This is exactly the kind of processing required by the ecologically-valid person perceiver.

In short, we have described here a person construal system that is rapidly adaptive and dynamic. It is able to perceive others in an ecologically-valid, real-time social environment, while also able to coordinate with the motor system to act on ongoing perceptions.

New Predictions and Future Directions

Beyond our model's ability to explain a wide range of phenomena, it also gives rise to a number of new and distinctive predictions, which future work could directly examine. Below are a few examples of important predictions derived from the model that could serve as testable hypotheses in the future.

Category Interactions due to Incidental Stereotypic and Phenotypic Overlap

Our model makes the novel prediction that any incidental overlap in the stereotype or phenotype content of two category memberships would lead the system to throw those categories into interaction. As shown in Phenomenon 3, overlapping stereotype content between the male and Black categories (e.g., aggressive) and between the female and Asian categories (e.g., docile) created top-down pressure that gave rise to sex–race interactions. Or, as shown in Phenomenon 4, overlapping phenotype content between male and angry faces (e.g., furrowed brow) and between female and happy faces (e.g., roundness) created bottom-up pressure that gave rise to sex–emotion interactions.

However, any number of category interactions are possible and, in fact, quite likely. Many stereotypes are likely to be incidentally shared by multiple categories. In fact, the very existence of some categories may be predicated on the stereotypes of other categories, such as sexual orientation categories and sex-category stereotypes (Kite & Deaux, 1987), and this is evident in perceptual construals (Freeman, Johnson, Ambady, & Rule, 2010a; Johnson, Gill, Reichman, & Tassinary, 2007). Future work could empirically estimate the degree of stereotype overlap between categories using explicit or implicit measures, and implement the estimated overlap into the stereotype and category levels. A variety of category interactions could arise in network simulations, and these could then be experimentally tested in the laboratory.

Similarly, the perceptual cues contained in the face, voice, and body are likely to, by chance, partly covary between categories. Future work could empirically estimate the degree of phenotype overlap between categories and then implement this estimated overlap into the cue and category levels. For example, face-modeling software can derive precise estimates of hundreds of facial cues from a facial photograph (e.g., Blanz & Vetter, 1999). Thus, researchers could derive estimates of cue overlap using representative samples of faces for specific category memberships, and then implement these estimates into instantiations of the model. If category interactions arose in network simulations, these could then be experimentally investigated in the laboratory.

Category interactions could also potentially be driven by both top-down and bottom-up overlap at the same time. For example, not only do male and angry cues and female and happy cues overlap (Becker et al., 2007), but also men are stereotyped as angry and women are stereotyped as happy (Fabes & Martin, 1991). Simulations with our model are uniquely poised to assess the relative contribution of potentially coexistent top-down and bottom-up forces in driving category interactions. Such simulations could also be used to tease apart the time-courses of these two forces' influence on perceptions.

Temporally Dynamic Influence of Top-Down Factors

As shown with Phenomena 2 and 3, high-level cognitive states and stereotype activations may readily exert top-down pressure on categorization. Although in the case of substantially ambiguous targets our model predicts that such pressure will lead categorizations to be pushed entirely one way (e.g., White) or another (e.g., Black), an ultimate categorization outcome is unlikely to be altered in the case of less ambiguous targets. Nonetheless, even though in many instances such outcomes may not be altered wholesale, our model predicts that top-down factors will still impose a variety of dynamic biases across the categorization process. Further, these will often result in the triggering of alternate category memberships (e.g., male category for a female target) that are partially-active in parallel. Consider the following example.

An individual on a job interview is told she will meet with a high-profile business executive. Her expectations of the executive trigger a host of stereotypes (e.g., business-oriented, dominant, high-status), which are mostly associated with the MALE category. As she enters the office and takes her initial glance at the executive, who is a woman, visual processing of female-specifying cues will activate cue nodes, which thereafter place excitatory pressure on the FEMALE category. Simultaneously, the activated stereotypes will exert excitatory pressure on the MALE category. Eventually, competition between the two categories will lead the system to converge on an attractor state involving strong activation of FEMALE and weak activation of MALE, thereby achieving a female construal.

Nonetheless, the MALE category would be partially-active in parallel for hundreds of milliseconds while top-down pressure from stereotype activation (biasing the competition toward a male construal) continuously interacts with the target's bottom-up perceptual accrual (biasing the competition toward a female construal). Thus, even though perceivers' ultimate construal outcomes might not be affected by top-down factors, our model predicts that the construal process will be considerably influenced by such factors.

Such temporally dynamic influences of top-down factors on the construal process have yet to be tested in the laboratory. Prior work investigating top-down influences has focused on ultimate construal outcomes rather than the temporal dynamics culminating in those outcomes. The studies that have examined temporal dynamics (e.g., via mouse-tracking), however, have investigated only bottom-up effects (e.g., manipulations of cues) to shed light on the nature of person construal processing, with little mention of top-down effects. Our model makes the novel prediction that even for person construals that are ultimately "veridical," top-down factors will still exert any number of subtle, temporally dynamic biases across the course of construal. Future work could investigate this, including measuring the time-course of these top-down biases and confirming its correspondence with network simulations.

Downstream Consequences of "Hidden" Parallel Activations

Anderson (2002) argued for the importance of bridging psychological phenomena across multiple orders of temporal magnitude. Here we provided a model of person construal that fleshes out the process by which an ultimate perception crystallizes on the order of hundreds of milliseconds. But how do these relatively low-level, fine-grained dynamics relate to higher-order phenomena on the order of hundreds of seconds or hours, such as aspects of social interaction or other behavioral outcomes? There are likely many relationships to be uncovered. For example, our model predicts that several unforeseen category and stereotype representations may be simultaneously and partially active before perceivers arrive at an ultimate construal. Subtle bottom-up overlap with an alternate category (e.g., slight feminine facial features on a man) can lead to partial parallel activation of that alternate category (e.g., female). Or, as discussed above, high-level cognitive states or stereotypes can exert top-down influences on category-level processing, in turn triggering partially-active representations of other candidate categories. Our model therefore predicts that, for a great many of our construals of others, a variety of "hidden" category and stereotype activations may be partially triggered in parallel—activations that are not reflected in an ultimate perceptual outcome.

Such subtle activations triggered during real-time construal could likely give rise to a variety of unforeseen

downstream consequences. The lasting effects of category and corresponding stereotype activation on higher-order social phenomena—even the briefest of kinds (e.g., priming)—have long been documented. Activated stereotypes change how we think about others, judge, and remember them (Bodenhausen, 1988; Brewer, 1988; Devine, 1989; Fiske & Neuberg, 1990). They also activate related attitudes and behavioral tendencies, in turn changing how we feel about others and evaluate them (Fazio, Sanbonmatsu, Powell, & Kardes, 1986) and how we interact with others and treat them (Bargh, Chen, & Burrows, 1996; Chen & Bargh, 1999). Thus, future work could investigate how "hidden" parallel activations of alternate categories and stereotypes computed during the construal process, or other aspects of this real-time process, relate to important downstream phenomena. Moreover, such work could test how variation in the presence of these parallel activations relates to measures of individual differences (e.g., levels of prejudice or motivation) or other behavioral outcomes.

Future Advances to the Model

Future work could advance the model and simulations presented here in several ways. First, our simulations were limited to focusing on how sensory information and high-level cognitive states temporally conspire to shape category and stereotype activations. However, any given change in one node of the system will lead to changes in all other nodes, as the system works over time to maximally satisfy all of its constraints in parallel. Thus, the model is highly interactive and inherently bidirectional. It therefore assumes that, beyond high-level cognitive states shaping lower levels of processing, lower levels of processing also shape high-level cognitive states. As such, the model predicts that sensory information and category and stereotype activations should all lead to a variety of changes in high-level cognitive states. However, in the present work our focus was on category and stereotype activations as the dependent measures of interest. Future work could develop the model further by testing the reverse relationship, making high-level states the dependent measure of interest (e.g., motivation, prejudice, top-down attention, affect) and examining how these states are shaped by a rich interaction with lower levels of processing, as the model predicts.

The model could also be advanced by deriving network parameters empirically (see Footnote 3), and experimental studies could be used to refine and expand the model. For example, data could be collected for estimating the connection weights between category nodes and potentially hundreds of stereotype nodes (e.g., via explicit or implicit measures) and hundreds of cue nodes (e.g., via face-modeling algorithms), and all these nodes and their weighting could be implemented in future versions. This would bring the model closer to the empirical rigor and level of quantification common to connectionist models of speech perception (e.g., McClelland & Elman, 1986). Moreover, future work could opt to replace the cue level

with more sophisticated approaches to modeling the uptake of sensory information, such as a pixel-based image processor (e.g., Burton, Bruce, & Hancock, 1999). This would make fewer assumptions about the role of specific features and instead rely more on the emergent properties inherent in other people's sensory information. Together, such advances would allow the model to better reflect the real-world interrelatedness among cues, categories, stereotypes, and high-level states.

Conclusion

A new approach to the study of person perception is on the rise, as evidenced by the two recent volumes, *The Science of Social Vision* (Adams, Ambady, Nakayama, & Shimojo, 2010) and *The Social Psychology of Visual Perception* (Balcetis & Lassiter, 2010). Social psychologists are working alongside researchers in the cognitive, neural, and vision sciences to provide a unified and more complete understanding of person perception. In the present work, we sought to open up the temporally extended, real-time process of person construal. In this real-time process, person construal is dynamic and interactive, and the connection between the "sensory" and the "social" an intimate one. Both our theory and the model we have described here show that many person construal phenomena may be accounted for by a dynamical system that permits lower-level sensory perception and higher-order social cognition to continually collaborate across multiple interactive levels of processing. Low-level sensory information and high-level social factors fluidly work together to give rise to stable and integrated perceptions of other people. Probabilistic and parallel construals gradually emerge through the ongoing interaction between categories, stereotypes, high-level cognitive states, and the low-level processing of facial, vocal, and bodily cues. Our hope is that a dynamic interactive framework for person construal will provide a helpful guiding force in the burgeoning interdisciplinary effort to understand the perception of our social worlds.

References

Adams, R. B., Ambady, N., Nakayama, K., & Shimojo, S. (2010). *The Science of Social Vision*. New York: Oxford University Press.

Allport, G. W. (1954). *The nature of prejudice*. Oxford: Addison-Wesley.

Anderson, J. R. (2002). Spanning seven orders of magnitude: a challenge for cognitive modeling. *Cognitive Science, 26*, 85-112.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies of the malleability of emotion perception. *Psychological Science, 19*, 724-732.

Balcetis, E., & Dunning, D. (2006). See what you want to see: Motivational influences on visual perception.

Journal of Personality and Social Psychology, 91, 612-625.

Balcetis, E., & Lassiter, D. (2010). *The Social Psychology of Visual Perception*. New York: Psychology Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition *(1994) Handbook of social cognition, Vol 1: Basic processes; Vol 2: Applications (2nd ed)* (pp. 1-40). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York: Guilford Press.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*, 230-244.

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology, 92*, 179-190.

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155-162.

Blair, I. V., Chapleau, K. M., & Judd, C. M. (2005). The use of Afrocentric Features as Cues for Judgment in the Presence of Diagnostic Information. *European Journal of Social Psychology, 35*, 59-68.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004a). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15*, 674-679.

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004b). The Automaticity of Race and Afrocentric Facial Features in Social Judgments. *Journal of Personality and Social Psychology, 87*, 763-778.

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology, 83*, 5-25.

Blanz, V., & Vetter, T. (1999). *A morphable model for the synthesis of 3D faces*. Talk presented at the SIGGRAPH'99. Los Angeles.

Blascovich, J., Wyer, N. A., Swart, L. A., & Kibler, J. L. (1997). Racism and racial categorization. *Journal of Personality and Social Psychology, 72*, 1364-1372.

Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology, 55*, 726-737.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition *Stereotype activation and inhibition* (pp. 1-52). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Bodenhausen, G. V., & Macrae, C. N. (2006). Putting a face on person perception. *Social Cognition, 24*, 511-515.

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience, 2*, 370-374.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *A Dual-Process Model of Impression Formation: Advances in Social Cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Erlbaum.

Bruce, V., & Young, A. W. (1986). A theoretical perspective for understanding face recognition. *British Journal of Psychology, 77*, 305-327.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From Pixels to People: A Model of Familiar Face Recognition. *Cognitive Science, 23*, 1-31.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding race recognition with an interactive activation model. *British Journal of Psychology, 81*, 361-380.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 641-651.

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition, 3*, 81-117.

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences, 11*, 535-543.

Castano, E., Yzerbyt, V. Y., Bourguignon, D., & Seron, E. (2002). Who may enter? The impact of in-group identification on in-group/out-group categorization. *Journal of Experimental Social Psychology, 38*, 315-322.

Chaiken, S., & Trope, Y. (1999). *Dual-Process Theories in Social Psychology*. New York: Guilford.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*, 215-224.

Churchland, P. S., & Sejnowski, T. J. (1989). Neural representation and neural computation. In L. Nadel (Ed.), *Neural Connections, Mental Computations*. Cambridge: MIT Press.

Cisek, P., & Kalaska, J. F. (2005). Neural Correlates of Reaching Decisions in Dorsal Premotor Cortex: Specification of Multiple Direction Choices and Final Selection of Action. *Neuron, 45*, 801-814.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience, 33*, 269-298.

Cohen, J. D., & Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umilta (Ed.), *Attention and*

Performance XV (pp. 453-476). Cambridge, MA: MIT Press.

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition, 35*, 15-28.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion, 14*, 289-311.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*, 5-18.

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Journal of Personality and Social Psychology, 11*, 1139-1150.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). The nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510-540.

Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron, 28*, 287-298.

Eberhardt, J. L., Dasgupta, N., & Banaszynski, T. L. (2003). Believing is seeing: The effects of racial labels and implicit beliefs on face perception. *Personality and Social Psychology Bulletin, 29*, 360-370.

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience 2*, 704-716.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition, 44*, 227-240.

Fabes, R. A., & Martin, C. L. (1991). Gender and age stereotypes of emotionality. *Personality and Social Psychology Bulletin, 17*, 532-540.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482-498.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013-1027.

Fazio, R. H., Sanbonmatsu, D., Powell, M., & Kardes, F. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229-238.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878-902.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23*, 1–74.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science, 20*, 1183-1188.

Freeman, J. B., & Ambady, N. (in press). When two become one: Temporally dynamic integration of the face and voice. *Journal of Experimental Social Psychology*.

Freeman, J. B., Ambady, N., Midgley, K. J., & Holcomb, P. J. (in press). The real-time link between person perception and action: Brain potential evidence for dynamic continuity. *Social Neuroscience*.

Freeman, J. B., Ambady, N., Rule, N. O., & Johnson, K. L. (2008). Will a category cue attract you? Motor output reveals dynamic competition across person construal. *Journal of Experimental Psychology: General, 137*, 673-690.

Freeman, J. B., Johnson, K. L., Ambady, N., & Rule, N. O. (2010a). Sexual orientation perception involves gendered facial cues. *Personality and Social Psychology Bulletin, 36*, 1318-1331.

Freeman, J. B., Pauker, K., Apfelbaum, E. P., & Ambady, N. (2010b). Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology, 46*, 179-185.

Freeman, J. B., Rule, N. O., Adams, R. B., & Ambady, N. (2010c). The neural basis of categorical face perception: Graded representations of face gender in fusiform and orbitofrontal cortices. *Cerebral Cortex, 20*, 1314-1322.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron, 54*, 677-696.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*, 509-517.

Goff, P. A., Thomas, M. A., & Jackson, M. C. (2008). "Ain't I a Woman?": Towards an intersectional approach to person perception and group-based harms. *Sex Roles, 59*, 392-403.

Grainger, J., & Jacobs, A. M. (1998). On localist connectionism and psychological science. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition*. Mahwah, N. J.: Erlbaum.

Halberstadt, J., Winkielman, P., Niedenthal, P. M., & Dalle, N. (2009). Emotional conception: how embodied emotion concepts guide perception and facial action. *Psychological Science, 10*, 1254-1261.

Halberstadt, J. B. (2005). Featural shift in explanation-biased memory for emotional faces. *Journal of Personality and Social Psychology, 88*, 38-49.

Halberstadt, J. B., & Niedenthal, P. M. (2001). Effects of emotion concepts on perceptual memory for emotional expressions. *Journal of Personality and Social Psychology, 81*, 587-598.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.

Hess, U., Adams, R. B., Jr., Grammer, K., & Kleck, R. E. (2009). Face gender and emotion expression: Are angry women more like men? *Journal of Vision, 9*, 1-8.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruganski (Eds.), *Social Psychology, Handbook of basic principles* (pp. 133-168). New York: Guilford Press.

Hill, H., Bruce, V., & Akamatsu, S. (1995). Perceiving the sex and race of faces: The role of shape and colour. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 261*, 367-373.

Ho, C., & Jackson, J. W. (2001). Attitudes toward Asian Americans: Theory and measurement. *Journal of Applied Social Psychology, 31*, 1553-1581.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America, 79*, 2554-2558.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14*, 640-643.

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in Social Categorization: The role of prejudice and facial affect in race categorization. *Psychological Science, 15*, 342-345.

Inzlicht, M., Kaiser, C. R., & Major, B. (2008). The face of chauvinism: How prejudice expectations shape perceptions of facial affect. *Journal of Experimental Social Psychology, 44*, 758-766.

Izhikevich, E. M. (2006). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. Cambridge, MA: MIT Press.

Johnson, K. L. (2009). *Race is Gendered: Overlapping Stereotypes Confound Person Construal*. Talk presented at the Society for Personality and Social Psychology Annual Meeting as part of the "Intersecting Identities and Expressions: On the Combinatorial Nature of Social Perception" symposium. Tampa, FL.

Johnson, K. L., Freeman, J. B., & Pauker, K. (invited revision). Race is gendered: How Covarying Phenotypes and Stereotypes Bias Sex Categorization.

Johnson, K. L., Gill, S., Reichman, V., & Tassinary, L. G. (2007). Swagger, sway, and sexuality: Judging sexual orientation from body motion and morphology. *Journal of Personality and Social Psychology, 93*, 321-334.

Johnson, K. L., Pollick, F., & McKay, L. (2010). Social constraints on the visual perception of biological motion. In R. B. Adams Jr., N. Ambady, K. Nakayama & S. Shimojo (Eds.), *The Science of Social Vision*. New York: Oxford University Press.

Johnson, S. L., Eberhardt, J. L., Davies, P. G., & Purdie-Vaughns, V. J. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*, 383-386.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge: MIT Press.

Kite, M. E., & Deaux, K. (1987). Gender belief systems: Homosexuality and implicit inversion theory. *Psychology of Women Quarterly, 11*, 83-96.

Ko, S. J., Judd, C. M., & Blair, I. V. (2006). What the Voice Reveals: Within- and Between-Category Stereotyping on the Basis of Voice. *Personality and Social Psychology Bulletin, 32*, 806-819.

Kovács, G., Gulyás, B., Savic, I., Perrett, D. I., Cornwell, R. E., Little, A. C., et al. (2004). Smelling human sex hormone-like compounds affects face gender judgment of men. *Neuroreport, 15*, 1275-1277.

Kunda, Z., Davies, P. G., Adams, B. D., & Spencer, S. J. (2002). The dynamic time course of stereotype activation: Activation, dissipation, and resurrection. *Journal of Personality and Social Psychology, 82*, 283-299.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103*, 284-308.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23*, 571-579.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72*, 275-287.

Li, W., Piëch, V., & Gilbert, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience, 7*, 651-657.

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82*, 5-18.

Locke, V., Macrae, C. N., & Eaton, J. L. (2005). Is person categorization modulated by exemplar typicality? *Social Cognition, 23*, 417-428.

MacLin, O. H., & Malpass, R. S. (2001). Racial Categorization of Faces: The Ambiguous Race Face Effect. *Psychology, Public Policy and Law, 7*, 98-118.

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*, 397-407.

Macrae, C. N., & Martin, D. (2007). A boy primed Sue: Feature-based processing and person construal. *European Journal of Social Psychology, 37*, 793-805.

Maddox, K. B., & Gray, S. A. (2002). Cognitive Representations of Black Americans: Reexploring the Role of Skin Tone. *Personality and Social Psychology Bulletin, 28*, 250-259.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Marsh, A. A., Adams, R. B., & Kleck, R. E. (2005). Why do fear and anger look the way they do? Form and social function in facial expressions. *Personality and Social Psychology Bulletin, 31*, 73-86.

Masuda, S., Tsujii, T., & Watanabe, S. (2005). An interference effect of voice presentation on face gender discrimination task: evidence from event-related potentials. *International Congress Series, 1278*, 156-159.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review, 90*, 215-238.

McClelland, J. L. (1991). Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology, 23*, 1-44.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 16518-16523.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, A. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology, 30*, 211-222.

Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology, 14*, 481-487.

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion, 9*, 128-133.

Pauker, K., Rule, N. O., & Ambady, N. (2010). Ambiguity and social perception. In E. Balcetis & D. Lassiter (Eds.), *The Social Psychology of Visual Perception*. New York: Psychology Press.

Pauker, K., Weisbuch, M., Ambady, N., Sommers, S. R., Adams Jr., R. B., & Ivcevic, Z. (2009). Not so Black and White: Memory for ambiguous group members. *Journal of Personality and Social Psychology, 96*, 795-810.

Pettigrew, T. F., Allport, G. W., & Bartnett, E. O. (1958). Binocular resolution and perception of race in South Africa. *British Journal of Psychology, 49*, 265-278.

Port, R. F., & van Gelder, T. (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.

Read, S. J., & Miller, L. C. (1993). Rapist or "Regular Guy": Explanatory Coherence in the Construction of Mental Models of Others. *Personality and Social Psychology Bulletin, 19*, 526-540.

Read, S. J., & Miller, L. C. (1998a). *Connectionist models of social reasoning and social behavior*. Mahwah, NJ: Erlbaum.

Read, S. J., & Miller, L. C. (1998b). On the dynamic construction of meaning: An interactive activation and competition model of social perception. In S. J. Read & L. C. Miller (Eds.), *Connectionist models of social reasoning and social behavior*. Mahwah, N. J.: Erlbaum.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review, 1*, 26-53.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic Cognition: A Parallel Distributed Processing Approach*. Boston: Bradford Books.

Rolls, E. T., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology, 73*, 713-726.

Rosch, E. H. (1978). Principles of categorization. In E. H. Rosch & B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). *A general framework for parallel distributed processing*. Cambridge, MA: MIT Press.

Sherif, M. (1948). *An outline of social psychology*. New York: Harper.

Sinclair, L., & Kunda, Z. (1999). Reactions to a black professional: motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology, 77*, 885-904.

Smith, E. L., Grabowecky, M., & Suzuki, S. (2007). Auditory-visual crossmodal integration in perception of face gender. *Current Biology, 17*, 1680-1685.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology, 70*, 893-912.

Smith, E. R. (2000). Subjective experience of familiarity: Functional basis in connectionist memory *The message within: The role of subjective experience in social cognition and behavior* (pp. 109-124). New York, NY: Psychology Press.

Smith, E. R., & DeCoster, J. (1998). Knowledge acquisition, accessibility, and use in person perception and stereotyping: Simulation with a recurrent connectionist network. *Journal of Personality and Social Psychology, 74*, 21-35.

Smith, E. R., & DeCoster, J. (1999). Associative and rule-based processing: A connectionist interpretation of dual-process models *Dual-process theories in social psychology* (pp. 323-336). New York, NY: Guilford Press.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*, 161-168.

Smolensky, P. (1989). Connectionist modeling: Neural computation/mental connections. In L. Nadel, A. Cooper, P. Culicover & R. M. Harnish (Eds.), *Neural connections, mental computations*. Cambridge, MA: MIT Press.

Spivey, M. J. (2007). *The continuity of mind*. New York: Oxford University Press.

Spivey, M. J., & Dale, R. (2004). The continuity of mind: Toward a dynamical account of cognition *Psychology of Learning and Motivation* (Vol. 45, pp. 87-142). San Diego: Elsevier.

Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science, 15*, 207-211.

Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review, 96*, 58-83.

Strangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*, 207-218.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature, 400*, 869-873.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25*, 79-97.

Trope, Y., & Cohen, O. (1989). Perceptual and inferential determinants of behavior-correspondent attributions. *Journal of Experimental Social Psychology, 25*, 142-158.

Usher, M., & McClelland, J. L. (2003). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*, 550-592.

Valentin, D., Abdi, H., O'Toole, A. J., & Cottrell, G. W. (1994). Connectionist models of face processing: A survey. *Pattern Recognition, 27*, 1209-1230.

van Overwalle, F. (2007). *Social connectionism: A reader and handbook for simulations*. New York: Psychology Press.

van Overwalle, F., & Labiouse, C. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review, 8*, 28-61.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72*, 262-274.

Zebrowitz, L. A. (2006). Finally faces find favor. *Social Cognition, 24*, 657-701.

Zebrowitz, L. A., Fellous, J.-M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review, 7*, 194-215.

Zebrowitz, L. A., Kikuchi, M., & Fellous, J.-M. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology, 98*, 175-189.

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass, 2*, 1497-1517.