A Small-World Network Model of Social Inferences: Rethinking Social Cognition Through Resource-Rational Representations

Junsong Lu[1], Chujun Lin[2*]

[1] Department of Psychology, University of California San Diego, La Jolla, U.S.

[2] Department of Psychology, Columbia University, New York, U.S.

Author Note

Junsong Lu https://orcid.org/0000-0001-6987-6228

Chujun Lin https://orcid.org/0000-0002-7605-6508

Code for the network simulations are available at https://github.com/Junsong798/Small-World-Minds-Create-the-Illusion-of-Low-Dimensional-Social-Cognition-Under-Constrained-Stimuli.

The authors have no conflict of interest to disclose.

*Corresponding author *E-mail address*: cl4767@columbia.edu

Abstract

People make social inferences using hundreds of concepts, yet how these concepts are mentally represented remain debated. One tradition proposes that social inferences are organized along a few psychological dimensions (e.g., warmth), whereas another argues that they are complex and require more dimensions to capture—a view recently revived by naturalistic designs. Motivated by this debate, we propose a network theory explaining how social concepts are mentally represented and how social inferences are generated consequently. Specifically, social concepts form a small-world network through two principles: growth, whereby the network expands as new concepts are acquired, and preferential attachment, whereby new concepts preferentially connect to well-connected concepts. Using information theory, we show that these processes reflect rational use of limited cognitive resources. We apply the theory to revisit four important topics in social cognition with simulations. First, concepts acquired earlier become more central and exert greater influence on social inference, explaining why attractiveness, an early-acquired concept, produces a strong halo effect. Second, psychological dimensions identified across domains are not fixed structures but emergent properties of the network: earlier-acquired concepts attract more connections and are therefore more likely to constitute dimensions. Third, the dimensionality of social inference emerges from the interplay between social cues and network, with richer information producing higher dimensionality, reconciling low- and high-dimensional accounts. Fourth, interrater consistency of inference increases with the degree of concepts, offering methodological implications. Overall, our theory provides a resource-rational account of how the mind organizes social concepts and generates social inferences under cognitive constraints. (250/250)

Keywords: social cognition, trait inference, person perception, network science, computational modeling

It is the way of heaven to take from what has in excess in order to make good what is deficient. The way of man is otherwise. It takes from those who are in want in order to offer this to those who already have more than enough.

— Tao Te Ching (道德经)

Introduction

Across domains ranging from micro-level cellular systems to macro-level social and transportation networks, complex systems exhibit robust structures that far exceed what could be centrally planned, most notably scale-free topologies (Barabasi & Albert, 1999). In airport networks, for example, a small number of hub airports connect to orders of magnitude more destinations than most others (Guida & Maria, 2007). Such regularities can be captured by the "rich-get-richer" dynamic, in which nodes (e.g., an airport) with greater connectivity (e.g., air route) increasingly attract more connections over time. At its core, this dynamic reflects constraints on resource allocation: routing connectivity through a small number of airport hubs allows most locations around the world to be linked using relatively few routes and limited infrastructural investment. As articulated in the *Tao Te Ching*, written in ancient China more than two millennia ago, the way of humanity is to take from what is insufficient to serve what has surplus. This highlighted how resource constraints in the human world give rise to imbalanced yet efficient structures across domains.

Although such resource allocation challenges may seem most relevant at the material or societal level, similar constraints operate in human psychology. In navigating the social world, individuals continually evaluate and make inferences about a wide range of people and groups. These social inferences vary across situations, goals, and targets, encompassing rich content at multiple levels of granularity, from transient mental states and momentary emotions to stable personality traits and broad group stereotypes (Amodio, 2025; S. T. Fiske et al., 2002; Freeman & Ambady, 2011; Uleman et al., 2008). Also, social information is rich, noisy, and multimodal, encompassing physical appearance, behavior, affective signals, and contextual factors (Carlston, 1994; Mileva, 2026). Given the diversity of inferences and the environmental complexity, a central question in social cognition concerns how people represent these concepts in their mind and how inferences are made based on mental representation. Crucially, under finite cognitive resources, any such representation must accommodate to the environmental complexity while being parsimonious.

Research on this question have argued for a low-dimensional account. This account assumes that only a few "core dimensions" underpin judgments of personality, trait inference, face attribution, emotions, and stereotypes toward groups (Ekman & Friesen, 1971; S. T. Fiske et al., 2002; Goldberg, 1990; C. Lin et al., 2021; Oosterhof & Todorov, 2008; Rosenberg et al., 1968; Sutherland et al., 2013; Thornton & Tamir, 2020). In a typical setting, participants make inference on a range of social concepts given a target like faces or social group labels. Researchers then examine how these judgments covary across targets and use dimensionality reduction methods (e.g., factor analysis) to characterize the underlying structure of these inferences. For example, face evaluation is captured by two dimensions known as trustworthiness and face dominance (Oosterhof & Todorov, 2008).

However, it is controversial whether a low-dimensional structure can capture the full complexity of social inferences. A growing body of recent research supports the high-dimensional account that social cognition is more complex than a few dimensions can capture (Brooks et al., 2024; Connor et al., 2024; Cowen et al., 2020, 2021, 2024; Cowen & Keltner, 2017; Freeman et al., 2020; Freeman & Lin, 2025; Lu & Lin, 2025; Nicolas et al., 2022a, 2025). For example, 12 to 28 dimensions are needed to capture emotion inferences from faces, body, and speech (Cowen & Keltner, 2019), and 40 dimensions capture judgments of social groups (Nicolas et al., 2022b)

These inconsistent findings on dimensionality across domains call for theoretical integration. The debate over how many dimensions exist and what they represent reflects a lack of understanding of the mental structures, cognitive processes, and social contexts that give rise to these empirical patterns. Motivated by this debate, we proposed a network model of social inference that draws on insights from complex network theory. Our model serves two primary goals.

First, in response to the divergent empirical findings, we developed a theory-driven model of how social knowledge is mentally structured and how social inferences are generated from this structure in response to social information. Many research questions in psychology are ill-posed, as the same empirical phenomenon can arise from multiple underlying processes (J. R. Anderson, 1991; Lieder & Griffiths, 2020). Additional constraints are therefore needed to justify a particular process model. Here, we imposed a computational constraint: the mind has finite cognitive resources and should use them efficiently, following the resource-rational approach to cognitive modeling (Bhui et al., 2021; Lieder & Griffiths, 2020). Rather than directly specifying mental processes as in many classic theories, we identified the optimal processes for developing a mental representation under finite cognitive resource. These developmental processes naturally imply a structure of mental representation. We then modeled social inference as a dynamic process operating over this structure. This formalization provides a mechanistic advance over traditional models, which have largely described the static content of psychological dimensions without specifying how social inferences are dynamically generated from incoming information.

Second, we used this new account to explain established findings and make new predictions related to four important topics in social cognition, all of which have figured prominently in the low- versus high-dimensionality debate: (a) what concepts are central to social perception and why they are central; (b) why researchers consistently observe psychological dimensions across domains and why their content shares some similarities, bearing on whether dimensional models provide a sufficient account of social cognition; (c) why some studies observe low dimensionality whereas others identify high dimensionality, addressing whether this discrepancy reflects more than methodological differences; and (d) which social inferences are more reliable (e.g., showing stronger target effects) and why they are more consistent across perceivers.

Specifically, we argued that mental representations of social concepts like "warmth" are structured as a small-world network that grows in a specific way: as people learn more social concepts, the network expands while preserving many distinct nodes (i.e., concepts) and maintaining short paths between them, such that one concept can trigger another through a few intermediate concepts. By analogy, social connections between people are also structured as a small-world network (Davidsen et al., 2002): although society contains thousands of people, we can reach nearly anyone through only a few intermediaries—a friend of a friend..

We developed this small-world network account across five sections. First, we introduce the ongoing debate between low- and high-dimensional accounts of social inference, clarify the theoretical question at its core, and explain why it matters. Second, motivated by this debate, we propose a small-world network account that reconciles the two perspectives while explaining other important findings in social cognition. We described the network structure and its dynamic interaction with environmental input and provided a resource-rational analysis based on information theory to explain how this structure can be learned through development under finite cognitive capacity. Third, we formally analyzed the statistical properties of the network representation, linking its developmental origins to its structure and laying the foundation for connecting network dynamics to observable behavior. Fourth, we applied the model to four important topics in social cognition through formal derivations and simulations, reinterpreting existing findings and generating new predictions. Finally, we concluded by advocating a shift from focusing on fixed, static structure toward the flexible, dynamic processes through which social inferences are formed, providing a better account of the complexity of social cognition in the real world.

The Debate: Low- and High-Dimensional Accounts of Social Cognition

In this section, we briefly reviewed six decades of evidence supporting both low- and high-dimensional accounts. A dimension refers to an evaluative axis along which multiple social inferences of a perceiver systematically vary together when judging targets. We defined the low-dimensional account as findings in which social inferences show high covariation, suggesting that one or a few key concepts or certain aspects of evaluation play a central role in shaping judgments. We highlighted evidence for the beauty-is-good effect and the emergence of core psychological dimensions. In contrast, the high-dimensional account argues that many more dimensions are needed to explain the associations among social inferences, implying a richer, multidimensional representation. As elaborated later, both accounts provide implications about the mental representation of social concepts and carry important theoretical implications.

Low-Dimensional Social Cognition

Dion and colleagues in 1972 demonstrated that people tend to attribute positive qualities (e.g., socially desirable) to physically attractive individuals—known as the beauty-is-good effect (Dion & Walster, 1972). In these studies, participants view face images that vary in attractiveness and rate them on multiple traits. Large-scale meta-analyses confirmed the robustness of this stereotype (Eagjy et al., 1991; Feingold, 1992; Jackson et al., 1995). For example, Eagjy and colleagues (1991) showed that this effect is more pronounced for interpersonal evaluations (e.g., sociability) and weaker for capability-related evaluations like intelligence and integrity. These findings have sparked research investigating why people hold such a "stereotype". Specifically, whether it reflects an objective correlation between attractiveness and positive traits. While some anecdotal links were found (Jackson et al., 1995; Kanazawa, 2011; Langlois et al., 2000; Sheehan & Hamermesh, 2024; Zebrowitz et al., 2002), particularly with objective intelligence and health (De Jager et al., 2018; Kanazawa, 2004; Thornhill & Gangestad, 1993; Weeden & Sabini, 2005; Zebrowitz et al., 2002, 2003), large-scale evidence indicated minimal correlations between ground-truth attractiveness and positive traits (Feingold, 1992; Mitchem et al., 2015). Therefore, the beauty-is-good effect reflects a cognitive heuristic in which even a single concepts (e.g., attractiveness) shapes multiple social inferences, illustrating a low-dimensional structure.

Whereas the beauty-is-good stereotype reflects low-dimensionality within a single association (e.g., the attractiveness-intelligence association), a parallel line of research investigates how low-dimensional structures emerge from the interrelations among many social inferences. This line of research dates back to Asch's seminal work on central traits on impression formation (Asch, 1946), where he showed that using certain concepts to describe a target can substantially alter perceivers' overall impression of them—a pattern similar to the beauty-is-good effect. Subsequent work showed that some trait judgments have greater influence on overall social inference because they are more strongly interconnected with other judgments (Wishner, 1960), and the complicated idea of central trait can be better captured by dimensions underlying these interrelated judgments derived from dimensionality reduction methods (Rosenberg et al., 1968). This tradition has since identified a small set of recurring dimensions across different social domains including stereotypes, trait inferences, facial attribution, most prominently warmth and competence. Although studies have proposed additional dimensions (e.g., morality) depending on the population, stimuli, and task (e.g., Abele & Bruckmüller, 2011; Brambilla et al., 2021; Lin et al., 2021; Goldberg, 1990), they generally agree that perceivers' social perception consist of numerous social inferences exhibits high covariance, and that this high covariance can be decomposed into a few core dimensions.

High-Dimensional Social Cognition

Despite the dominance of low-dimensional accounts in the field, others have long argued that social inferences are more complex and captured by more dimensions since the 1940s. For instance, Cattell's trait sphere theory proposes that the correlations between perceived traits may not be driven by a small number of latent dimensions but their shared components (e.g., overlapping behavior); based on people's ratings on familiar others, he identified 16 personality dimensions (Cattell, 1943). While this high-dimensional account was less popular than the low-dimensional account, recent studies have increasingly revealed high-dimensional findings (Brooks et al., 2024; Cowen et al., 2020, 2021, 2024; Cowen & Keltner, 2017; Freeman et al., 2020; Freeman & Lin, 2025; Lu & Lin, 2025; Nicolas et al., 2022a).

These recent studies shares two key features. First, rather than starting with a small set of predefined dimensions and constructing judgment terms to confirm them as in prior research, they uses data-driven approaches that are agnostic to the expected dimensions and utilizes diverse judgment terms or response formats (Cowen et al., 2024; Lu & Lin, 2025; Stamkou et al., 2024). For instance, Brooks and colleagues (2024) asked participants to rate 4,659 face images on 48 emotions and mental states, and found 28 facial expression dimensions. Using open-ended responses, Nicolas and colleagues (2025) identified 40 distinct dimensions that capture free descriptions of social groups. Second, rather than using constrained stimuli such as face images or short verbal descriptions, recent work employs more naturalistic stimuli ranging from ambient images with contextual backgrounds to dynamic videos that convey richer multisensory information. For example, Cowen and Keltner (2017) showed participants 2,185 naturalistic video clips that depicted significant life events (e.g., weddings), and identified 27 emotion dimensions.

Dimensional Accounts as Mental Representations

While psychological dimensions identified using multivariate statistical methods are statistical summaries of social inferences, findings from both low- and high-dimensional accounts can offer a preliminary indication of how concepts may be mentally represented and how social inferences may be generated from their representational structure.

Low-dimensional findings such as the beauty-is-good effect suggest that people systematically treat certain characteristics, such as attractiveness, as more informative about others than other characteristics. This implies that certain concepts and related inferences have structural advantages in shaping social perception, potentially grounded on an evolutionary basis. Take attractiveness as an example: selection pressures have favored heightened sensitivity to cues of low mate quality like anomalous, unattractive faces. This heightened sensitivity may lead perceivers to attribute negative characteristics to unattractive targets and positive characteristics to attractive ones, even at the cost of overgeneralizing these cues to characteristics for which attractiveness is not actually diagnostic (Zebrowitz, 2017; Zebrowitz et al., 2003).

Similarly, dimensional models make a broader claim about the structure of social concepts. The stereotype content model (SCM), for example, proposes that social information is organized across two axes: warmth and competence (S. T. Fiske et al., 2002, 2007). Thus, an inference about a target (e.g., that they are extraverted) can inform the target's perceived position along these two latent axes, which in turn shapes inferences about other characteristics as well as broader cognitive and behavioral consequences. This idea is further elaborated in the extension of SCM, the behaviors from intergroup affect and stereotypes (BIAS) map, which proposes that perceiving a target as high in competence but low in warmth elicits envy and corresponding hostile behavior (Cuddy et al., 2008). Consistent with this account, perceived warmth and competence from faces have been shown to predict consequential judgments, including evaluations of political candidates and voting behavior (Giacomin & Rule, 2020; Todorov, 2005), even among children (Antonakis & Dalgas, 2009).

If mental representations shape how social inferences are generated and their corresponding behavioral consequences, evidence for high-dimensional social cognition raises a fundamental question: to what extent do the few dimensions identified in previous research capture the underlying structure of social concepts? As mentioned above, evolutionary theories have been invoked to explain why particular dimensions, including warmth, competence, gender, attractiveness, and even the Big Five personality traits, are prominent in social cognition (S. T. Fiske et al., 2007; A. E. Martin & Slepian, 2021; Nettle, 2006; Sutherland et al., 2013; Zebrowitz et al., 2002). Such explanations, however, are often post hoc and therefore provide limited guidance for predicting which dimensions should emerge and how they should be organized. A high-dimensional account may instead emphasize the complexity and uncertainty of social environments, raising the possibility that richer representations are needed to capture the structure of social concepts (Connor et al., 2024; Lu & Lin, 2025). Yet the strong predictive power of a small number of dimensions also raises the question of why additional dimensions would be necessary if two dimensions were sufficient to predict consequential real-world outcomes, including electoral outcomes (Antonakis & Dalgas, 2009).

These inconsistent findings and unresolved questions can be partly attributed to how researchers have typically approached the study of social inference. The summarizing nature of dimensions provides only a coarse-grained view of how social inferences are structured. Moreover, statistical methods used to identify dimensions, such as factor analysis, impose specific assumptions about the structure and distribution of the data that may not necessarily align with the substantive theories of interest (Haines et al., 2025). When a dimensional approach is implemented using latent variable models such as factor analysis, the data are necessarily represented in terms of an underlying dimensional structure. Yet it remains unclear whether this assumed structure faithfully reflects the organization of concepts in the mind, and therefore whether it can provide insight into how many dimensions actually exist or what their contents are. Fundamentally, this debate reflects our limited understanding of how concepts relevant to social cognition are mentally represented and how social inferences are generated from this representational structure. It therefore leaves open several possibilities: the mind may rely on a simple but efficient representation, a complex and comprehensive representation, or a flexible representation that adapts to the structure of the environment.

Moving from Methodological Discrepancy to Theoretical Implications

Given the inconsistent findings regarding how many core dimensions underpin social cognition, researchers have sought to reconcile these divergent accounts. One prominent explanation attributes the apparent divergence to methodological differences.

Most studies on psychological dimensions rely on dimensionality reduction techniques that reduce the covariance structure among social inferences to a few latent factors. However, there lacks a hard criterion for determining how many dimensions are adequate in recovering the covariance structure: adding more dimensions will always account for the data better, leaving the choice to researchers' judgments (C. Lin et al., 2025). Beyond statistical methods, differences in stimulus sampling and design may also contribute to divergent findings. As implied by the Continuum Model of impression formation, when more relevant information is available, perceivers may shift from coarse-grained categorical processing toward more fine-grained and individualized representations of the target (S. T. Fiske & Neuberg, 1990). Indeed, evidence supporting low-dimensional accounts has exhibited some level of constraints. When the emotion cues are removed, participants outside the United States do not consistently categorize emotions according to the six basic emotions (Gendron et al., 2014). This may partly reflect limited stimulus sampling in social cognition, where White faces are disproportionately represented and facial expressions are often limited to highly prototypical forms (Satchell et al., 2023). Similar constraints arise in factor-analytic studies: including a set of preselected, correlated traits can induce high correlations among judgments and thereby produce an apparent latent dimension. Thus, the dimensionality observed in a given study may depend on the stimuli, statistical method, and researchers' analytic decisions.

However, treating the debate as purely methodological leaves important findings from the past decades unexplained and risks imposing overly simplified assumptions about the mind. Such simplification may therefore hinder a more complete understanding of social cognition.

First, attributing the debate to methodological differences cannot explain why certain dimensions consistently emerge across domains and paradigms. Importantly, studies across distinct and even disconnected areas of psychology (e.g., developmental, personality, social) have identified the "Big Two" dimensions, agency/competence and communion/warmth, as capturing how people process, perceive, and navigate their social worlds. This convergence has been framed as "the primacy of gender", which argues that gender is the most important social category shaping the lens through which we understand the social world (A. E. Martin & Slepian, 2021). Moreover, dimensions identified under naturalistic designs include "core dimensions" previously identified in low-dimensional studies (Connor et al., 2024; Lu & Lin, 2025). Thus, high-dimensional accounts appear to identify additional core dimensions beyond the "Big Two," rather than an entirely different set of dimensions. Indeed, if dimensions serve evolutionary purposes, as suggested by some studies (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Zebrowitz, 2017), their number and content should reflect key challenges in the evolution history and how we solve them, rather than being arbitrary and determined by the elicitation method.

Second, considering dimensionality as a function of stimuli type and cue complexity makes an unrealistic assumption that the mind simply maps social information onto social inferences. As illustrated by the beauty-is-good effect, empirical associations between attractiveness and perceived intelligence did not reflect the true covariation of cues related to these concepts. For example, Talamas et al. (2016) found a correlation of .81 between attractiveness and perceived intelligence, despite the absence of a comparably strong empirical association between attractiveness and measured intelligence (Mitchem et al., 2015). Thus, social inference cannot be understood as merely mapping available cues onto corresponding judgments. Even if the mind can reconstruct social information, cognitive constraints may limit both how and how much it can reconstruct (Lieder & Griffiths, 2020). A satisfactory account of the debate should therefore consider both how much social information is available and how it interacts with limited cognitive resources to produce variation in the number and content of dimensions.

Our model development is thus anchored in a theoretical perspective while also incorporating concerns raised by accounts emphasizing methodological differences. As a starting point, we agreed that dimensions are statistically derived and can therefore serve as summaries of social inferences. Latent factor models were originally developed in psychometrics from the idealist perspective, in which high covariance between measures was taken as evidence for an underlying psychological entity (Borsboom et al., 2004). More specifically, these models assume that individual concepts, such as perceived hospitality of a target, are influenced by a set of latent factors, such as warmth and competence, without unique causal relationships among the concepts themselves. However, recent evidence challenges this assumption. Using methods for determining the appropriate number of dimensions, Connor and colleagues (2024) found that 14 dimensions accounted for only 38% of the total variance in social attributions. Similarly, Lu and Lin (2025) showed that even 25 dimensions accounted for less than 15% of the variance in unconstrained social inferences from videos. The limited variance explained suggests that a substantial portion of the variation in social inferences is not captured by common dimensions and may instead reflect local associations, such as shared cue use (Cattell, 1943; Cramer et al., 2012; Epskamp et al., 2018).

In modern psychometrics, such unique associations are often modeled as psychometric networks, in which nodes represent concepts and links represent their correlations (Borsboom et al., 2021; Epskamp et al., 2018). Unlike latent-variable models, network models allow concepts to be directly influenced by their neighboring concepts rather than by an underlying common dimension. We adopt this network perspective but formalize it as a generative model of social inference rather than a statistical model of behavioral measures. This is because empirical measures of social inference (e.g., perceived warmth) reflect both mental processes and experimental stimuli and therefore cannot directly reveal the underlying mental representation. In addition, we distinguish between representational structure and inferential processes by specifying how social concepts are organized and how social inferences emerge from this structure, respectively. These components can then be evaluated against empirical findings.

A Small-World Network Model of Social Cognition

From Early Network Models to Small-World Network Structure

Cognitive scientists have long viewed the mind as inherently associative and thus have adopted a network approach to investigate mental processing (Collins & Loftus, 1975; Siew et al., 2019). Early work used networks to model semantic memory. In their seminal study, Collins and Quillian (1969) proposed that concepts are represented as nodes within a hierarchical tree. While elegant, this model is best suited for hierarchically organized concepts (e.g., animal categories) and does not easily generalize to capture mental representations of other concepts, such as color, faces, or social relationships (Kemp et al., 2007; Steyvers & Tenenbaum, 2005).

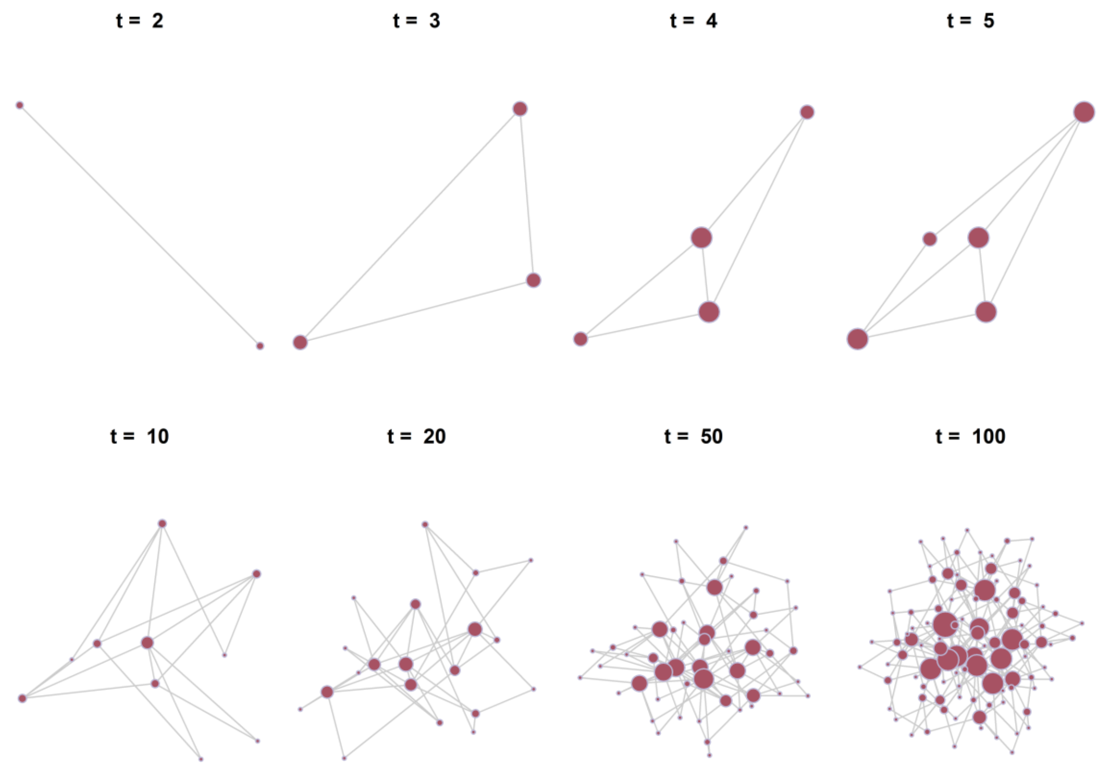
Researchers in sociology and physics provided insights into an alternative network structure that might inform mental representations: the small-world network. In the early 1960s, sociologist de Sola Pool and mathematician Kochen introduced the concept of "small world" in their analysis of social connections (de Sola Pool & Kochen, 1978). Their manuscript has been widely circulated and inspired the famous small-world experiment by Milgram (1967). The central idea, popularized by playwright John Guare as "six degrees of separation," posits that anyone in a social network can be reached through a few intermediary connections.

Meanwhile, in a separate line of work unknown to sociologists at the time, Erdős and Rényi developed a random graph model (1960). They showed that when the average number of connections per node in this network increases beyond a certain threshold, a giant component emerges—an extensive cluster that connects the majority of nodes in the network. In this configuration, the maximum distance between any two nodes (the network's diameter) decreases sharply. Relative to the size of the network (the number of nodes), this diameter becomes so small that nearly any pair of two nodes become connected through only a few steps. This random graph model provides a useful demonstration of how small-world properties can arise once a network becomes sufficiently connected.

The random graph model assumes connections between nodes are formed randomly, typically resulting in a Poisson degree distribution. However, this assumption often fails to capture real-world systems, where link formation is driven by non-random mechanisms (Barabási, 2009; Barabasi & Albert, 1999; M. Newman, 2018). Instead, they often exhibit a power-law distribution, also known as a scale-free architecture: a few nodes (hubs) have many connections, while most have few. This architecture commonly arises from two generative mechanisms observed in a range of real-world networks, including cellular networks, social structures, neural networks, and the Internet (Barabási, 2009): growth, whereby networks expand by adding new nodes, and preferential attachment, where new nodes are more likely to connect with already well-connected ones (Barabasi & Albert, 1999) (Figure 1). These two mechanisms not only produce scale-free networks but also lead to small-world properties (see Appendix A): hubs act as shortcuts that link distant parts of the network, drastically reducing the average path length between any two nodes.

Figure 1.

Evolution of the Network Generated by Growth and Preferential Attachment



Notes. Simulation of a growing network with 100 nodes, illustrating the two generative mechanisms: growth and preferential attachment. Each new node forms two links to existing nodes. Over time (t = 2, 3, 4, 5, 10, 20, 50, 100), this generative process produces a scale-free architecture with a few highly connected hubs. These hubs act as shortcuts that drastically reduce the average path length between nodes, giving the network small-world properties.

Generating the Structure of a Small-World Mind: Growth and Preferential Attachment

We propose that the same generative mechanisms—growth and preferential attachment—also govern mental representations of social inferences. While the idea that conceptual knowledge in social cognition may be structured as a network in the mind is not new (Freeman & Ambady, 2011; Freeman & Lin, 2025; Stolier et al., 2018, 2020), the specific structural characteristics and underlying generative processes of such network remain unclear. Our account fills this gap by arguing that mental representations of social inferences are generated through growth and preferential attachment, giving rise to a network with small-world properties that balance complexity and efficiency.

The first generative mechanism, growth, addresses a fundamental computational challenge: how finite human minds can build increasingly complex social representations while keeping them tractable. This balance between complexity and tractability aligns with Piaget's theory of cognitive development, in which individuals assimilate new information into existing cognitive structures or accommodate it by modifying those structures (Piaget, 2013). Over time, these processes progressively build richer representations in a way that supports navigating complex social environments. Importantly, the time at which a concept enters the network is shaped by the general sequence of our life experiences. For example, children encounter parents from the earliest stages of life, whereas concepts related to romantic partners emerge only later as new forms of social relationships become relevant. Thus, concepts related to family relationships are likely to be acquired earlier than concepts related to romantic relationships. At a more fine-grained level, concepts associated with readily accessible social information are also likely to be acquired earlier. For example, concepts related to physical attractiveness may be learned earlier than more abstract personality concepts because cues relevant to attractiveness, such as facial appearance, are directly available in everyday social interactions, providing frequent information that facilitates the formation of the corresponding concept.

The second generative mechanism, preferential attachment, characterizes how new information (e.g., new concepts) is incorporated into established mental representations. Intuitively, it follows a “rich-get-richer” pattern, where the probability of a new node (concepts) connecting to a exist node i is proportional to the number of connections, or degree, of node i (k_i):

$$\Pi(i) = \frac{k_i}{\sum_j k_j},$$

Although preferential attachment has not yet been applied to research on social inferences, it offers a clear prediction: new social concepts tend to be integrated with concepts that are already central in the network.

Together, growth and preferential attachment provide generative mechanisms through which mental representations of social inferences can develop into small-world networks. This framework differs from dimensional accounts, especially low-dimensional accounts that posit a small set of fixed, core dimensions underlying social cognition. Although the mind has limited cognitive capacity, and this constraint may favor low-dimensional representations, a fixed representation may not be well suited to accommodating the complexity of the social environment. Instead, we take a developmental perspective and explain how a rich mental representation can be built gradually as learning proceeds.

A Resource-Rational Analysis of Growth and Preferential Attachment

While growth and preferential attachment specify how a small-world mental representation develops, it remains unclear why these principles, borrowed from network science, are justified for a mental model, or whether other learning mechanisms are more appropriate. The resource-rational framework provides a way to address this question by evaluating mechanisms in terms of both cognitive limitations and rationality. Among multiple plausible processes, a theoretically more reasonable mechanism would be the one favoring efficient use of limited cognitive resources (Lieder & Griffiths, 2020). We therefore evaluate growth and preferential attachment from this perspective.

Growth is computationally advantageous because adaptive knowledge does not need to be acquired all at once. Even if a representation is useful—for example, for inferring the traits of a potential mate or collaborator—it can be acquired when it becomes relevant rather than being fully specified from birth (Boyer & Barrett, 2015). This strategy avoids the cognitive cost of maintaining complex knowledge before it is needed. It is also consistent with the proposed hierarchy of social needs across the lifespan (Kenrick et al., 2010; Neel et al., 2016). For example, self-protection precedes status seeking, which in turn precedes mate seeking. Concepts relevant to social threat (e.g., warmth) may therefore be acquired earlier, whereas concepts related to mating and social status may develop later.

Growth thus implies a sequential learning process. This leads to a further question: how should newly acquired knowledge be incorporated into an existing representation while cognitive resources remain limited? We propose that preferential attachment provides a resource-rational solution. To illustrate this, we revisited the analysis by Dasgupta and Griffiths (2022), who derived an optimal learning strategy under limited cognitive capacity. We showed that their optimal strategy is equivalent to preferential attachment.

Generally, learning can be understood as mapping new knowledge onto knowledge that has already been acquired, such as classifying new objects into existing categories or linking new concepts to concepts the learner already knows. A resource-rational learner should favor a mapping π that requires fewer representational cost. Consider classification as an example. According to information theory, the representational cost of a mapping π that assigns N objects onto a set of categories c can be quantified by the entropy of the resulting category distribution:

$$H(\pi) = -\sum_i P_\pi(c_i) log P_\pi(c_i)$$

where the probability of a given category c_i is simply the proportion of objects assigned to it,

$$P_\pi(c_i) = \frac{n_i}{N}.$$

Here, entropy measures the average number of bits needed to specify which category a new object is mapped to under an optimally efficient code. Higher entropy means that the mapping π requires more information to represent and is therefore more cognitive demanding. Because the mapping π is probabilistic rather than a set of fixed rules between objects and clusters, the assigned cluster c_i is considered as a random variable over which entropy is calculated.

If cognitive resources are limited and available in a fixed average amount, a resource-rational learner should favor mappings that can be represented around this budget. This can be formalized using an exponential distribution:

$$P(\pi) \propto exp(-kH(\pi)),$$

which assigns exponentially lower probability to mappings with higher entropy, given a fixed average mental capacity of $1/k$. This distribution implies a conditional probability of assigning any new object to a cluster given the current mapping. The probability of a particular assignment j is determined by the additional entropy it introduces,

$$P(\pi_j \mid \pi) \propto exp(-k(H_j - H)),$$

where H_j denotes the entropy of the mapping after the item joins cluster j (π_j). Intuitively, the less costly a way of integrating new information into existing knowledge, the more likely it is to be used. Dasgupta and Griffiths (2022) show that this conditional probability has a precise consequence for how any single piece of new knowledge should be integrated: mapping it onto a cluster that already contains n_j items lower entropy more than mapping it onto a sparsely populated or newer cluster:

$$P(\pi_j \mid \pi) \propto n_j.$$

Mathematically, this formula is equivalent to a discrete-time stochastic process called the Chinese restaurant process, which has been widely used in nonparametric Bayesian models to explain categorization (Aldous, 1985; J. R. Anderson, 1991; Gershman & Blei, 2012) and stereotyping (Gershman & Cikara, 2023). In stereotype formation, this manifests the tendency to categorize new individuals into more prevalent groups. Importantly, this formula is also similar to preferential attachment, if we analogize assigning links to existing nodes as assigning the number of objects into clusters:

$$\Pi(i) \propto k_i.$$

Thus, growth and preferential attachment are principled mechanisms following information theory. Growth justified the basic setup of sequential learning, whereas preferential attachment prescribes how new knowledge is integrated across time points. Together, they generate mental representations that efficiently organize knowledge under a fixed cognitive budget.

Dynamics in a Small-World Mind: Spreading Activation and Environmental Inputs

Having established the structure of social concept representations, we now turn to the dynamics that operate within this structure. Specifically, we illustrate how the network interacts with environmental inputs (e.g., social cues) to generate social inferences.

In network models of cognitive processing, such as semantic memory, a central mechanism is spreading activation, in which activation of one concept spreads to and activates adjacent concepts in the network (Collins & Loftus, 1975). Spreading activation typically has two important characteristics. First, activation propagates recursively: initially activated nodes can receive feedback from related nodes that become activated later (Siew, 2019). This recursive property creates dynamic patterns of mutual reinforcement within local regions of the network. Second, activation carries a limited cognitive resource: its signal strength decays over time and distance, eventually returning to baseline resting levels (J. R. Anderson, 1983; Dell, 1986). Therefore, in a large network with long average path lengths—meaning many steps are required to connect distant nodes—activation can quickly dissipate before reaching distant parts of the network. Small-world networks, on the other hand, keep average path lengths short by linking distant regions through well-connected hubs, allowing activation to reach a large portion of the network efficiently before the signal fades.

While network structure determines how far and how quickly activation can spread, environmental inputs determine where activation begins. For instance, prior work conceptualizing mental representations of social inferences using networks, the dynamic interactive theory (Freeman & Ambady, 2011), proposes that nodes indicating social concepts differ in terms of their sensitivity to environmental inputs. Specifically, valence-related inferences are more sensitive to visual and auditory inputs resembling emotional expressions, enabling rapid automatic detection of social threat or safety. In contrast, more abstract inferences, such as personality traits, typically require detailed observation of behavior and sustained processing.

The interaction between environmental inputs and spreading activation can be thought of as a two-step process: environmental inputs selectively activate particular nodes in the network, and these nodes then transmit activation to their neighbors through recursive spreading. In a small-world network, this means that even activation that begins in a highly localized region—such as valence-related inferences triggered by emotional expressions—can quickly propagate to distant regions of the network, reaching conceptually unrelated inferences. Because this activation originates from the same source, the resulting spread produces similar patterns and levels of activation across a wide variety of nodes. As a result, in a network with small-world properties, limited environmental input can create the appearance of broad, coordinated activations in mental representations that resemble low-dimensional structures. On the contrary, rich environmental input can create heterogeneous activations in mental representations that exhibit high-dimensional structures.

Statistical Properties of the Small-World Mind

While we have elaborated the development of the small-world network and network dynamics, it is unclear what these processes imply for social cognition and whether they can ultimately provide insight into how social inferences are made. In this section, we bridge this gap by linking the generative processes to the resulting mental structure and network dynamics. Specifically, we formally analyze the statistical properties of the network, laying the foundation for deriving and interpreting the model's predictions about social cognition in the next section.

To provide a first glance of the network's macroscopic structure, we mapped the two generative processes onto the degree distribution, which characterizes the number of connections associated with each concept. We also discussed the additional constraints that a finite mind imposes, which create densely connected regions resembling clusters. Second, we derived the probability that two concepts are connected at a given distance under this degree distribution, showing that concepts can remain extremely close to each other despite the large network size. Finally, based on these derived distances, we calculated the correlation between the activations of two concepts under spreading activation, mimicking the empirical correlations between social inferences. This analysis connects the network's statistical properties to its dynamics, and ultimately to observable behaviors.

Degree Distribution and Communities: From Development to the Structure

In a small-world network generated through growth and preferential attachment, a node's degree reflects when it entered the network. Earlier-acquired concepts have more opportunities to accumulate connections, while preferential attachment makes this process increasingly uneven across nodes. Specifically, for a network that has been developed for *t* steps where each step a new node is acquired, the degree of a node entered at t_i is:

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta},$$

where m is the number of links that each new concept brings to the network, and β is the dynamical exponent that determines how quickly degree accumulates, which equals 0.5 (See Appendix B for derivation). The larger t_i is, the smaller the node's degree at *t*. When $t_i = t$, the node has only the m links it brings to the network.

Importantly, the degree increases in a sublinear manner (β < 1), preventing any single node from accumulating an unbounded number of connections that the mind cannot accommodate. An example of the resulting network is shown in panel (a) of Figure 2. Panel (b) shows how the degrees of five social concepts evolve over time.

Because entering time uniquely maps onto degree, we can derive the probability of nodes with a specific degree k_i from the probability of entering the network at time t_i (see Appendix B). This yields the power-law degree distribution:

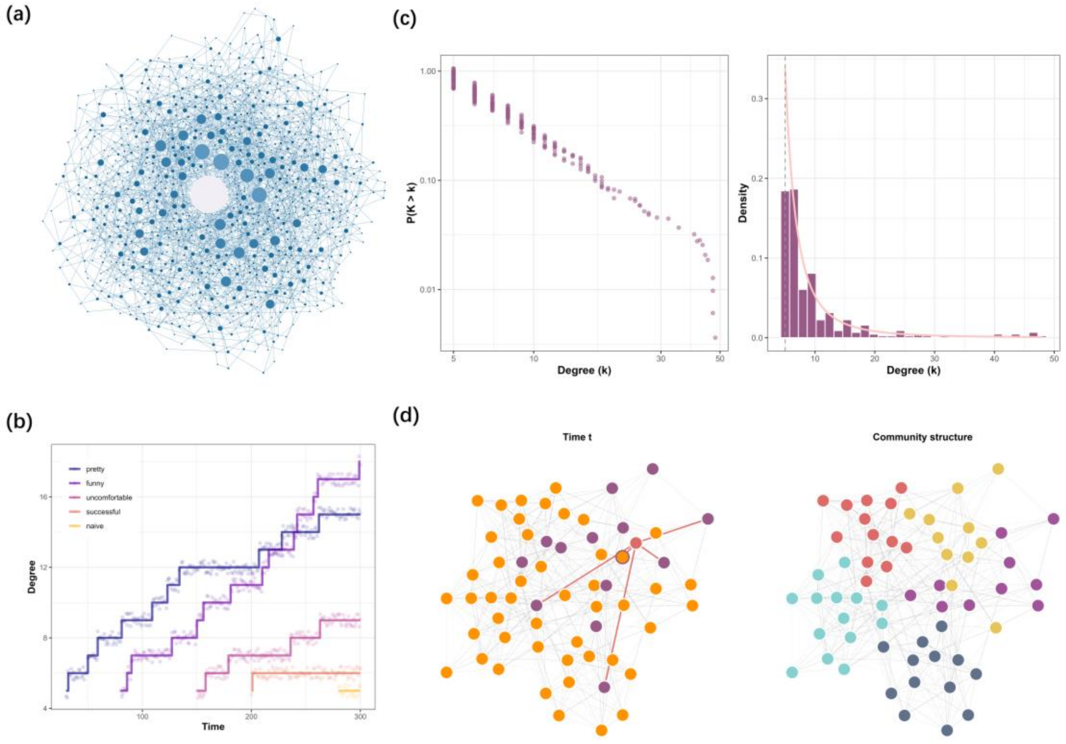
$$P(k) \propto k^{-\gamma},$$

where $\gamma = 1 + 1/\beta$ = 3. This power-law distribution characterizes the scale-free property: most concepts have relatively few connections, whereas a small number of concepts have exceptionally high degree, spanning a much broader range of connectivity. Specifically, the probability of observing high-degree nodes decreases drastically as degree increases (see panel c of Figure 2). For example, a node with a degree of 1000 would be approximately $1/1000^3$ as likely as a node with one degree. Therefore, a few high-degree concepts can serve as hubs that link otherwise weakly connected parts of the network, facilitating connections between concepts that might otherwise remain disconnected or only indirectly connected.

Beyond producing an extremely imbalanced degree distribution, preferential attachment can also generate locally clustered regions. The original formulation of preferential attachment requires a new concept to access the degree of all existing nodes in the network when selecting where to form its links. Such global access, however, is computationally costly. A cognitive system may instead have access to only a fraction of the degree information available in the network, corresponding to the information-filtering principle in complex network theory. This constraint resonates with the idea of local preferential attachment in psychology: once a new concept brings m links to the network and forms its first link with a particular node *i*, its remaining m-1 links are attached to neighbors of node *i*, with the probability of attachment proportional to their degrees (see panel d of Figure 2). Thus, except for the first link, the new concept needs to evaluate the degrees of only a small, locally accessible set of nodes rather than those of the entire network, substantially reducing the computational cost of network growth. Because this process distributes links within a small, local neighborhood, nodes within that neighborhood become densely interconnected, giving rise to "communities" that resemble clustered structures.

Figure 2

Network Structure Emerging from Growth and Preferential Attachment



Notes. (a) An example of the network generated through sequential growth and local preferential attachment. Each node represents a social concept, with larger nodes indicating concepts with higher degree. The heterogeneous node sizes illustrate the highly uneven distribution of connections produced by preferential attachment. (b) Degree accumulation for five social concepts (pretty, funny, uncomfortable, successful, naive) acquired at different points during network development (t = 300). Entering time for each of these concepts is based on the age-of-acquisition dataset of English words (Kuperman et al., 2012). Earlier-acquired concepts accumulate connections over a longer period and consequently are more likely to reach higher degrees. (c) The resulting degree distribution. The left panel shows the complementary cumulative distribution, $P(K \geq k)$, as a function of degree $k$on log–log axes. The approximately linear relationship indicates a power-law tail, consistent with a scale-free degree distribution. The right panel shows the corresponding empirical degree density, with the power-law distribution overlaid in red. Most concepts have relatively few connections, whereas a small number of concepts have substantially higher degrees. (d) Emergence of community structure through local preferential attachment. The left panel shows the network at an intermediate stage of development (time = t). A new node, annotated in red, enters the network and first establishes a link to an existing node, highlighted by a purple border. Its remaining links are then distributed among the neighbors of this purple-bordered node, as indicated by the red links. The right panel shows the resulting community structure, with nodes assigned to different communities based on their densely interconnected local regions (Christensen et al., 2023; Fortunato & Hric, 2016; Kritschgau et al., 2024; Traag et al., 2019). The comparison illustrates how repeatedly restricting preferential attachment to the neighborhood of an already connected node concentrates links locally and gives rise to distinct communities.

Node Distances

The scale-free property derived above has direct impact on the distance between nodes. Specifically, scale-free networks exhibit the ultra-small-world property (R. Cohen & Havlin, 2003; Maslov et al., 2003). While randomly connected networks can display small-world properties that characterize a short maximum distance between nodes, scale-free networks compressed connectivity even further. Specifically, the largest distance between nodes in scale-free networks scales with *ln(ln N)* rather than the typical scaling *lnN* observed in random networks, where *N* represents network size (see Appendix A). The average distance also behave similarly (R. Cohen & Havlin, 2003). Because *ln(ln N)* grows extremely slowly, even very large scale-free networks maintain remarkably short distances between nodes. This unusually slow increase in the largest path length relative to network size arises from the presence of hubs—a few nodes with exceptionally high connectivity that can link distant network regions in remarkably few steps (panel a of Figure 2). Applying this to mental representations of social concepts, it means that even conceptually distant concepts can potentially influence one another despite their conceptual separation.

To understand how closely connected nodes are within a small-world network, we can derive the probability that two nodes with degree k_i and k_j are connected at a distance *d* (e.g., with *d*-1 intermediate nodes). In large networks, this probability can be approximated by the expected number of links between these two nodes (see derivations in Appendix C). We derived this quantity based on the configuration model in complex network theory (Bertotti & Modanese, 2019; Courtney & Bianconi, 2016), which is a simplified mathematical framework for analyzing a complex network. The result showed that the expected number of links between two nodes at distance *d* is proportional to the product of their degrees (Appendix C):

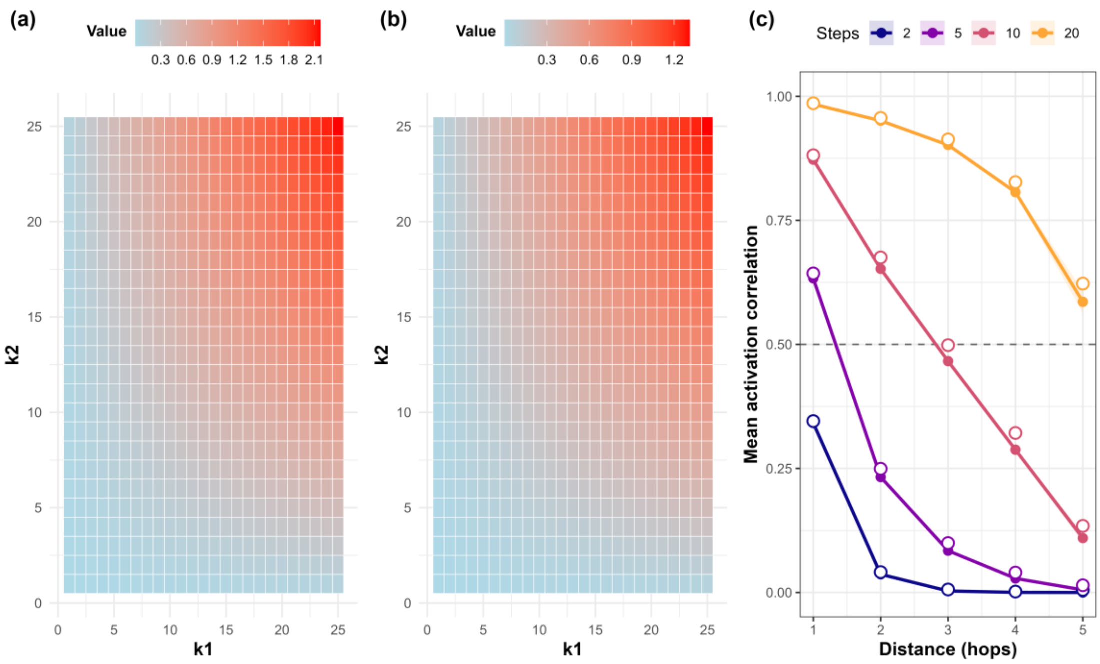
$$E\left[\#\mathrm{edges}_{i \to j|d}\right] \approx \frac{k_i k_j}{2m} \cdot \left(\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}\right)^{d-1},$$

where $\langle k \rangle$ and $\langle k^2 \rangle$ refer to the first and second moment of the degree distribution, and *m* is the number of links introduced by each new concept. Thus, nodes with higher degrees are more likely to be connected by short paths. Also, this quantity is amplified exponentially with the distance *d*, at a rate determined by the degree distribution. This indicates that even nodes with relatively few connections become increasingly reachable as the distance increases.

Panel a and b in Figure 3 shows the expected number of links between two nodes at distance 1 (e.g., directly connected) as a function of their degrees in two different networks. In both networks, the expected number of links rapidly exceeds 1 as node degree increases. The expected number of links is high even for low-degree nodes (e.g., degree 5-15). In real-world networks, the expected number of links can be even higher due to assortative mixing, where high-degree nodes tend to connect preferentially with other high-degree nodes (M. E. J. Newman, 2002).

Figure 3

Degree-Dependent Connectivity and Activation Correlations Across Network Distance



Notes. (a–b) Expected number of links connecting two nodes at d = 1 as a function of their degrees, k1 and k2. Warmer colors indicate a greater expected number of links. Panel (a) shows a network of 500 nodes with 3 links added at each time step (m = 3), whereas panel (b) shows a network of 200 nodes with 2 links added per step. In both networks, the expected number of links increases with the degrees of both nodes, such that highly connected nodes are more likely to be directly connected. (c) Activation correlation between two nodes as a function of their network distance (number of hops) at different time steps. Filled points are based on the analytical formula for the correlation, whereas hollow points are based on simulations. Activation correlations decrease with distance but increase as spreading activation proceeds, before eventually declining as activation dissipates.

Activation Correlations

The degree distribution and node distances constrain the topology over which activation spreads, giving rise to a structured covariance in node activations. Following the literature on spreading activation, we assumed that behavioral measures of a specific concept, including reaction times, Likert-scale ratings, and accuracy, are linear functions of the concept's activation (Siew, 2019; Siew et al., 2019). When activation is expressed through behavioral judgments, this covariance structure naturally translates into correlations among behavioral measures. Such correlations are central to research in social cognition. In trait inference, correlations between trait judgments can reflect how social inferences are formed (e.g., whether an attractive person is intelligent) or the strength of a stereotype is (e.g., whether a race group is warm). Regarding psychological dimensions, correlations among judgments are used to identify the core dimensions along which social cognition is organized. Thus, understanding the principles that give rise to correlations among node activations can provide a mechanistic account of why correlated patterns emerge in social judgments.

Under spreading activation, we derived the correlation between the activation levels of two nodes at time *t*, given an initial activation pattern triggered by social information (see Appendix D for the exact formula). We plotted the correlation between pairs of nodes as a function of their distance. We also evaluated the correlation at different time points to examine how it changes as spreading activation progresses.

The results revealed two properties as shown in panel c of Figure 3. First, the correlation between the activation levels of two nodes decreases with their distance—nodes that are more closely connected exhibit stronger correlations because activation can propagate between them more readily before decaying. The network topology also determines how broadly the activation can spread from an initially activated one. For a network of 5000 concepts, the small-world network with ultra-small-world property has an average distance of 2.14 (InIn5000), whereas a random network has an average path length of 8.52 (In5000). Under a moderate decay rate, activation originating from one node in the latter network has substantially less opportunity to reach distant nodes before dissipating.

Second, the overall correlation between node activations changes systematically over time, although the precise temporal pattern depends on the decay and spreading parameters. As activation propagates through the network, initially distinct activation patterns can become increasingly coordinated before the overall activation eventually decays. A concrete example is that after seeing a masculine male, both the concept of male and physical strength can be activated. As mental processing goes by, the perceiver would consider the target as both higher in male typicality and physical strength. Mechanistically, this is because the network structure acts as compression and maps the numerous pieces of information it acquires into a coherent and consistent perception. From the raw social information, the concept of male may trigger a lot of different stereotypes, such as confident and unemotional; it does not necessarily imply physical strength. However, when both cues are activated upon seeing the target, the small-world mind makes the impression more coherent to reduce representational costs. Thus, in this case, male and physical strength are compressed into a highly correlated manner. Mathematically, this mental process is similar to denoising in machine learning and has been applied in dimensionality-reduction methods such as principal component analysis. We provided more detailed psychological explanations and derivation in the next section.

Model Predictions: Rethinking Social Cognition from the Computations upon a Small-World Network

The basic analysis of the statistical properties of the small-world network model provided the foundation for deriving predictions about broader aspects of social inference. In this section, we leveraged published data and simulations to reevaluate four important topics in social cognition. For each topic, we either provided new directions for future studies or reinterpreted existing findings within our unifying computational framework. First, we discussed the longstanding concept of centrality in social inference and showed that centrality is a natural consequence of network growth and preferential attachment. This property can shape subsequent social inference and produce the emergent phenomenon of the broader halo effect. Second, we reinterpreted psychological dimensions and argued that they are emergent properties of network dynamics. The model also predicts which psychological dimensions are more likely to emerge (e.g., the content of dimensions) across contexts or domains. Third, we revisited the low- versus high-dimensionality debate introduced at the beginning of the paper. We argued that the dimensionality of social inference is context-dependent, resolving the recent debate between low- and high-dimensional accounts. Fourth, we predicted that some inferences are systematically more consistent across perceivers as a result of network structure, offering methodological implications.

On the Origin of Centrality of Concepts

As noted by Asch back in 1946, impressions of a person can be drastically altered by adding or removing certain words from a list of traits used to describe them. For example, replacing "warm" with "cold" in a list of seven traits describing a target person overturns a large number of judgments about other traits. Asch, however, argues that the centrality of these concepts is not an innate attribute but is determined by context, defined by which traits are used to describe the person and which traits are evaluated.

The investigation of centrality has been an important aspect of social cognition research across personality, trait inference, face attribution, mental states, emotions, and stereotypes (Abele et al., 2016; Cramer et al., 2012; Ekman & Friesen, 1971; S. T. Fiske et al., 2007; Sutherland & Young, 2022; Thornton & Tamir, 2020; Zebrowitz, 2017). Although many subsequent studies did not explicitly use the term of centrality, they continued to ask which judgments are important for social perception and which aspects constitute its core content. A commonly used method for identifying key contents, or dimensions, of social cognition is latent factor analysis. While we discussed on psychological dimensions in greater detail in the next section ("On the origin of psychological dimensions), we emphasized here that the centrality observed in Asch's study is fundamentally different from what factor analysis captures. One potential misunderstanding is that centrality can be quantified by the magnitude of a factor loading, with a larger loading indicating that an item is more important to a latent dimension. However, factor loadings primarily indicate how reliably an item reflects the underlying construct being measured (Cramer et al., 2012; Dunn et al., 2014). More fundamentally, factor models assume that measured items, such as perceived extraversion, are determined by underlying latent factors, thereby ruling out direct interactions among perceived concepts. This assumption cannot readily account for the directional influence between concepts observed by Asch, in which changing one concept can alter judgments of other concepts.

The divergence from Asch's original definition of centrality in subsequent studies may therefore partly reflect the lack of a clear and quantitative account of what makes a concept central. Asch developed his theory of person perception within the Gestalt tradition, emphasizing that the whole is more than the sum of its parts. However, as indicated by Rosenberg (1968), his theory did not clearly specify how the centrality of a concept is determined. We therefore need a more rigorous and quantitative approach to understanding centrality.

Prediction 1: Early Acquired Concepts Dominate Social Inference Under Spreading Activation

Our small world network model developed from complex system theory naturally aligns with Gestaltism adopted by Asch but from a modern perspective (P. W. Anderson, 1972). While we agree that context is important in shaping social inference (see "Scaling Dimensionality of Social Inference with Contexts"), we take a different view of centrality from Asch: the centrality of a trait is a structural attribute that can be estimated without referring to the "context" defined by Asch and his successors. More specifically, we argued that centrality is rooted in the network structure and reflects a concept's structural advantage, whereas contextual effects arise from the dynamic interaction among structure, processing, and incoming social information.

More specifically, we characterize a concept's centrality by its degree, reflecting the extent to which it can influence other concepts on average. Because the network is constructed sequentially during development, earlier-acquired concepts have more opportunities to attract connections and therefore exert stronger overall influence on other inferences.

To formalize and test this prediction, we built a network based on the 300 most frequent social concepts from a prior naturalistic experiment. Where participants freely generated concepts to describe the target in a video (Lu & Lin, 2025). At each time step, a new concept entered the network, with the order of entry determined by its mean age of acquisition from a published benchmark (Kuperman et al., 2012), in which participants recalled and reported the age at which they acquired each concept. Network dynamics were modeled using the recursive, resource-limited spreading-activation process introduced above. At time 0, an external cue (e.g., a face image) activates one concept in the network and assigns it an initial level of activation. We repeated this process for each of the 300 nodes. At each subsequent time step, part of the activation was retained, while the remaining activation was transmitted to neighboring nodes in proportion to connection strength and lost through decay.

Residual activation below a small threshold was treated as having returned to baseline rather than lingering indefinitely.

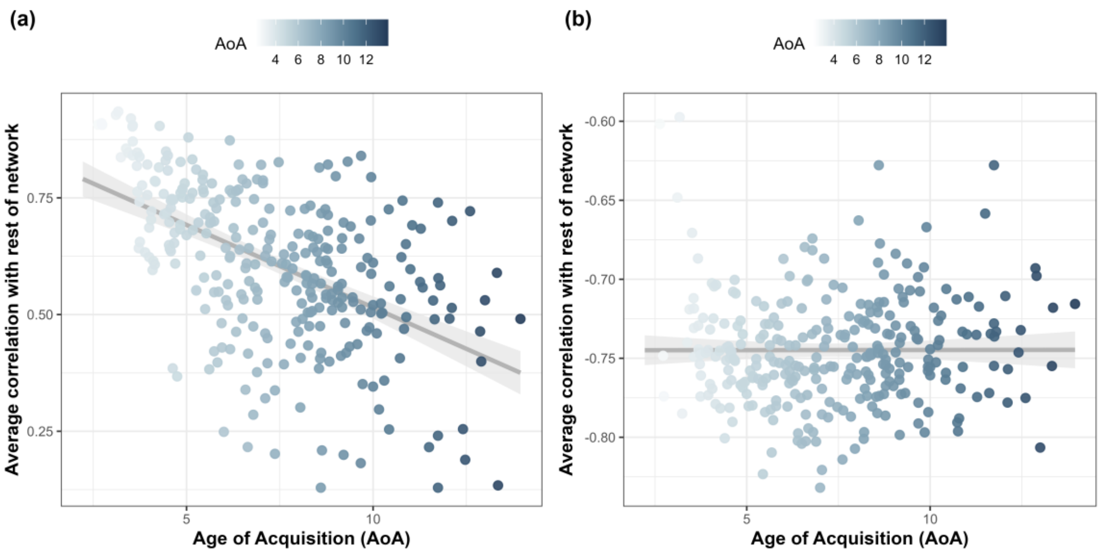
As a dependent variable indicating the extent to which the currently activated node can influence other inferences, we calculated the average correlation between its activation trajectory and those of all other nodes within a time window before most activation faded out (t = 2–15). This average correlation captures the extent to which the node can, on average, generate activation dynamics in other concepts that are synchronous with its own.

The simulation showed that average correlation with the rest of the network was significantly higher for concepts acquired earlier in life (Spearman's $\rho = -.54$, $p < .001$; linear regression: $b = -.035$, $t(298) = -10.64$, $p < .001$). Correlations ranged from .13 to .93 ($M = .61$), with the earliest-acquired concepts producing the greatest network-wide synchrony. In addition, this effect was largely explained by node degree, $b = .30$, $t(297) = 6.16$, $p < .001$. After degree was added to the regression, the age effect was largely attenuated, $b = -.014$, $t(297) = -2.94$, $p = .004$.

Importantly, the effect of centrality depends on the combination of small-world structure and limited mental activation. On the one hand, if the network lacks the small-world property and has long average paths between pairs of nodes, activation may fade before reaching a broad range of concepts or fall below the activation threshold. On the other hand, if the mind has unlimited capacity to maintain activation, centrality will no longer differentiate the influence of different concepts. Given sufficient time, any initial activation will eventually reach every other node and produce similar activation patterns across the network.

Figure 4

Average Activation Correlation with the Rest of the Network as a Function of Age of Acquisition Under Finite Activation



Notes. Panel (a) shows the finite-activation condition, in which activation is subject to a threshold and decays as it propagates through the network. Panel (b) shows the infinite-activation condition, in which activation is not subject to a threshold and does not decay as it propagates. Average correlation indicates the extent to which a node generates activation dynamics in other concepts that are synchronous with its own.

We tested this prediction by simulating infinite activation on the same network. Specifically, activation was not subject to a threshold and did not decay as it passes through the network. It turned out that average correlation with the rest of the network was not correlated with age of acquisition (Spearman's $\rho = .05$, $p = .39$; $b = 0$, $t(298) = 0.02$, $p = .98$). Also, the correlations themselves became uniformly and strongly negative ($M = –.74$, range –.83 to –.59). Because activation was modeled as a fixed-sum resource that does not decay, any gain in activation at the source node depletes the activation available to the rest of the network and introduces negative correlations.

Therefore, while centrality is a consequence of network growth and preferential attachment, its influence on other inferences is also constrained by the limited capacity of mental activation.

Reinterpreting the Beauty-is-Good Effect as Network Dynamics

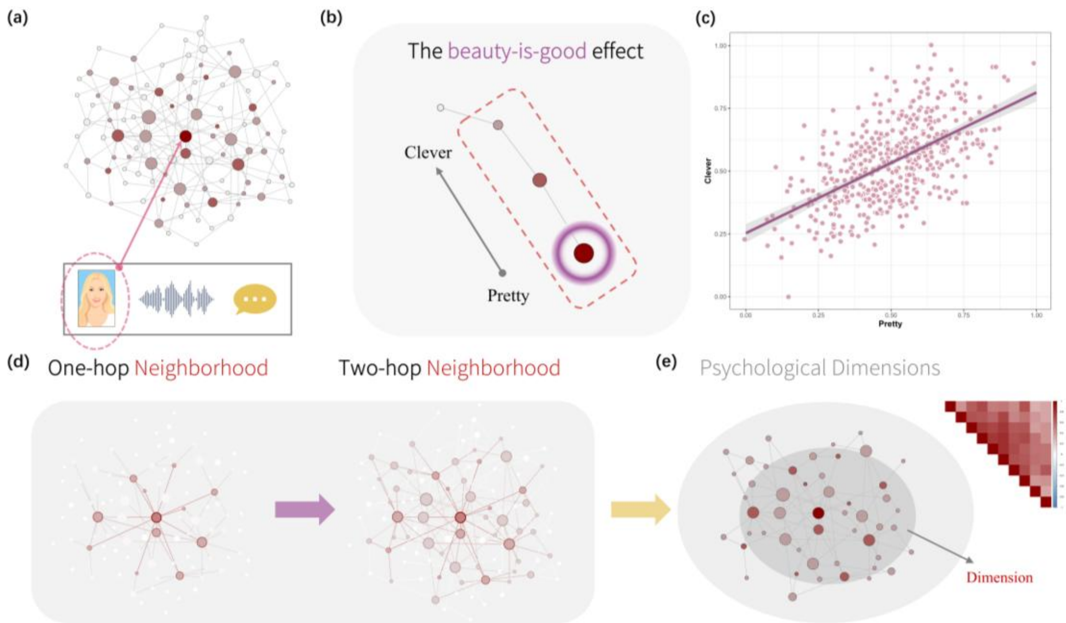
As mentioned previously, one crucial finding supporting the low-dimensional account of social cognition is the beauty-is-good effect, whereby beautiful people are systematically evaluated more positively (Zebrowitz, 2017; Zebrowitz et al., 2002; Zebrowitz & Rhodes, 2004). The small-world network model explained this well-established effect from a different perspective, building on the centrality discussed above. Specifically, the previously found high correlations between perceived attractiveness and other positive traits may arise from efficient activation spreading in the mental representation. When an attractive face is presented, the concept "pretty" is activated, as illustrate in panel A of Figure 5. Due to the short average path length induced by the ultra-small-world property, activation of "pretty" can quickly passed to other traits like "clever" and "kind" before fading out (panel B of Figure 5). In this sense, the beauty-is-good effect is a computational byproduct of network dynamics that occur spontaneously and automatically, without necessarily corresponding to a conscious psychological process that perceivers can explicitly recognize.

Consider the commonly reported correlation between perceived attractiveness and perceived intelligence (Talamas et al., 2016; Zebrowitz et al., 2002). If these two concepts are sufficiently close in the mental representation, their activation levels, and consequently their behavioral measures, should be highly correlated. As derived from our statistical analysis of the network, concepts with higher degrees are more likely to be connected through shorter paths. Indeed, perceived attractiveness and related concepts such as "pretty," "beautiful," and "attractive" often have high degrees because of their early acquisition. For example, the age of acquisition of "pretty" falls within the 5.7th percentile, making it earlier acquired than most other English concepts. Perceived intelligence, shaped by cultural norms and folk theories (Gould, 1996; Talamas et al., 2016), is also broadly connected to various concepts, contributing to its high degree.

To simulate this effect, we constructed a network of 635 concepts, targeting the association between “pretty” and “clever”. The 635 concepts were sampled from the 2926 concepts that people used to describe others in Lu and Lin (2025), taking every fourth concept in the age-of-acquisition ordering. This procedure ensured broad coverage of the age-of-acquisition range while keeping the simulation computationally tractable. We then simulated 500 stimuli, with each stimulus activating “pretty” to a different initial level to mimic the presentation of faces varying in attractiveness. In this network, the distance between “pretty” and “clever” is three, meaning that two intermediate nodes separate them along their shortest path. The activation levels of “pretty” and “clever” were strongly correlated across the 500 stimuli (panel C of Figure 5), Spearman’s $\rho = .57$, $p < .001$, reproducing the correlation characteristic of the beauty-is-good effect.

Figure 4

The Beauty-is-Good Effect and Psychological Dimensions in a Small-World Mental Representation.



Notes. (a) In typical studies, constrained stimuli, such as isolated face images varying in attractiveness, are used to elicit the beauty-is-good effect. (b) The concept related to attractiveness like “pretty” can be easily activated by social cues (Scheib et al., 1999; Tovée et al., 1999), and its activation is rapidly transmitted to neighboring nodes, such as “clever.” (c) Our simulation revealed a strong correlation between “pretty” and “clever”. (d) The initial activation of the central node (dark red) spreads to one-hop neighbors (directly connected nodes), then to two-hop neighbors (nodes that are indirectly connected through one intermediate node). Due to the short paths between nodes in a small-world network, similar activation levels efficiently spread across the entire network. (e) These similar activation patterns lead to high shared covariance across nodes (upper-right heatmap). When applied dimension reduction analyses to the activation levels of the nodes, only a few core dimensions (e.g., one indicated by the central grey region) are needed to explain most of covariance in the activations across nodes.

On the Origin of Psychological Dimensions

Whereas research on central traits and the beauty-is-good effect examines how a single concept can influence other inferences, research on psychological dimensions has investigated the interrelations among many social inferences since the 1960s. By examining the intercorrelations among social judgments, these studies have sought to identify the latent dimensions along which multiple social inferences systematically vary together. Thus, both lines of research can be traced back to Asch's work on person perception (Asch, 1946) and address a similar fundamental question: which concepts or dimensions are most important in social cognition?

Reinterpreting Psychological Dimensions in the Small-World Network Framework

Despite extensive research on psychological dimensions, this literature remains largely descriptive and leaves open how these dimensions arise in the first place and why social judgments are often highly correlated. Some researchers treat dimensions as statistical summaries of social judgments, whereas others have proposed psychological explanations, arguing that different dimensions reflect distinct underlying mechanisms (S. T. Fiske et al., 2007; Hehman et al., 2017; Oosterhof & Todorov, 2008; Sutherland & Young, 2022).

As briefly mentioned before, a prominent explanation is the evolutionary account, which proposes that social perception is organized around a few key dimensions because of their importance for survival and mating (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Sutherland & Young, 2022; Todorov et al., 2008; Zebrowitz et al., 2003; Zebrowitz & Montepare, 2008). For example, warmth and competence are considered important because they help us evaluate whether a target has good intentions and whether they are capable of acting on those intentions (S. T. Fiske et al., 2007). This evolutionary account is supported by developmental, cross-cultural, and cross-species evidence. Across cultures, studies have found some consistency in the dimensions that emerge (B. C. Jones, 2021; C. Lin et al., 2021). Within the same population, children show similar dimensions in their social perception, which can predict real-world outcomes such as election results (Antonakis & Dalgas, 2009; Cogsdill et al., 2014; Collova et al., 2019). Even macaque monkeys show implicit preferences for trustworthy-looking human faces (Costa et al., 2018).

Although the evolutionary account is plausible, a broader evolutionary perspective suggests that humans faced many recurrent challenges beyond survival and mating throughout evolutionary history (Gigerenzer & Hug, 1992; Kruger et al., 2007; Li et al., 2018). Parental care, status seeking, and mate retention, for example, represent important challenges during development (Kenrick et al., 2010; Neel et al., 2016; Schaller, 2018). It is therefore unclear whether a two-dimensional space, or even a small set of dimensions, is sufficient for navigating the social environment. Moreover, the notion of a "fundamental dimensions" sits uneasily with the domain-specificity assumption in evolutionary psychology. Evolutionary psychology proposes that knowledge is organized and applied in relation to domains that correspond to recurrent evolutionary problems, such as mating. Because acquiring and maintaining social knowledge is computationally costly, such knowledge may develop only when it becomes relevant rather than being fully specified from an early age (Boyer & Barrett, 2015).

The small-world network provides a different account of the origin of psychological dimensions. We argue that because dimensions are derived from social judgments using multivariate statistical methods, they are not necessarily psychological entities or modules that correspond to distinct functional mechanisms. Instead, they are computational byproducts of network dynamics, reflecting patterns of correlated activation among concepts within the network.

Statistically, a prerequisite for dimensions to emerge is substantial shared variance among measures like ratings of different social inferences (Barrett & Kline, 1981; Mulaik, 1987). This can be evaluated using the Measure of Sampling Adequacy (MSA), which ranges from 0 to 1 and indicates whether a given inference shares sufficient variance with other inferences to support a latent-dimension structure (Cerny & Kaiser, 1977). A low MSA indicates that an inference has relatively distinctive associations with other inferences that cannot be readily captured by the shared variance accounted for by latent dimensions.

Following Prediction 1, when a few central nodes are activated, activation propagates broadly through their neighborhoods (panel D in Figure 4), creating highly correlated activation patterns across nodes. If node activations map onto social inferences in behavioral measures, these correlations provide the shared variance necessary for dimensions to emerge (panel E in Figure 4). In addition, as described above, local preferential attachment creates densely connected communities. Activation of a node is therefore more likely to spread to other nodes within the same community, preserving differences in covariance between communities. In behavioral measures, this prevents activation patterns from collapsing into a single dimension.

Empirical findings support the view that apparent latent dimensions reflect correlated activation among nodes within a community. Both personality and person-perception research have demonstrated semantic similarities between traditional dimensions derived from factor analysis and network communities identified solely from patterns of connections (Cramer et al., 2012; Golino & Epskamp, 2017; Lu & Lin, 2025). More importantly, empirical evidence on social inferences challenges the conditional independence assumption of latent factor models (Lu & Lin, 2025). This assumption holds that, after the variance attributable to latent dimensions is accounted for, individual social inferences should no longer be correlated (Christensen & Golino, 2021; Kan et al., 2020; Van Bork et al., 2021). However, a substantial amount of unique correlation remains among social inferences, violating this assumption (Lu & Lin, 2025). Thus, dimensions can emerge from correlated activation patterns without requiring corresponding functional dimensions in the mind.

This network account of "dimensions" also provides a new perspective on developmental changes in dimensions. If dimensions emerge from communities of interconnected concepts, their development can be understood as changes in the composition and boundaries of these communities. This perspective also provides a natural explanation for developmental changes in the consensus of social judgments. Empirical findings show that, for judgments of warmth-related traits from faces varying in warmth (within-concept judgments), agreement between individuals increases with age (Cogsdill et al., 2014). As the warmth-related community becomes richer through development, with more concepts and stronger connections among them, individuals may increasingly converge on the same set of warmth-related inferences. At the same time, when warmth-related judgments are made from faces varying in competence (between-concept judgments), agreement decreases with age. From a network perspective, the communities underlying different dimensions may be less differentiated early in development, allowing activation to spread more readily across them. Consequently, competence-related cues may influence warmth judgments in a relatively consistent way across children. With development, stronger community boundaries increasingly constrain such cross-concept activation, making warmth and competence more distinct and allowing greater individual variation in between-concept judgments.

Prediction 2: Early Acquired Concepts Are More Factorizable

If psychological dimensions are indeed computational byproducts of network dynamics, a natural follow-up question is why certain dimensions, such as the "Big Two," are consistently observed across different domains of social cognition, and how the network model can account for cross-cultural, cross-species, and developmental evidence.

We argue that the answer lies in differences in degree across concepts and the general developmental trajectory of learning from the social world. Because early-acquired concepts have higher degrees, they are more capable of producing shared covariance with other nodes in the network. As a result, they should have higher MSA values and be more likely to contribute to a psychological dimension when factor analysis is performed. Based on age-of-acquisition ratings for over 30,000 English words (Kuperman et al., 2012), we found that concepts related to the previously identified "valence" dimension are acquired early in development, such as "nice" (3.95 years) and "kind" (4.89 years), whereas words related to more abstract inferences are acquired later, such as "sensitive" (8.19 years) and "extroverted" (13.35 years).

Consistent with these differences in age of acquisition, social concepts acquired earlier in development tend to correlate more strongly with other social inference concepts. As noted above, valence-related concepts are acquired early in life and tend to show stronger and more numerous correlations with other social inferences than competence-related concepts (i.e., greater explained variance when they form a factor) (Abele & Bruckmüller, 2011; Abele & Wojciszke, 2007; S. T. Fiske et al., 2002; A. L. Jones et al., 2024; Liebenow et al., 2024; A. E. Martin & Slepian, 2021). More broadly, concepts related to dimensions identified in prior research also tend to be acquired early. For example, "smart," related to the competence dimension, is estimated to be acquired at age 5.50, whereas "pretty," related to attractiveness, is acquired at around age 4.09. Their early acquisition allows these concepts to accumulate more connections and thereby gives them greater capacity to generate covariance with other concepts. Our theory therefore predicts that early-acquired concepts are more likely to form psychological dimensions.

We conducted simulations to test this prediction. We again extracted the 300 most frequently mentioned social inferences from our previous naturalistic study (Lu & Lin, 2025) and constructed networks using growth and local preferential attachment. At each step, we added one concept to the network, with the order of entry determined by empirical age-of-acquisition ratings (Kuperman et al., 2012). Specifically, we constructed two networks. In the factual network, concepts entered the network in their empirical order of acquisition. In the counterfactual network, this order was reversed, such that concepts acquired earliest in reality entered the network last.

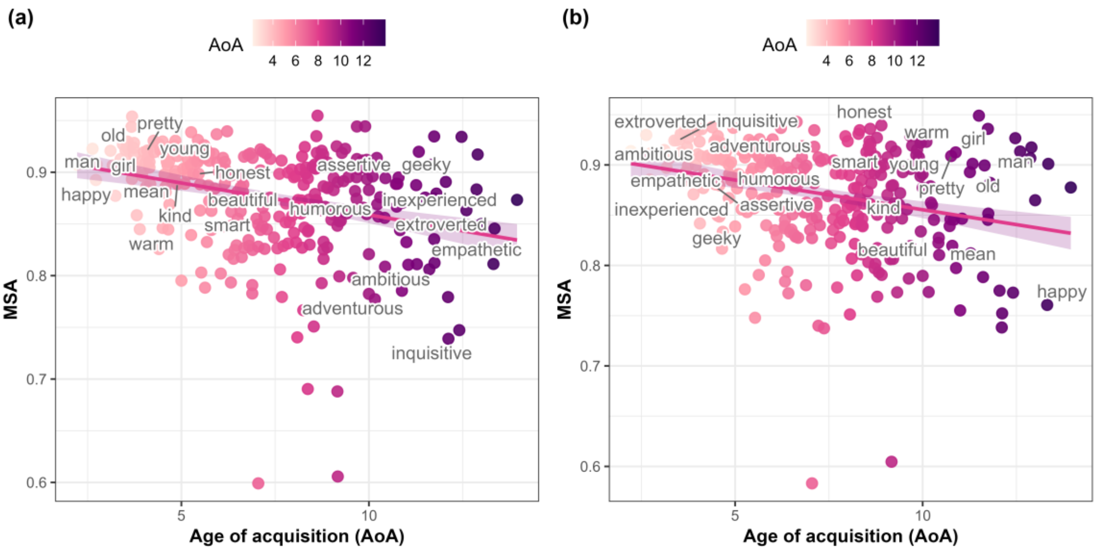
On each trial, an external cue (e.g., a piece of social information) activated one concept in the network, after which activation spread through the network. We then calculated the average activation of each concept over the readout window (steps 2–10). Repeating this process for 1,000 cues produced a stimulus-by-concept data matrix resembling those obtained in typical studies of social cognition. From the simulated data, we calculated each concept's MSA based on the correlations among activation patterns. We then used MSA as the dependent variable, indexing the extent to which a concept can be represented by a latent factor, and age of acquisition as the independent variable.

Note that we analyze MSA rather than performing a factor analysis because, under this statistical setting, which dimensions emerge also depend on which concepts are included, thereby limiting the generality of the results. For example, although "extraversion" is acquired later, constructing a measure that includes multiple items semantically similar to extraversion can force an "extraversion" dimension to emerge. A second concern is that real-world learning involves mechanisms beyond growth and preferential attachment. Concepts that frequently co-occur in the same social contexts (e.g., concepts related to mating) or appear in similar syntactic contexts (e.g., sharing the same preceding and following words) are more likely to become connected. Consequently, words such as "pretty" and "attractive" may be connected despite differences in their age of acquisition. This can be understood as a stronger form of local preferential attachment, in which preferential attachment occurs within a small subset of concepts that share semantic contexts. Both complexities are unnecessary for modeling our main argument and are therefore not incorporated in the current simulation.

In the factual network, MSA declined with age of acquisition as predicted (Spearman's $\rho = -.35$, $p < .001$; linear regression: $b = -.006$, $t(298) = --5.52$, $p < .001$, $R^2 = .09$; panel A of Figure 5). MSA ranged from .60 to .95 ($M = .88$). We then added degree to the regression to test the proposed mechanism. Degree independently predicted MSA ($b = .001$, $t(297) = 2.33$, $p = .02$) and reduced the age-of-acquisition coefficient by roughly a quarter, indicating that the factorizability of a concept is largely driven by the number of connections it has. In the counterfactual network, we replicated the same relationship. MSA decreased with counterfactual age of acquisition (Spearman's $\rho = -.31$, $p < .001$; linear regression: $b = -.006$, $t(298) = -5.08$, $p < .001$; panel B of Figure 5). However, the concepts with high MSA were reversed between the two networks. In the factual network, concepts related to gender, warmth, intelligence, and beauty had higher MSA, whereas later-acquired concepts such as ambition and inquisitiveness had lower MSA. In contrast, the counterfactual network would favor an ambition-related dimension, while concepts related to gender and warmth would become less factorizable. Thus, it is not the identity of a specific concept that determines whether it contributes to a psychological dimension; rather, its position in the developmental sequence of acquisition determines how many connections it accumulates and, consequently, how readily it generates the covariance structure from which dimensions emerge.

Figure 5

Early Acquired Concepts are more Factorizable.



Note. (a) In the factual network, concepts entered the network in their empirical age-of-acquisition order. (b) In the counterfactual network, this order was reversed. Each point represents one of the 300 social concepts, with its MSA calculated from simulated activation patterns across 1,000 external cues. Lines show the linear regression of MSA on age of acquisition. In both networks, MSA decreased with age of acquisition.

Scaling Dimensionality of Social Inference with Contexts

With the network account of psychological dimensions, we finally return to the recent debate between low- and high-dimensional accounts of social cognition. As mentioned before, empirical studies have disagreed on the number of core dimensions underlying social cognition. Because determining how many dimensions adequately describe the data is influenced by the method applied and subject to researchers' own decisions (C. Lin et al., 2025; Mondloch et al., 2023; Satchell et al., 2023), we here considered the debate over dimensionality as a disagreement about the overall complexity of social perception, rather than about statistical adequacy, to guide our analysis.

Inference as Searching for Equilibrium

The small-world network model accounts for both varying dimensionality across studies and the recurrence of a few core dimensions regardless of overall dimensionality. In this section, we first elaborated on how dimensionality varies as the same cognitive system responds to different levels of cue complexity and social information. We then argued that the mind does not passively reproduce the complexity of the social information it receives. Instead, network dynamics constitute a process of mental compression that gradually pushes activation toward a coherent and less complex perception.

First, to understand why dimensionality changes with input complexity, we revisited the beauty-is-good effect in Prediction 1. Classical paradigms studying the beauty-is-good effect have predominantly used constrained stimuli such as face images (Dion & Walster, 1972; Zebrowitz et al., 2002, 2003). When constrained stimuli are presented without auditory, semantic, or behavioral information, only a limited number of social concepts are activated. As discussed above in relation to network growth, many concepts are acquired early because they are more salient and accessible in social environments. Therefore, under constrained settings, early acquired concepts are more likely to be activated. In such cases, activation of attractiveness can quickly spread to other concepts such as intelligence without substantial decay because of the short paths in a small-world representation. This can result in high correlations between inferences. In such constrained settings, abstract social inferences (e.g., intelligence, morality) receive little direct activation from environmental inputs but can nevertheless obtain efficient activation through their activated neighbors.

Correspondingly, the small-world network suggests that this stereotype should weaken or even disappear when richer social information is available. This is because when multiple concepts are activated, activation from different concepts competes and eventually reaches an equilibrium. By equilibrium, we refer to a relatively stable state that is resistant to small perturbations. As activation spreads in a recursive and recurrent manner, strongly activated concepts pass activation to less activated nodes until the net activation between them becomes stable. In a constrained setting, this influence is more unilateral. Under rich social cues, many concepts receive unique patterns of activation and are therefore less susceptible to the influence of any single concept. Consistent with this hypothesis, when naturalistic images including context such as body posture, movement, clothing, and situational cues are used, the correlation between attractiveness and perceived warmth is relatively modest ($r = .12$; Zheng & Lin, 2024). Similarly, the attractiveness-intelligence association became statistically insignificant and approached zero ($r = -0.04$) in social inferences elicited by video clips (Lu & Lin, 2025). The same logic explains why a small number of latent dimensions cannot account for most of the variance in social inferences under naturalistic stimuli. Because visual, auditory, and behavioral cues are not necessarily strongly correlated in naturalistic settings (e.g., facial appearance and verbal articulation), more dimensions are required to account for the complex activation patterns and unique variance across the network.

From this perspective, our model predicts that the constraint on stimuli has stronger effect on dimensionality than the constraint on measurements, a distinction that has received less attention in literature. For example, prior studies have often combined constrained stimuli, such as cropped face images, with a limited set of pre-selected trait items to investigate the dimensions of trait inference. These practices have raised concerns that the identified dimensions may reflect inadequate stimulus sampling (A. L. Jones et al., 2024; Mondloch et al., 2023; Satchell et al., 2023). Indeed, many studies have relied on a priori assumptions, such as the warmth–competence framework, and selected trait words to correspond to these dimensions. This practice introduces semantic redundancy and can inflate inter-trait correlations. However, even when trait words are selected comprehensively without targeting specific dimensions, a low-dimensional structure emerges when the stimuli remain constrained (A. L. Jones et al., 2024; C. Lin et al., 2021). This finding follows from our Predictions 1 and 2: because highly central concepts can influence a broad range of other concepts, even comprehensive sampling of measures may fail to capture the more fine-grained structure of the mental representation when the stimulus itself provides limited information. Our own work (Lu & Lin, 2025) and recent evidence (Connor et al., 2024) show that when constraints on both stimuli and trait words are simultaneously relaxed, evidence for a low-dimensional representation vanishes. This supports our network mechanism in which the richness of environmental input plays a critical role in eliciting more comprehensive social inferences.

Our second argument explains why certain dimensions consistently emerge across contexts by proposing that network dynamics serve as a form of compression computation (see Appendix E for derivations). This argument is closely related to the finding in the “Activation Correlations” section that the overall correlation between activations increases with time step regardless of distance. An intuitive interpretation of this increasing correlation is that the mind coordinates different sources of social information to form a unified perception. Such compression is necessary because perception cannot be fully determined by cue-induced activation alone: concepts are often ambiguous, and their precise semantic interpretation depends on context. For example, activation of "quick" alone is less informative for understanding a person because it can refer to different meanings without additional information. When additional information is provided, "quick" can be integrated into a more holistic perception. In one case, when both "skillful" and "quick" are activated, the perceiver may interpret "quick" as referring to the person's efficiency in completing their work because they are skillful. Together, the two concepts contribute to a unifying concept of "competence." On the other hand, "quick" can be activated together with "clumsy." In this case, "quick" may be interpreted as evidence of unthoughtful behavior, and together with "clumsy," it indicates "incompetence". This pattern was reported in Experiment X of Asch (1946).

Mathematically, the network model has several characteristic "directions" along which activation patterns gradually evolve (see Appendix E). These directions, represented by the network's eigenvectors, are solely determined by the connections between concepts rather than by any particular pattern of activation induced by social cues. Because only a few central concepts have a greater capacity to influence covariation across the network, these concepts and their surrounding communities largely determine the characteristic directions, regardless of which social information is available or whether the input is sparse or rich. As interpreted in Prediction 2, these central nodes are closely related to previously identified dimensions. This reasoning therefore explains why, across studies supporting either high- or low-dimensional accounts, researchers consistently observe some dimensions, such as warmth, competence, and attractiveness.

Prediction 3: Dimensionality of Social Inference Increases with Cue Complexity

If activation correlation indeed reflects network dynamics in response to social information, then the dimensionality of perception is not a fixed property of the mind but is influenced by both the cues and the network structure. In this section, we used simulations to test (a) whether dimensionality changes with cue complexity and (b) whether, rather than merely mapping cue complexity, the mental representation compresses information over time.

We first defined our measure of dimensionality. As mentioned earlier, the exact number of dimensions identified depends on the methods used and researchers' own discretion. Thus, we calculated overall complexity of perception using effective dimensionality which does not rely on the number of dimensions:

$$D_{\text{eff}} = \frac{(\sum_a \theta_a)^2}{\sum_a \theta_a^2},$$

where θ_a stands for the eigenvalues of the activation correlation matrix. The index can be understood as the number of independent dimensions required to describe the covariance structure (Del Giudice, 2021; Pirkl et al., 2012).

We used the same 300-concept network as before. We manipulated input complexity to mimic constrained and naturalistic experimental settings. Specifically, we manipulated the number of concepts initially activated and the duration of their activation: (a) a single constrained cue (“pretty”) was activated at time 0; (b) a cue pool consisting of two concepts (“pretty” plus another sampled cue) was activated at time 0; and (c) four naturalistic settings in which all nodes were activated at time 0, but activation was sustained for 5, 10, 15, or 20 steps, representing shorter or longer exposure to social information. Activation across nodes was regulated by a graph Gaussian process over the network, such that nearby concepts received correlated input.

In each setting, we sampled 500 stimuli by sampling the initial activation of the selected cue(s). For example, in the single-cue condition, we sampled an initial activation level for the node “pretty” to mimic variation in the attractiveness of faces. We averaged the activation of each node from $t = 2$ to $t = 20$. This resulted in a 500 × 300 matrix of concept activation, with rows representing stimuli and columns representing concepts. Based on this matrix, we derived an activation correlation matrix for analyzing dimensionality.

Table 1

Dimensionality of Simulated Social Inference Across Stimulus Environments

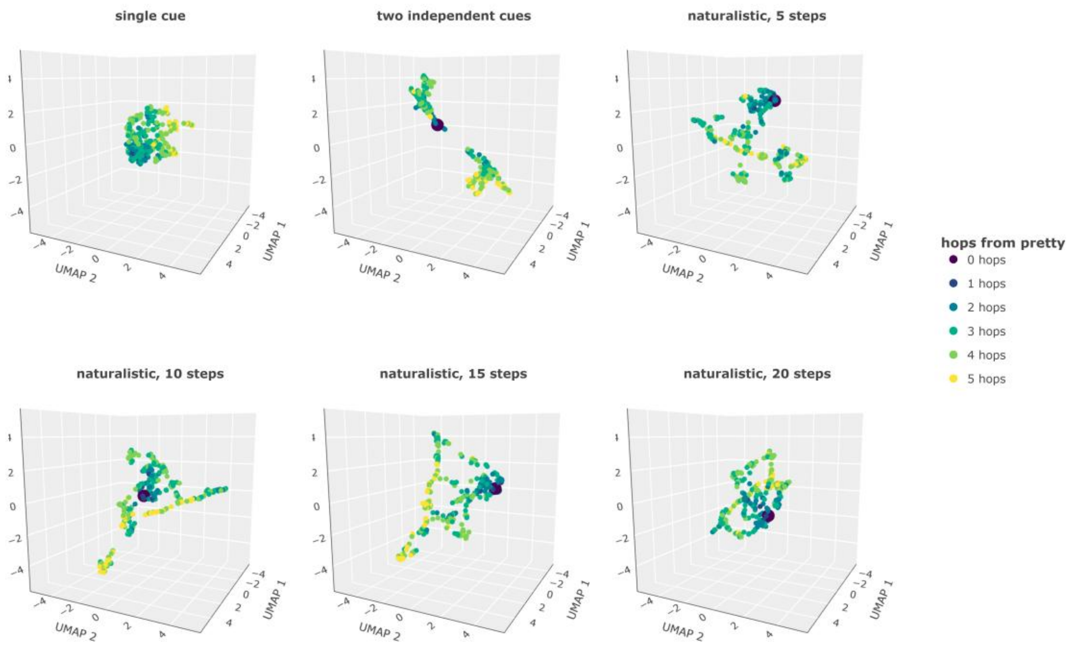
Cue complexity	First dimension (% of variance)	Effective dimensionality
Single cue	51.2	3.8
Two independent cues	35.1	5.0
Naturalistic input, 5 steps	16.2	22.1
Naturalistic input, 10 steps	14.0	27.1
Naturalistic input, 15 steps	12.8	31.1
Naturalistic input, 20 steps	9.9	39.9

Notes. The second column shows the proportion of variance explained by the first dimension, which accounts for most of the variance relative to the remaining dimensions.

We compared the effective dimensionality across the 6 conditions. The results are shown in Table 1. Under the single-cue condition, the first dimension explained 51.2% of the variance, and effective dimensionality was 3.8. This is comparable to previous studies on face-attribution dimensions using face images alone, in which the number of dimensions range from 2 to 4 and the first factor explained 30-60% of the total variance (C. Lin et al., 2021; Oosterhof & Todorov, 2008). Activating two relatively independent cues reduced the variance explained by the first dimension to 35.1% and increased effective dimensionality to 5.0. In addition, 18% of concept pairs were negatively correlated, reflecting competition between cues during activation spreading. Naturalistic input pushed dimensionality further, ranging from 22.1 after 5 steps of social information exposure to 39.9 after 20 steps. This pattern is consistent with more recent studies using videos or naturalistic images to elicit social inferences (Connor et al., 2024; Lu & Lin, 2025; Nicolas et al., 2025). The same network and dynamics therefore produced substantially different levels of dimensionality, ranging from 2 to 40, depending on cue complexity.

Figure 6

UMAP Visualization of Activation Patterns Across Cue Complexity and Spreading Hops



Notes. Each panel represents one of the six experimental conditions. Each point represents a concept, positioned according to the similarity of its activation pattern across stimuli. Colors indicate the number of spreading hops (0–5) from the initial activation. Greater spatial proximity between points indicates more similar activation patterns.

To provide an intuitive view of the covariance structure, we used Uniform Manifold Approximation and Projection (UMAP) to visualize the activation patterns of all 300 nodes in a three-dimensional space (McInnes et al., 2020). Specifically, each node was represented by its activation pattern across stimuli, and UMAP placed nodes with similar activation patterns closer together in the projected space. As shown in Figure 6, activation patterns under a single cue clustered together across concepts, indicating high correlations and low dimensionality. With two relatively independent cues, two clusters emerged, each primarily reflecting activation driven by one cue. Under naturalistic input, activation patterns were more dispersed but remained concentrated around several regions of the space rather than being distributed arbitrarily. This structure reflects the influence of central nodes and network communities. Moreover, nodes activated most strongly by the initial concept “pretty” were its network neighbors, recovering a pattern consistent with the beauty-is-good stereotype and the broader halo effect.

The results also supported the idea that network dynamics compress social information. To illustrate this process, we analyzed effective dimensionality at different time points in the naturalistic five-step environment rather than averaging across time points as above. After the stimulus stopped at step 5, we tracked dimensionality from step 5 to 25. As shown in Table 2, effective dimensionality decreased from 68.4 to 18.0, while the variance explained by the first component increased from 5.2% to 19.0%. Thus, continued activation spreading reduced dimensionality even when the stimulus provided abundant information at the beginning. In real-world perception, however, computational constraints would limit the number of iterations participants could perform before forming a perceptual judgment. These constraints may prevent activation from converging to a low-dimensional state within the time window of an experiment. Nevertheless, the gradual compression of activation toward directions characterized by network eigenvectors helps explain why studies using different designs consistently observe some common core dimensions.

Together, the simulation results suggest that dimensionality depends on both the amount of social information available and the time available for activation to spread.

Table 2

Effect of Spreading Duration on Dimensionality Under Naturalistic Input Sustained for 5 Steps

Time point	**First component (% of variance)**	**Effective dimensionality**
t = 5	5.2	68.4
t = 10	8.9	41.9
t = 20	16.2	22.1
t = 25	19.0	18.0

Note. All time points use the identical naturalistic setting reported in Table 1 for a 5-step input duration. When t = 20, the results reproduce the naturalistic input of 5 steps of Table 1.

Emergence of Interrater Consistency and the Target Effect

While we have primarily discussed the generative process of social inference, a closely related topic concerns the reliability of these inferences across perceivers. That is, when different perceivers encounter the same person, to what extent do they converge on similar inferences? This question is particularly important because social-cognitive findings are often based on judgments aggregated across perceivers. Aggregation can reduce idiosyncratic noise, but it is meaningful only to the extent that perceivers share some common structure in their inferences.

Empirically, evidence for such cross-perceiver reliability has been reported (Kramer et al., 2018; Little & Perrett, 2007; Oosterhof & Todorov, 2008). This finding generalized to inference about individuals from different racial groups made by raters from different ethnic or cultural groups (Albright et al., 1997; Zebrowitz et al., 1993). Even children between 3 and 10 years of age show significant consensus on well-established psychological dimensions such as trustworthiness and competence (Cogsdill et al., 2014). At the same time, interrater agreement varies systematically across different inferences. For example, judgments of gender typicality and age show higher agreement than judgments of trustworthiness and dominance (Kramer et al., 2018).

Prediction 4: Higher Interrater Consistency is Underpinned by High Degree

Despite empirical evidence for high reliability in social inference, at least for some concepts, there is no unifying theory explaining how such reliability emerges from computations based on social knowledge and mental processing. Here, we use reliability specifically to refer to *interrater consistency*: the extent to which different perceivers converge on similar inferences when evaluating the same target.

Understanding interrater consistency of social inference carries both theoretical and methodological significance. Theoretically, consensus can reveal whether an inference is driven primarily by bottom-up information from the target—the target effect—or by top-down cognitive projections from the perceiver—the perceiver effect. This distinction sheds light on how perceivers derive social inferences and whether different inferences may arise from distinct causal processes (Hehman et al., 2017).

Methodologically, trait-specific interrater consistency should inform experimental design and model selection. When test length is constrained, concepts with higher interrater consistency within the same domain may be preferable. High-consensus traits can also be estimated efficiently with conventional models, whereas low-consensus traits may require cross-classified multilevel models or Social Relations Models to account for perceiver and dyadic effects, including target-by-perceiver interactions (Back & Kenny, 2010).

The small-world network model provides precise predictions about which inferences should have higher interrater consistency. Because interrater consistency reflects the proportion of observed variance attributable to shared target information rather than idiosyncratic perceiver effects, the higher interrater consistency also implies a stronger target effect. Specifically, our model predicts that concepts acquired at earlier ages are more reliable and associated with larger target effects. The reason is two-fold. First, as elaborated in the section on growth, concepts that are more accessible and directly grounded in environmental cues are acquired earlier, providing the foundation upon which more abstract concepts are subsequently constructed. Consequently, the evidence underlying these inferences is more directly accessible, requiring fewer multi-step and potentially noisy inferential processes. This makes the resulting inferences more consistent across perceivers.

The second reason is more fundamental and can be derived from the network structure. The basic idea is that interrater consistency of a social inference should be higher when the idiosyncratic variation in its activation is lower. For example, if an emotion concept such as “happiness” is acquired early and becomes connected to many other concepts such as “warm,” “joyful,” and “smiling,” activation of one or more of these neighboring concepts is sufficient to infer the extent of “happiness” in response to a target. Because the concept with high degree can be reached through multiple related sources, its activation is less dependent on any particular social cue and should therefore be relatively consistent across perceivers. In contrast, an emotion concept acquired later has a lower degree, and the particular concepts to which it becomes connected may differ substantially across perceivers given the large number of existing concepts in the network. For example, one perceiver may associate “nostalgia” primarily with “childhood” and “memory,” whereas another may connect it more strongly to “home” and “loss.” Consequently, the activation of such a concept depends more heavily on a perceiver-specific subset of cues, making the inference more sensitive to individual differences in network structure and experience.

To illustrate the relationship between interrater consistency and degree, we can represent the activation of a concept analogously to a measurement model based on classical test theory (Mair, 2018). Under this analogy, the inference associated with a social concept like “happiness” can be represented as the sum of inputs from its *n* neighboring concepts:

$$X = \sum_{i=1}^{n} X_i,$$

where

$$X_i = T_i + e_i,$$

Here X_i is the input from neighbor *i* of the focal concept “happiness”, and T_i is its signal (e.g., reliable information relevant to the focal concept). The error term e_i consists of differences in network structure across perceivers, spreading activation parameters (e.g., decay), and how individual perceivers respond to the same social cue. In more general cases, X_i may also include direct environmental input, which corresponds to the first reason discussed above. This extension, however, does not affect the main argument here.

For mathematical convenience, we used Cronbach's α to characterize the reliability of the focal concept's activation:

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n}\sigma_i^2}{\sigma_X^2}\right).$$

Here, *n* is the number of neighbors of the focal concept, analogous to the number of test items in a psychological scale, σ_i^2 is the variance of input from neighbor *i*, and σ_x^2 is the variance of the focal node's activation X. Assuming equal variance across inputs such that $\sum_i \sigma_i^2 = n\sigma_0^2$ and a common correlation between neighboring inputs, one can derive the derivative of α with respect to *n*:

$$\frac{d\alpha}{dn} = \frac{\rho(1-\rho)}{[1+(n-1)\rho]^2}.$$

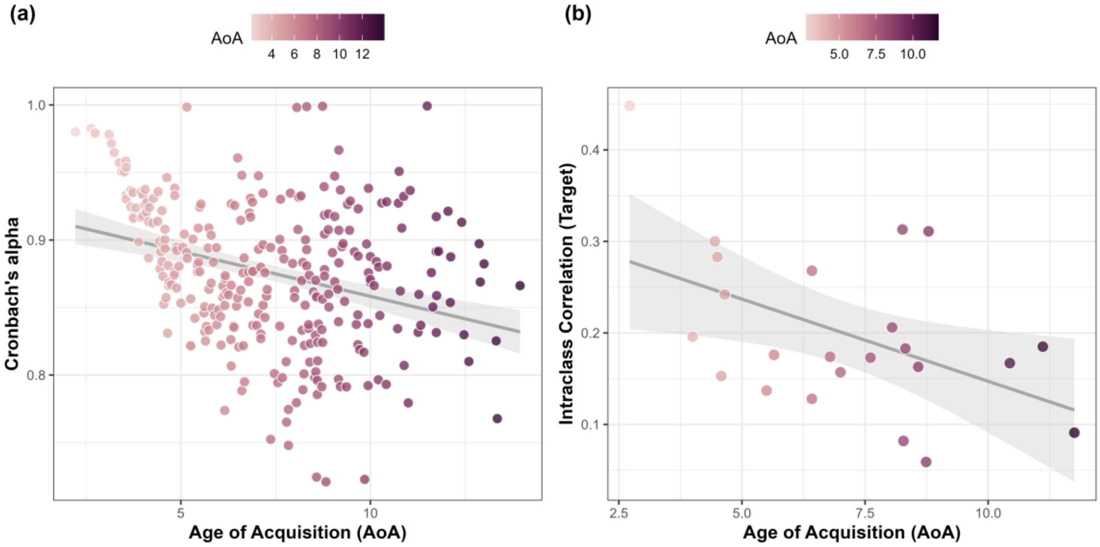
Here, ρ is the common correlation between inputs from neighbors. Because ρ lies typically between 0 and 1, as neighboring activations are on average positively correlated as in Figure 3, the derivative is positive, indicating that reliability increases with degree *n*. Therefore, earlier acquired concepts with higher degrees are structurally more reliable because their activation can be supported by information from more channels, allowing random noise and idiosyncratic variation associated with any individual cue to be averaged out.

To simulate this effect, we estimated Cronbach's α for each of the 300 social concepts in the network used in previous sections. Importantly, the specification above should be understood as a mechanistic explanation rather than as the procedure by which reliability is actually computed. To generate variation across perceivers, we simulated 50 independently grown networks using the same set of 300 concepts. We modeled each cue (e.g., a face image) by assigning initial activation to five nodes in the network. We repeated this procedure 500 times, randomly sampling the initial activation levels of the five cues on each trial and recording the activation profile of each node across time points. For each node in the network, we computed Cronbach's α from its cues-by-perceiver matrix, treating perceivers as items and cues as observations as in prior studies (Kramer et al., 2018).

Results showed that Cronbach's α varied across concepts (α ranged from .75 to .99, $M = .87$) and was predicted by age of acquisition: earlier acquired concepts showed significantly greater consistency across perceivers (Spearman $\rho = -.36$, $p < .001$; linear regression $b = -.007$, $t(298) = -6.40$, $p < .001$, $R^2 = .12$). The relationship was shown in panel A of Figure 7. The effect of age of acquisition became insignificant once node degree entered the regression (degree: $b = .004$, $t(297) = 6.49$, $p < .001$; age: $b = -.0009$, $t(297) = -0.64$, $p = .53$), providing supports for the proposed mechanism.

Figure 7

Interrater Reliability of Social Inferences as a Function of Age of Acquisition



Notes. Panel A shows the relationship between age of acquisition and Cronbach's α estimated from the simulated networks. Each point represents one concept, with higher α indicating greater consistency across independently grown networks. Panel B shows the empirical relationship between age of acquisition and target intraclass correlation coefficients for 23 social traits reported by Hehman et al.

While we used Cronbach's α for illustrative purposes, it is not an ideal measure of interrater consistency because it is jointly determined by the number of perceivers and the average interrater correlation (T.-Y. Lin et al., 2026; Müller & Büttner, 1994). More recent studies have instead adopted intraclass correlations (ICCs), which more directly quantify the extent to which ratings converge across perceivers. Because our model predicts greater consistency for early acquired concepts across perceivers, we expect the same pattern to generalize to ICC. To provide additional support for this prediction, we obtained target ICCs for 23 traits from Hehman et al. (2017):

$$ICC(A,1) = \frac{\sigma^2_{Target}}{\sigma^2_{Target} + \sigma^2_{Perceiver} + \sigma^2_e},$$

where σ^2_{Target} refers the variance attributable to the targets (between-target variance), $\sigma^2_{Perceiver}$ is the variance attributable to the raters/perceivers (between-perceiver variance), and σ^2_e is the residual error variance. As predicted and shown in panel B of Figure 7, target ICC was negatively associated with age of acquisition, *b* = –.018, *t*(21) = -2.49, *p* = .02, even after controlling for the frequency of the word in the English corpus, *b* = –.033, *t*(21) = -2.84, *p* = .009. Using the same bootstrap correlation procedure as in Hehman et al., we likewise found a strong negative association between target ICC and age of acquisition, bootstrapped *r* = -.48, *p* = .021.

Synonyms and Item Selection

Previous studies have typically treated the reliability of social inferences as a property of the measurement instrument and therefore something that must be determined empirically. In contrast, we argue that reliability can, at least in part, be predicted from the structure of the conceptual network. This provides a principled basis for selecting among concepts that capture similar or overlapping aspects of social perception. Importantly, even synonymous concepts are unlikely to occupy exactly the same position in the network. Different words are typically used in somewhat different contexts and experiences, leading them to acquire different connections and degrees despite substantial overlap in meaning.

Consider *beautiful* and *attractive*. Both capture positive evaluations of appearance, but they occupy different positions in the conceptual network. *Beautiful* is acquired earlier (mean age = 5.72) than *attractive* (mean age = 8.58). By the time *attractive* enters the network, many more concepts have already been established and are available as potential connections, including *smart*, *beautiful*, and *kind*. Moreover, because *attractive* is used across a broader range of contexts, the particular concepts to which it becomes connected may vary more across perceivers. Its meaning and activation may therefore be more susceptible to individual differences. In contrast, the earlier acquired *beautiful* should have a more consistently established network position and therefore yield more consistent inferences.

This distinction has a direct methodological implication. When several concepts capture broadly similar content, researchers need not treat them as interchangeable measurement items. Instead, concepts with greater centrality or earlier acquisition may provide more consistent measurements without requiring additional items. Conversely, later-acquired and sparsely connected concepts may contain greater perceiver-specific variation and therefore require models that explicitly account for target, perceiver, and dyadic effects rather than treating disagreement simply as measurement error (Back & Kenny, 2010).

Moving Towards a Resource-Rational, Process-Oriented Account of Social Cognition

Taken together, our analyses provided a formal theory of how social concepts are mentally represented and how social inferences can arise under the limited cognitive resources. Resonating with the opening of our study, we reconciled the low versus high dimensional debate by showing that they are the special cases of network activation under sparse and rich social information. Our theory has implications for broad domains of social cognition at both theoretical and meta-theoretical levels. Theoretically, it illustrates how to advance our understanding of social cognition from a modeling approach while incorporating the cognitive constraints that limit the space of plausible underlying processes. At the meta-theoretical level, it addresses how researchers' own biases toward simplicity should be overcome to achieve more ecologically valid theories. We next expand these implications in two sections.

Navigating the Complex World with a Scalable Mind

Broadly speaking, the low- versus high-dimensional debate resonates with a long-standing question in psychology concerning the nature of human rationality. Early work, particularly following the rise of axiomatic rationality in economics and decision theory, treated deviations from formal principles (e.g., axioms in expected utility theory) as evidence of cognitive bias (Frederick et al., 2002; Kahneman, 1992; Tversky & Kahneman, 1974). Similar assumptions permeated social cognition, where low-dimensional models were often interpreted as evidence that people rely on mental shortcuts (S. Fiske, 2018; Olivola & Todorov, 2010). This view, however, has been substantially revised in various domains. Research on ecological rationality demonstrated that humans exploit environmental structure through frugal heuristics well-adapted to statistical regularities in their environments (Gigerenzer, 2008, 2025; Todd & Gigerenzer, 2007). Although such heuristics may fail when environmental conditions change, their success under typical conditions reflects adaptive intelligence rather than irrationality.

Our network account of social inference fits naturally with this theoretical development in related fields, and with recent resource-rational analysis framework characterizing cognition as adaptive given constraints on time, computation, and representational capacity (Lieder & Griffiths, 2020). In this sense, learning efficient representations is striking a balance between frugality that fits in with our limited cognitive capacity and complexity that accommodates rich statistical regularities in complex social environments. Simply attributing our mind and perception as a low or high dimensional structure dismisses the flexibility of our inference processes.

With this challenge, we emphasize that theory development of social cognition relies on finding scalable mental representations. In machine learning, a system is scalable if its computational costs increase linearly or sub-linearly, rather than polynomially or exponentially, with input complexity (Buchanan et al., 2025). Oversimplified models cannot accommodate environmental richness or change, whereas complex models are costly to learn and use. Effective representations must balance complexity and efficiency. The small-world network model satisfies this requirement. First, network structure expands with development rather than remaining fixed, allowing the representation to scale. Second, the mathematical form of preferential attachment equivalent to representing information with minimal cognitive cost, giving rise to clustering and "rich-get-richer" dynamics (Dasgupta & Griffiths, 2022). This property resonates with proverbs cited at the beginning of this paper in Tao Te Ching: The way of heaven does not induce the rich-get-richer mechanisms; it is humans, other creatures, and society that rely on this due to the limited resource in our mind, body, and the social world. Third, the limited activation the mind can maintain ensures only a proportion of learned concepts are involved in making inferences, allowing flexible responding to external cues.

Together, our theory shows that incorporating resource-rational structure and processing upon the structure allows explaining a wide range of findings in social cognition. Future refinement or related theory developments could expand our approach by adding more fine-grained cognitive constraints to account for other empirical findings.

Moving Beyond the Low-Dimensional Thinking

Beyond the theoretical concerns, the persistence of low-dimensional conclusions in the field may reflect deeper cognitive and cultural biases toward simplicity.

Humans have a strong motivation to understand and find meaning in their experiences (A. R. Cohen et al., 1955). Given the inherent complexity of human psychology, this motivation compels researchers to create mental placeholders that establish feelings of certainty and understanding (Strevens, 2000)—the low-dimensional models serve as good candidates. Beyond these cognitive tendencies, cultural norms in social science research favoring verbal descriptions over mathematical representations when communicating research findings also contribute to the popularity of low-dimensional solutions (Watts, 2017). Results that are easier to interpret, communicate, remember, and reproduce tend to persist during social transmission; whereas, results that are more difficult to interpret, communicate, or remember are more often criticized and distrusted (Jolly & Chang, 2019). Indeed, transmission of social knowledge (stereotypes) over generations of participants compresses knowledge representations from high- to low-dimensional (D. Martin et al., 2014), which corresponds to our argument that spreading activation on a small-world network is equivalent to mental compression. Such processes may operate in scientific production and transmission, suggesting that our innate cognitive tendencies may systematically distort the facts and favor simplified models that are more readily communicated even if they are less accurate.

Moving beyond this low-dimensional "wishful thinking" is key to advancing an ecologically valid understanding of social cognition (Jolly & Chang, 2019). We believe that progress requires two complementary approaches: incorporating the complexity of real-world social interactions into experimental designs and developing quantitative mental models that make precise predictions about behavior. Specifically, we recommend reconceptualizing experiments as simplified models of real-world interactions rather than artificial manipulations of isolated variables (Smaldino, 2023). Social cognitive processes such as person perception typically unfold across time rather than within a single time slice or glimpse of a face. Allowing for a reasonable level of complexity can therefore reveal the flexible inferences that emerge from the mind as a complex system. When this complexity is overly constrained, however, our observations may capture only a limited subset of the processes through which the mind operates. In addition, to better understand cognitive processes, simply manipulating social cues may not be sufficient, as the way cues are integrated and combined can largely influence behavioral outcomes (Goel et al., 2024; Ma et al., 2023; Zaki, 2013). We recommend developing explicit mental representation models that can inform cognitive processes in naturalistic contexts, such as how complex environmental inputs are processed and transformed into behavioral responses.

Conclusions

In summary, we propose a computational framework for understanding the dynamic and context-dependent nature of social cognition. Specifically, we argue that mental representations of social concepts are structured as a small-world network that develops through growth and preferential attachment. Through mathematical derivations and simulations, we provide a framework for reconciling the long-standing debate between low- and high-dimensional accounts of social cognition. Crucially, we apply the model to four central topics in social cognition, offering new explanations of existing findings, concepts, and theories from a resource-rational perspective. By jointly specifying mental structure and the processes operating over that structure, our model moves beyond static dimensional accounts and provides a mechanistic account of how social inferences dynamically emerge and depend on context. This theoretical foundation provides a basis for a more ecologically valid understanding of how people acquire, organize, and use social knowledge while adapting to the complexity of the real world.

Appendix A

The Small World Property

The "small world" property, popularized as six degrees of separation, was first explored in sociology rather than in the context of psychological representations. It refers to the striking observation that any two people on Earth are likely connected through only a handful of acquaintances. In network science terms, this means that the path length between two randomly chosen nodes in a network is surprisingly short.

Why "surprisingly"? Much of our intuition about distance comes from physical space, where connections grow slowly with size. Imagine people arranged along a line: to reach someone at the far end, you would need to pass through nearly everyone in between. Even if people are arranged in a grid, the number of steps between two individuals still scales with the square root of the total population. By contrast, in a random network where each person has on average $\langle k \rangle$ acquaintances, the number of reachable nodes grows rapidly with distance: approximately $\langle k \rangle$ nodes at distance 1, $\langle k \rangle^2$ nodes at distance 2, $\langle k \rangle^3$ at distance 3, and in general $\langle k \rangle^d$ at distance d. As a result, the average separation between individuals increases only logarithmically with population size: $d \sim \ln N / \ln\langle k \rangle$. For example, if the global social network contains 8 billion people ($N = 8 \times 10^9$) and each has 500 acquaintances ($\langle k \rangle = 500$), the expected average distance is just over four steps (i.e., four acquaintances).

Beyond random networks, the small-world property also appears in scale-free networks. These networks contain hubs, nodes with disproportionately many connections, that further shorten distances. A familiar example is the airline network: major hubs such as Atlanta or Dubai dramatically reduce the number of flights needed to travel between distant airports. In mathematical terms, hubs can reduce the scaling of average distance from *lnN* to *ln(lnN)*, a phenomenon known as the "ultra-small" property. Yet the effect of hubs is not always dominant. In some cases, hubs are not influential enough to substantially alter path lengths, and scale-free networks effectively approximate the behavior of random networks.

Appendix B

Deriving the Degree Distribution Based on Degree Dynamics

The structure of the network, characterized by its degree distribution, is implied by how edges are accumulated. Based on preferential attachment, we have the probability of any single link of the new node arriving node *i* at time *t*:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} = \frac{k_i}{2m(t-1)},$$

where k_i is the current degree of node *i*, m refers to the number of links brought by the new node. Because the new node adds m new links to the network, this formula can be multiplied by m and consider as the rate of change in the degree of node *i*:

$$\frac{dk_i}{dt} = m\,\Pi(k_i) = m\frac{k_i}{2m(t-1)} = \frac{k_i}{2(t-1)}.$$

When t is large, we can simplify the t-1 as t and obtain:

$$\frac{dk_i}{k_i} = \frac{1}{2}\frac{dt}{t}.$$

Integrating each side gives:

$$\int_m^{k_i(t)} \frac{dk}{k} = \frac{1}{2}\int_{t_i}^{t} \frac{ds}{s}$$

Solving this equation and exponentiating both sides yields the formula in the main text:

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta},$$

where β equals 0.5, represents the dynamical exponent that governs how fast a concept's degree accumulates. This suggests that the degree of a node is controlled by the time it enters the network. Thus, the probability of nodes smaller than a certain value is now equivalent to the proportion of nodes entering the network later than a specific time point:

$$P(k_i(t) < k) = P\left(t_i > \left(\frac{m}{k}\right)^{\frac{1}{\beta}} t\right)$$

Thus:

$$P\left(t_i > \left(\frac{m}{k}\right)^{\frac{1}{\beta}} t\right) = \int_{\left(\frac{m}{k}\right)^{\frac{1}{\beta}} t}^{t} \frac{ds}{t} = 1 - \left(\frac{m}{k}\right)^{\frac{1}{\beta}}$$

Given this cumulative density function, the probability density function is:

$$P(k) = \frac{d}{dk}\left[1 - m^{\frac{1}{\beta}} k^{-\frac{1}{\beta}}\right] = \frac{1}{\beta} m^{\frac{1}{\beta}} k^{-\left(\frac{1}{\beta}+1\right)}$$

which is a power law in k, with exponent:

$$\gamma = \frac{1}{\beta} + 1 = 3,$$

Therefore:

$$P(k) \propto k^{-\gamma}$$

Appendix C

Deriving the Linking Probability Between Nodes

We derive the probability that two nodes are connected at a distance d. This section provides an overview; more detailed descriptions can be found in the supplementary materials.

We first show that, in large networks, this probability approximates the expected number of links between the two nodes. Thus, the derivation converts to deriving the expected number of links between two nodes. In the network, each node is assigned with a number of stubs (half-edges), resulting in a total of $\sum k_i = 2m$ stubs across the network. At each step, two stubs are selected uniformly at random and connected to form an edge. This process continues until all stubs are paired and the network is fully constructed.

We start with a simple case. Let k_i and k_j denote the degrees of two nodes, the probability that at least one edge exists between them is:

$$p(edge \geq 1) = 1 - (1 - \frac{k_j}{N\langle k \rangle})^{k_i}$$

Based on the Taylor series, we have:

$$p(edge \geq 1) = 1 - (1 - k_i \frac{k_j}{N\langle k \rangle} + (\frac{k_i k_j}{N\langle k \rangle})^2 \frac{k_i(k_i - 1)}{2!} + O(x^3)) \approx \frac{k_i k_j}{2m}$$

where N is the number of nodes, $\langle k \rangle$ is the average degree, $\langle k^2 \rangle$ is the second moment of the degree distribution. For large N, the second-order term vanishes quickly. We can also derive the expected number of direct links (e.g., distance = 1) between the two nodes $E[\#\text{edges}_{i \to j | d=1}]$ using the linearity of expectation:

$$E[\#\text{edges}_{i \to j | d=1}] = \sum_{a=1}^{k_i} P\,(\text{stub } a \to \text{node } j) = k_i \frac{k_j}{2m - 1} \approx \frac{k_i k_j}{2m}$$

where P is the probability of a specific stab from node i that links to node j. Thus, in the limit of large m, the probability that an edge exists between two nodes converges to the expected number of edges between them.

Next, we derive the expected number of links between two nodes at an arbitrary distance d. As a first step, we derive the expected number of links between two nodes with degrees k_i and k_j that are separated by a distance d. The simplest nontrivial case occurs when the two nodes are separated by two steps (i.e., distance $d = 2$).

This quantity is the product of the number of edges between i and l times that of between l and j:

$$
\begin{aligned}
E[\#\text{edges}_{i\to j|d=2}] &= \sum_{l} \frac{k_i k_l}{2m} \frac{(k_l-1)k_j}{2m-1} \\
&= \frac{k_i k_j}{2m} \frac{\sum_l k_l\,(k_l-1)}{2m-1} \\
&= p_{ij} \frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle}
\end{aligned}
$$

The derivation is similar to a 3-step pathway, in which node i passes through node l_1 and l_2 before reaching j:

$$
\begin{aligned}
E[\#\text{edges}_{i\to j|d=3}] &= \sum_{l_1}\sum_{l_2} \frac{k_i k_{l_1}}{2m} \frac{(k_{l_1}-1)k_{l_2}}{2m} \frac{(k_{l_2}-1)k_j}{2m} \\
&= \frac{k_i k_j}{2m} \left(\sum_{l_1} \frac{k_{l_1}(k_{l_1}-1)}{2m}\right)\left(\sum_{l_2} \frac{k_{l_2}(k_{l_2}-1)}{2m}\right) \\
&= p_{ij}\left(\frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle}\right)^2
\end{aligned}
$$

Generalizing the same logic to a d-step pathway, we have:

$$
\mathbb{E}[\#edges_{i\to j|d}] \approx p_{ij} \cdot \left(\frac{\langle k^2\rangle - \langle k\rangle}{\langle k\rangle}\right)^{d-1}
$$

This suggests that the expectation, or the probability of linking, increases exponentially with distance and converges rapidly to 1, leaving no node pairs separated beyond this distance.

Appendix D

Spreading Activation and the Correlation Between Activations

We defined and implemented spreading activation as follows: at each time step, a concept retains a proportion *r* of its current activation; it passes the remaining proportion 1-*r* to its neighbors, divided among them in proportion to the strengths of the connections leaving it; and a proportion δ of all activation is lost due to decay.

The connections between concepts are characterized by a matrix *P*, whose rows are normalized so the sum is one. The entry in row *i* and column *j* is the share of the activation leaving concept *i* that arrives at concept *j*. Let n denote the number of concepts and *I* the n × n identity matrix. The vector of activations $\mathbf{x}_t$ representing the activation level of all concepts at step *t* is:

$$
\mathbf{x}_t = (1-\delta)[(rI + (1-r)P^T)]\mathbf{x}_{t-1},
$$

where the *i*th element of $\mathbf{x}_t$, $x_{t,i}$, is the activation of concept i at step t. The matrix in brackets describes how activation is redistributed at each step. Its diagonal gives the proportion of activation that a concept retains, while its off-diagonal entries give the proportions passed to its neighbors. The term δ then determines the proportion of activation that remains after decay. We can therefore define the matrix for spreading one step as *M*:

$$M = [(1-\delta)(rI + (1-r)P^T)].$$

The matrix for spreading *t* steps is then the *t*-th power of *M*:

$$\Phi = [(1-\delta)(rI + (1-r)P^T)]^t.$$

Substitute this into the formula of activation at step *t*, we have:

$$\mathbf{x}_t = \Phi \mathbf{x}_0$$

Typically, the correlation between activation levels of two concepts *i* and *j* is calculated across multiple stimuli at a given time point *t*. To simplify the derivation, we assume that the initial activation induced by a set of stimuli is a random vector with diagonal covariance. The variance of initial activation of concept *l* is $\sigma^2_{0,l}$, which is the *l*th diagonal entry of that covariance. Based on the properties of covariance, we can derive the covariance between $x_{t,i}$ and $x_{t,j}$:

$$\mathrm{Cov}\left(x_{t,i}, x_{t,j}\right) = \sum_{l=1}^{n} \Phi_{il}\, \Phi_{jl} \sigma_{0,l}^2.$$

The correlation then follows by normalizing by the corresponding variances:

$$\rho\left(x_{t,i}, x_{t,j}\right) = \frac{\sum_{l=1}^{n} \Phi_{il}\, \Phi_{jl} \sigma_{0,l}^2}{\sqrt{\left(\sum_{l=1}^{n} \Phi_{il}^2\, \sigma_{0,l}^2\right)\left(\sum_{l=1}^{n} \Phi_{jl}^2\, \sigma_{0,l}^2\right)}}$$

Thus, given the network structure and how concepts are initially activated by social cues, we can derive the activation correlation between any pair of concepts at any time point.

Appendix E

Spreading Activation as Compression

To understand why spreading activation operates as compression, we showed that as time goes on, activation gradually converges toward a low-dimensional structure that is uniquely determined by the network structure, regardless of the type and complexity of social information that initially activates the network.

Recall that the matrix form of spreading activation running for one step is:

$$M = [(1-\delta)(rI + (1-r)P^T)]$$

As an n-by-n matrix, *M* has n eigenvalues that can be ranked in descending order:

$$\lambda_1 > \lambda_2 > \cdots > \lambda_n$$

The activation at time t is:

$$x_t = M^t x_0$$

Using the eigendecomposition of *M*, we have:

$$M^t = QM^tQ^{-1} = \sum_i {\lambda_i}^t v_i \, {u_i}^T,$$

where, v_i and u_i represent right and left eigenvectors, respectively. Now, the activation at time *t* can be written as:

$$x_t = {\lambda_1}^t \left(v_1 {u_1}^T + \sum_{i \geq 2} \left(\frac{\lambda_i}{\lambda_1}\right)^t v_i \, {u_i}^T \right) x_0.$$

The term $\left(\frac{\lambda_i}{\lambda_1}\right)^t$ approaches 0 as time goes by when $|\lambda_i| < |\lambda_1|$. When *t* becomes sufficiently large, the activation therefore approaches:

$x_t \approx {\lambda_1}^t x_0,$

which is a projection onto v_1. Thus, in the limit, the activation becomes effectively one-dimensional. In more realistic settings, limited cognitive capacity would prevent spreading from continuing for a large number of steps. In this case, activation is characterized primarily by a few leading eigenvectors rather than by the full network structure, regardless of the initial activation pattern (e.g., which social cues are provided). This provides a mathematical explanation for why studies using different designs can nevertheless identify a small number of common dimensions using dimensionality-reduction methods.

The specifications above also show that the eigenvectors are determined by the network structure *P* and the parameters governing spreading activation, rather than by any particular initial activation pattern. These eigenvectors can be understood in terms of the community structure of the network. For a right eigenvector, we have:

$$Mv_i = \lambda_i v_i.$$

Because the main structural component of M is P^T, we can simplify the argument by considering:

$$P^T v_i = \lambda_i v_i.$$

The *i*th row of P^T contains the weights connecting node *i* to its neighbors. Therefore, the *i*th element of $P^T v_i$ is a weighted sum of the values of v_i among the neighbors of node *i*. Consequently, when nodes within a community are closely connected, their values on the same eigenvector tend to be similar and often have the same sign. In this sense, an eigenvector defines a direction along which strongly connected nodes tend to move together.

Because the *n* eigenvectors are orthogonal, different eigenvectors would capture different patterns of variation in the network. In a simple case, if a network contains only two clearly separated communities, satisfying the orthogonality constraint that their inner product is zero naturally gives rise to two distinct directions, one represented by $(1,1,1,1,0,0,0,0)^T$ and another by $(0,0,0,0,1,1,1,1)^T$, with each direction corresponding primarily to one community. In more realistic networks, communities are not perfectly separated, so the corresponding eigenvectors can include both positive and negative values and show partial overlap across communities.

References

Abele, A. E., & Bruckmüller, S. (2011). The bigger one of the "Big Two"? Preferential processing of communal information. *Journal of Experimental Social Psychology*, *47*(5), 935–948. https://doi.org/10.1016/j.jesp.2011.03.028

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the Fundamental Content Dimensions: Agency with Competence and Assertiveness—Communion with Warmth and Morality. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01810

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Albright, L., Malloy, T. E., Dong, Q., Kenny, D. A., Fang, X., Winquist, L., & Yu, D. (1997). Cross-cultural consensus in personality judgments. *Journal of Personality and Social Psychology*, *72*(3), 558–569. https://doi.org/10.1037/0022-3514.72.3.558

Aldous, D. J. (1985). Exchangeability and related topics. In *Lecture Notes in Mathematics* (pp. 1–198). Springer Berlin Heidelberg. https://doi.org/10.1007/bfb0099421

Amodio, D. M. (2025). A learning and memory account of impression formation and updating. *Nature Reviews Psychology*. https://doi.org/10.1038/s44159-025-00445-x

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*(3), 409.

Anderson, P. W. (1972). More Is Different. *Science*, *177*(4047), 393–396.

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science*, *323*(5918), 1183–1183. https://doi.org/10/cqp5nz

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258. https://doi.org/10.1037/h0055756

Back, M. D., & Kenny, D. A. (2010). The Social Relations Model: How to Understand Dyadic Processes: The Social Relations Model. *Social and Personality Psychology Compass*, *4*(10), 855–870. https://doi.org/10.1111/j.1751-9004.2010.00303.x

Barabási, A.-L. (2009). Scale-Free Networks: A Decade and Beyond. *Science*, *325*(5939), 412–413. https://doi.org/10.1126/science.1173299

Barabasi, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*, 509–512.

Barrett, P. T., & Kline, P. (1981). The observation to variable ratio in factor analysis. *Personality Study & Group Behaviour*, *1*(1), 23–33.

Bertotti, M. L., & Modanese, G. (2019). The configuration model for Barabasi-Albert networks. *Applied Network Science*, *4*(1). https://doi.org/10.1007/s41109-019-0152-1

Bhui, R., Lai, L., & Gershman, S. J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, *41*, 15–21.

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., Van Borkulo, C. D., Van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 58. https://doi.org/10.1038/s43586-021-00055-w

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Boyer, P., & Barrett, H. C. (2015). Intuitive ontologies and domain specificity. *Handbook of Evolutionary Psychology (2nd Ed., Pp. 168–180). Hoboken, NJ: Wiley*.

Brooks, J. A., Kim, L., Opara, M., Keltner, D., Fang, X., Monroy, M., Corona, R., Tzirakis, P., Baird, A., Metrick, J., Taddesse, N., Zegeye, K., & Cowen, A. S. (2024). Deep learning reveals what facial expressions mean to people in different cultures. *iScience*, *27*(3), 109175. https://doi.org/10.1016/j.isci.2024.109175

Buchanan, S., Pai, D., Wang, P., & Ma, Y. (2025). *Learning Deep Representations of Data Distributions*. Online.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, *38*(4), 476–506. https://doi.org/10.1037/h0054116

Cerny, B. A., & Kaiser, H. F. (1977). A Study Of A Measure Of Sampling Adequacy For Factor-Analytic Correlation Matrices. *Multivariate Behavioral Research*, *12*(1), 43–47. https://doi.org/10.1207/s15327906mbr1201_3

Christensen, A. P., Garrido, L. E., Guerra-Peña, K., & Golino, H. (2023). Comparing community detection algorithms in psychometric networks: A Monte Carlo simulation. *Behavior Research Methods*, *56*(3), 1485–1505. https://doi.org/10.3758/s13428-023-02106-4

Christensen, A. P., & Golino, H. (2021). On the equivalency of factor and network loadings. *Behavior Research Methods*, *53*(4), 1563–1580. https://doi.org/10.3758/s13428-020-01500-6

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring Character From Faces: A Developmental Study. *Psychological Science*, *25*(5), 1132–1139. https://doi.org/10.1177/0956797614523297

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology*, *51*(2), 291.

Cohen, R., & Havlin, S. (2003). Scale-Free Networks Are Ultrasmall. *Physical Review Letters*, *90*(5), 058701. https://doi.org/10.1103/PhysRevLett.90.058701

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240–247.

Collova, J. R., Sutherland, C. A. M., & Rhodes, G. (2019). Testing the functional basis of first impressions: Dimensions for children's faces are not the same as for adults' faces. *Journal of Personality and Social Psychology*, *117*(5), 900–924. https://doi.org/10.1037/pspa0000167

Connor, P., Nicolas, G., Antonoplis, S., & Koch, A. (2024). Unconstrained Descriptions of Facebook Profile Pictures Support High-Dimensional Models of Impression Formation. *Personality and Social Psychology Bulletin*, *0*(0), 1–17.

Costa, M., Gomez, A., Barat, E., Lio, G., Duhamel, J.-R., & Sirigu, A. (2018). Implicit preference for human trustworthy faces in macaque monkeys. *Nature Communications*, *9*(1), 4529. https://doi.org/10.1038/s41467-018-06987-4

Courtney, O. T., & Bianconi, G. (2016). Generalized network structures: The configuration model and the canonical ensemble of simplicial complexes. *Physical Review E*, *93*(6). https://doi.org/10.1103/physreve.93.062311

Cowen, A. S., Brooks, J. A., Prasad, G., Tanaka, M., Kamitani, Y., Kirilyuk, V., Somandepalli, K., Jou, B., Schroff, F., Adam, H., Sauter, D., Fang, X., Manokara, K., Tzirakis, P., Oh, M., & Keltner, D. (2024). How emotion is experienced and expressed in multiple cultures: A large-scale experiment across North America, Europe, and Japan. *Frontiers in Psychology*, *15*, 1350631. https://doi.org/10.3389/fpsyg.2024.1350631

Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences*, *117*(4), 1924–1934. https://doi.org/10.1073/pnas.1910704117

Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, *114*(38). https://doi.org/10.1073/pnas.1702247114

Cowen, A. S., & Keltner, D. (2019). What the Face Displays: Mapping 28 Emotions Conveyed by Naturalistic Expression. *The American Psychologist*, *75*(3), 349. https://doi.org/10.1037/amp0000488

Cowen, A. S., Manokara, K., Fang, X., Sauter, D., Brooks, J. A., & Keltner, D. (2021). *Facial movements have over twenty dimensions of perceived meaning that are only partially captured with traditional methods*. PsyArXiv. https://doi.org/10.31234/osf.io/hc93t

Cramer, A. O. J., Van Der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., & Borsboom, D. (2012). Dimensions of Normal Personality as Networks in Search of Equilibrium: You Can't like Parties if you Don't like People. *European Journal of Personality*, *26*(4), 414–431. https://doi.org/10.1002/per.1866

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 61–149). Elsevier. https://doi.org/10.1016/S0065-2601(07)00002-0

Dasgupta, I., & Griffiths, T. L. (2022). Clustering and the efficient use of cognitive resources. *Journal of Mathematical Psychology*, *109*, 102675. https://doi.org/10.1016/j.jmp.2022.102675

Davidsen, J., Ebel, H., & Bornholdt, S. (2002). Emergence of a Small World from Local Interactions: Modeling Acquaintance Networks. *Physical Review Letters*, *88*(12), 128701. https://doi.org/10.1103/PhysRevLett.88.128701

De Jager, S., Coetzee, N., & Coetzee, V. (2018). Facial Adiposity, Attractiveness, and Health: A Review. *Frontiers in Psychology*, *9*, 2562. https://doi.org/10.3389/fpsyg.2018.02562

de Sola Pool, I., & Kochen, M. (1978). Contacts and influence. *Social Networks*, *1*(1), 5–51.

Del Giudice, M. (2021). Effective Dimensionality: A Tutorial. *Multivariate Behavioral Research*, *56*(3), 527–542. https://doi.org/10.1080/00273171.2020.1743631

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283.

Dion, K., & Walster, E. (1972). WHAT IS BEAUTIFUL IS GOOD. *Journal of Personality and Social Psychology*, *24*(3), 285–290. https://doi.org/10.1037/h0033731

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, (105), 399–412.

Eagjy, A. H., Ashmore, R. D., & Makhijani, M. G. (1991). What Is Beautiful Is Good, But. . .: A Meta-Anatytic Review of Research on the Physical Attractiveness Stereotype. *Psychological Bulletin*, *110*(1), 109–128.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129. https://doi.org/10.1037/h0030377

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10/gc32sh

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, *111*(2), 304–341. https://doi.org/10.1037/0033-2909.111.2.304

Fiske, S. (2018). *Social cognition: Selected works of Susan Fiske*. Routledge. https://www.google.com/books?hl=zh-CN&lr=&id=OlkPEAAAQBAJ&oi=fnd&pg=PP1&dq=Fiske+shortcut&ots=5eXz86MeEE&sig=2iVxXGHrI_CowJwnnCYXkSoz1Jg

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., 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*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In *Advances in experimental social psychology* (Vol. 23, pp. 1–74). Elsevier.

Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, *659*, 1–44. https://doi.org/10.1016/j.physrep.2016.09.002

Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279. https://doi.org/10.1037/a0022327

Freeman, J. B., & Lin, C. (2025). A high-dimensional model of social impressions. *Trends in Cognitive Sciences*, S136466132500110X. https://doi.org/10.1016/j.tics.2025.04.011

Freeman, J. B., Stolier, R. M., & Brooks, J. A. (2020). Dynamic interactive theory as a domain-general account of social perception. In *Advances in Experimental Social Psychology* (Vol. 61, pp. 237–287). Elsevier. https://doi.org/10.1016/bs.aesp.2019.09.005

Gendron, M., Roberson, D., Van Der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion*, *14*(2), 251–262. https://doi.org/10.1037/a0036052

Gershman, S. J., & Blei, D. M. (2012). A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology*, *56*(1), 1–12. https://doi.org/10.1016/j.jmp.2011.08.004

Gershman, S. J., & Cikara, M. (2023). Structure learning principles of stereotype change. *Psychonomic Bulletin & Review*, *30*(4), 1273–1293. https://doi.org/10.3758/s13423-023-02252-y

Giacomin, M., & Rule, N. O. (2020). How static facial cues relate to real-world leaders' success: A review and meta-analysis. *European Review of Social Psychology*, *31*(1), 120–148. https://doi.org/10.1080/10463283.2020.1771935

Gigerenzer, G. (2008). Why Heuristics Work. *Perspectives on Psychological Science*, *3*(1), 20–29. https://doi.org/10.1111/j.1745-6916.2008.00058.x

Gigerenzer, G. (2025). The rationality wars: A personal reflection. *Behavioural Public Policy*, *9*(3), 495–515. https://doi.org/10.1017/bpp.2024.51

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*(2), 127–171. https://doi.org/10.1016/0010-0277(92)90060-U

Goel, S., Jara-Ettinger, J., Ong, D. C., & Gendron, M. (2024). Face and context integration in emotion inference is limited and variable across categories and individuals. *Nature Communications*, *15*(1), 2443. https://doi.org/10.1038/s41467-024-46670-5

Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229. https://doi.org/10.1037//0022-3514.59.6.1216

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, *12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Gould, S. J. (1996). *Mismeasure of man*. WW Norton & company. https://www.google.com/books?hl=zh-CN&lr=&id=WTtTiG4eda0C&oi=fnd&pg=PA19&dq=the+mismeasure+of+man&ots=881UNMy7Yr&sig=xEBGelO4MpuxQMVKAwm2dtKKYRU

Guida, M., & Maria, F. (2007). Topology of the Italian airport network: A scale-free small-world network with a fractal structure? *Chaos, Solitons & Fractals*, *31*(3), 527–536. https://doi.org/10.1016/j.chaos.2006.02.007

Haines, N., Kvam, P. D., Irving, L., Smith, C. T., Beauchaine, T. P., Pitt, M. A., Ahn, W.-Y., & Turner, B. M. (2025). A tutorial on using generative models to advance psychological science: Lessons from the reliability paradox. *Psychological Methods*. https://doi.org/10.1037/met0000674

Hehman, E., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology*, *113*(4), 513–529. https://doi.org/10.1037/pspa0000090

Jackson, L. A., Hunter, J. E., & Hodge, C. N. (1995). Physical Attractiveness and Intellectual Competence: A Meta-Analytic Review. *Social Psychology Quarterly*, *58*(2), 108–122. https://doi.org/10.2307/2787149

Jolly, E., & Chang, L. J. (2019). The Flatland Fallacy: Moving Beyond Low–Dimensional Thinking. *Topics in Cognitive Science*, *11*(2), 433–454. https://doi.org/10.1111/tops.12404

Jones, A. L., Shiramizu, V., & Jones, B. C. (2024). Decoding the language of first impressions: Comparing models of first impressions of faces derived from free-text descriptions and trait ratings. *British Journal of Psychology*, bjop.12717. https://doi.org/10.1111/bjop.12717

Jones, B. C. (2021). To which world regions does the valence–dominance model of social perception apply? *Nature Human Behaviour*, *5*. https://doi.org/10.1038/s41562-020-01007-2

Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes*, *51*(2), 296–312. https://doi.org/10.1016/0749-5978(92)90015-Y

Kan, K.-J., De Jonge, H., Van Der Maas, H. L. J., Levine, S. Z., & Epskamp, S. (2020). How to Compare Psychometric Factor and Network Models. *Journal of Intelligence*, *8*(4), 35. https://doi.org/10.3390/jintelligence8040035

Kanazawa, S. (2004). Why beautiful people are more intelligent. *Intelligence*, *32*(3), 227–243. https://doi.org/10.1016/j.intell.2004.03.003

Kanazawa, S. (2011). Intelligence and physical attractiveness. *Intelligence*, *39*(1), 7–14. https://doi.org/10.1016/j.intell.2010.11.003

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*(3), 307–321. https://doi.org/10.1111/j.1467-7687.2007.00585.x

Kenrick, D. T., Neuberg, S. L., Griskevicius, V., Becker, D. V., & Schaller, M. (2010). Goal-Driven Cognition and Functional Behavior: The Fundamental-Motives Framework. *Current Directions in Psychological Science*, *19*(1), 63–67. https://doi.org/10.1177/0963721409359281

Kramer, R. S. S., Mileva, M., & Ritchie, K. L. (2018). Inter-rater agreement in trait judgements from faces. *PLOS ONE*, *13*(8), e0202655. https://doi.org/10.1371/journal.pone.0202655

Kritschgau, J., Kaiser, D., Alvarado Rodriguez, O., Amburg, I., Bolkema, J., Grubb, T., Lan, F., Maleki, S., Chodrow, P., & Kay, B. (2024). Community detection in hypergraphs via mutual information maximization. *Scientific Reports*, *14*(1), 6933. https://doi.org/10.1038/s41598-024-55934-5

Kruger, D. J., Wang, X. T., & Wilke, A. (2007). Towards the Development of an Evolutionarily Valid Domain-Specific Risk-Taking Scale. *Evolutionary Psychology*, *5*(3), 147470490700500. https://doi.org/10.1177/147470490700500306

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390–423. https://doi.org/10.1037/0033-2909.126.3.390

Li, N. P., van Vugt, M., & Colarelli, S. M. (2018). The Evolutionary Mismatch Hypothesis: Implications for Psychological Science. *Current Directions in Psychological Science*, *27*(1), 38–44. https://doi.org/10/gc2tjg

Liebenow, H. A., Boucher, K. L., & Cassidy, B. S. (2024). Trait inferences from the “big two” produce gendered expectations of facial features. *Journal of Experimental Social Psychology*, *111*, 104585. https://doi.org/10.1016/j.jesp.2023.104585

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1. https://doi.org/10.1017/S0140525X1900061X

Lin, C., Han, Y., Keles, U., Xu, Y., Lu, J., Zheng, R., & Adolphs, R. (2025). Investigating the Structure of Emotion: Tools, Pitfalls and Recommendations. *Affective Science*. https://doi.org/10.1007/s42761-025-00337-6

Lin, C., Keles, U., & Adolphs, R. (2021). Four dimensions characterize attributions from faces using a representative set of English trait words. *Nature Communications*, *12*(1), 5168. https://doi.org/10.1038/s41467-021-25500-y

Lin, T.-Y., Tuerlinckx, F., & Vanbelle, S. (2026). *An intraclass correlation coefficient is a correlation: A tutorial*.

Little, A. C., & Perrett, D. I. (2007). Using composite images to assess accuracy in personality attribution to faces. *British Journal of Psychology*, *98*(1), 111–126. https://doi.org/10.1348/000712606X109648

Lu, J., & Lin, C. (2025). Network models reveal high-dimensional social inferences in naturalistic settings beyond latent construct models. *Communications Psychology*, *3*(1). https://doi.org/10.1038/s44271-025-00275-w

Ma, W. J., Kording, K. P., & Goldreich, D. (2023). *Bayesian Models of Perception and Action: An Introduction*. The MIT Press.

Mair, P. (2018). *Modern Psychometrics with R*. Springer International Publishing. https://doi.org/10.1007/978-3-319-93177-7

Martin, A. E., & Slepian, M. L. (2021). The Primacy of Gender: Gendered Cognition Underlies the Big Two Dimensions of Social Cognition. *Perspectives on Psychological Science*, *16*(6), 1143–1158. https://doi.org/10.1177/1745691620904961

Martin, D., Hutchison, J., Slessor, G., Urquhart, J., Cunningham, S. J., & Smith, K. (2014). The Spontaneous Formation of Stereotypes via Cumulative Cultural Evolution. *Psychological Science*, *25*(9), 1777–1786. https://doi.org/10.1177/0956797614541129

Maslov, S., Sneppen, K., Alon, U., Bornholdt, S., & Schuster, H. G. (2003). *Handbook of graphs and networks*. Wiley-VCH New York:

McInnes, L., Healy, J., & Melville, J. (2020). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction* (arXiv:1802.03426). arXiv. http://arxiv.org/abs/1802.03426

Milgram, S. (1967). The small world problem. *Psychology Today*, *2*(1), 60–67.

Mitchem, D. G., Zietsch, B. P., Wright, M. J., Martin, N. G., Hewitt, J. K., & Keller, M. C. (2015). No relationship between intelligence and facial attractiveness in a large, genetically informative sample. *Evolution and Human Behavior*, *36*(3), 240–247. https://doi.org/10.1016/j.evolhumbehav.2014.11.009

Mondloch, C. J., Twele, A. C., & Thierry, S. M. (2023). We need to move beyond rating scales, white faces and adult perceivers: Invited Commentary on Sutherland & Young (2022), understanding trait impressions from faces.

British Journal of Psychology, *114*(2), 504–507. https://doi.org/10.1111/bjop.12619

Mulaik, S. A. (1987). A Brief History of the Philosophical Foundations of Exploratory Factor Analysis. *Multivariate Behavioral Research*, *22*(3), 267–305. https://doi.org/10.1207/s15327906mbr2203_3

Müller, R., & Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine*, *13*(23–24), 2465–2476. https://doi.org/10.1002/sim.4780132310

Neel, R., Kenrick, D. T., White, A. E., & Neuberg, S. L. (2016). Individual differences in fundamental social motives. *Journal of Personality and Social Psychology*, *110*(6), 887–907. https://doi.org/10.1037/pspp0000068

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*(6), 622–631. https://doi.org/10/c2ckwb

Newman, M. (2018). *Networks*. Oxford university press. https://www.google.com/books?hl=zh-CN&lr=&id=YdZjDwAAQBAJ&oi=fnd&pg=PP1&dq=network+science+newman&ots=V0KY-Hmbuy&sig=MrBgIt_UTMeIPY9Ouww6sf2SzW0

Newman, M. E. J. (2002). Assortative Mixing in Networks. *Physical Review Letters*, *89*(20). https://doi.org/10.1103/physrevlett.89.208701

Nicolas, G., Bai, X., & Fiske, S. T. (2022a). A spontaneous stereotype content model: Taxonomy, properties, and prediction. *Journal of Personality and*

Social Psychology, *123*(6), 1243–1263. https://doi.org/10.1037/pspa0000312

Nicolas, G., Bai, X., & Fiske, S. T. (2022b). A spontaneous stereotype content model: Taxonomy, properties, and prediction. *Journal of Personality and Social Psychology*, *123*(6), 1243–1263. https://doi.org/10.1037/pspa0000312

Nicolas, G., Uddenberg, S., & Todorov, A. (2025). Spontaneous Content of Impressions of Naturalistic Face Photographs. *Social Cognition*, *43*(2), 114–143. https://doi.org/10.1521/soco.2025.43.2.114

Olivola, C. Y., & Todorov, A. (2010). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*(2), 315–324. https://doi.org/10.1016/j.jesp.2009.12.002

On the evolution of random graphs. (1960). *The Structure and Dynamics of Networks, 5*, *17*, 38–82. https://doi.org/10.1515/9781400841356.38

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Piaget, J. (2013). *The construction of reality in the child*. Routledge. https://www.taylorfrancis.com/books/mono/10.4324/9781315009650/construction-reality-child-jean-piaget

Pirkl, R. J., Remley, K. A., & Patane, C. S. L. (2012). Reverberation Chamber Measurement Correlation. *IEEE Transactions on Electromagnetic Compatibility*, *54*(3), 533–545. https://doi.org/10.1109/TEMC.2011.2166964

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*(4), 283–294. https://doi.org/10.1037/h0026086

Satchell, L., Jaeger, B., Jones, A. L., Lopez, B., & Schild, C. (2023). Beyond reliability in first impressions research: Considering validity and the need to “mix it up with folks.” *Social Psychological Bulletin*, *18*, 1–21. https://doi.org/10.32872/spb.10211

Schaller, M. (2018). The Parental Care Motivational System and Why It Matters (for Everyone). *Current Directions in Psychological Science*, *27*(5), 295–301. https://doi.org/10.1177/0963721418767873

Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1431), 1913–1917. https://doi.org/10.1098/rspb.1999.0866

Sheehan, C. M., & Hamermesh, D. S. (2024). Looks and longevity: Do prettier people live longer? *Social Science & Medicine*, *354*, 117076. https://doi.org/10.1016/j.socscimed.2024.117076

Siew, C. S. Q. (2019). spreadr: An R package to simulate spreading activation in a network. *Behavior Research Methods*, *51*(2), 910–929. https://doi.org/10.3758/s13428-018-1186-5

Siew, C. S. Q., Wulff, D. U., Beckage, N. M., Kenett, Y. N., & Meštrović, A. (2019). Cognitive Network Science: A Review of Research on Cognition through the Lens of Network Representations, Processes, and Dynamics. *Complexity*, *2019*, 1–24. https://doi.org/10.1155/2019/2108423

Smaldino, P. E. (2023). *Modeling social behavior: Mathematical and agent-based models of social dynamics and cultural evolution*. Princeton University Press.

Stamkou, E., Keltner, D., Corona, R., Aksoy, E., & Cowen, A. S. (2024). Emotional palette: A computational mapping of aesthetic experiences evoked by visual art. *Scientific Reports*, *14*(1), 19932. https://doi.org/10.1038/s41598-024-69686-9

Steyvers, M., & Tenenbaum, J. B. (2005). The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, *29*(1), 41–78. https://doi.org/10.1207/s15516709cog2901_3

Stolier, R. M., Hehman, E., & Freeman, J. B. (2018). A Dynamic Structure of Social Trait Space. *Trends in Cognitive Sciences*, *22*(3), 197–200. https://doi.org/10.1016/j.tics.2017.12.003

Stolier, R. M., Hehman, E., & Freeman, J. B. (2020). Trait knowledge forms a common structure across social cognition. *Nature Human Behaviour*, *4*(4), 361–371. https://doi.org/10.1038/s41562-019-0800-6

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition*, *74*(2), 149–175.

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Sutherland, C. A. M., & Young, A. W. (2022). Understanding trait impressions from faces. *British Journal of Psychology*, *113*(4), 1056–1078. https://doi.org/10.1111/bjop.12583

Talamas, S. N., Mavor, K. I., & Perrett, D. I. (2016). Blinded by Beauty: Attractiveness Bias and Accurate Perceptions of Academic Performance. *PLOS ONE*, *11*(2), e0148284. https://doi.org/10.1371/journal.pone.0148284

Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human Nature*, *4*(3), 237–269. https://doi.org/10.1007/BF02692201

Thornton, M. A., & Tamir, D. I. (2020). People represent mental states in terms of rationality, social impact, and valence: Validating the 3d Mind Model. *Cortex*, *125*, 44–59. https://doi.org/10.1016/j.cortex.2019.12.012

Todd, P. M., & Gigerenzer, G. (2007). Environments That Make Us Smart: Ecological Rationality. *Current Directions in Psychological Science*, *16*(3), 167–171. https://doi.org/10.1111/j.1467-8721.2007.00497.x

Todorov, A. (2005). Inferences of Competence from Faces Predict Election Outcomes. *Science*, *308*(5728), 1623–1626. https://doi.org/10.1126/science.1110589

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460. https://doi.org/10.1016/j.tics.2008.10.001

Tovée, M. J., Maisey, D. S., Emery, J. L., & Cornelissen, P. L. (1999). Visual cues to female physical attractiveness. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. (world). https://doi.org/10.1098/rspb.1999.0624

Traag, V. A., Waltman, L., & Van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, *9*(1), 5233. https://doi.org/10.1038/s41598-019-41695-z

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, New Series*, *185*(4157), 1124–1131.

Uleman, J. S., Adil Saribay, S., & Gonzalez, C. M. (2008). Spontaneous Inferences, Implicit Impressions, and Implicit Theories. *Annual Review of Psychology*, *59*(1), 329–360. https://doi.org/10.1146/annurev.psych.59.103006.093707

Van Bork, R., Rhemtulla, M., Waldorp, L. J., Kruis, J., Rezvanifar, S., & Borsboom, D. (2021). Latent Variable Models and Networks: Statistical Equivalence and Testability. *Multivariate Behavioral Research*, *56*(2), 175–198. https://doi.org/10.1080/00273171.2019.1672515

Watts, D. J. (2017). Should social science be more solution-oriented? *Nature Human Behaviour*, *1*(1), 0015.

Weeden, J., & Sabini, J. (2005). Physical Attractiveness and Health in Western Societies: A Review. *Psychological Bulletin*, *131*(5), 635–653. https://doi.org/10.1037/0033-2909.131.5.635

Zaki, J. (2013). Cue Integration: A Common Framework for Social Cognition and Physical Perception. *Perspectives on Psychological Science*, *8*(3), 296–312. https://doi.org/10.1177/1745691613475454

Zebrowitz, L. A. (2017). First Impressions From Faces. *Current Directions in Psychological Science*, *26*(3), 237–242. https://doi.org/10.1177/0963721416683996

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*(3), 194–215. https://doi.org/10.1207/S15327957PSPR0703_01

Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking Smart and Looking Good: Facial Cues to Intelligence and their Origins.

Personality and Social Psychology Bulletin, *28*(2), 238–249. https://doi.org/10.1177/0146167202282009

Zebrowitz, L. A., & Montepare, J. M. (2008). Social Psychological Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*, *2*(3), 1497–1517. https://doi.org/10.1111/j.1751-9004.2008.00109.x

Zebrowitz, L. A., Montepare, J. M., & Lee, H. K. (1993). They don’t all look alike: Individuated impressions of other racial groups. *Journal of Personality and Social Psychology*, *65*(1), 85–101. https://doi.org/10.1037/0022-3514.65.1.85

Zebrowitz, L. A., & Rhodes, G. (2004). Sensitivity to “Bad Genes” and the Anomalous Face Overgeneralization Effect: Cue Validity, Cue Utilization, and Accuracy in Judging Intelligence and Health. *Journal of Nonverbal Behavior*, *28*(3), 167–185. https://doi.org/10.1023/B:JONB.0000039648.30935.1b

Zheng, R., & Lin, C. (2024). Cues driving trait impressions in naturalistic contexts are sparse. *Preprint at Https://Doi. Org/10.31234/Osf. Io/V39tk*. https://osf.io/v39tk/download