Information From Spectating Can Facilitate the Emergence of Conventions

Jon Atwell

jon@finegrained.io

Abstract

Conventions are a core component of social organization, but network-based models do not account for their emergence; the clustering present in social networks creates sustained competition between alternative behaviors. I argue that including actors' observation of the activities of non-alters can lessen that competition and facilitate group-wide coordination. Such *spectating* behaviors–observation without direct engagement—have long been theorized as relevant, but network models do not account for this potentially rich source of information. In large-group experiments (62 groups, 1488 participants), the addition of spectating-based information facilitates the emergence of group-wide conventions. This pathway yields more information than a variety of alternatives, including additional bridging ties. Its availability rapidly reduces the diversity of alternatives, which, frequently, then drives a continuous increase in coordination success until there is a single convention. These results suggest that spectating behaviors are a thrifty means of acquiring social information and can be instrumental in facilitating the emergence of conventions, from role structures to strong organizational cultures.

Introduction

Conventions—consensus expectations for behaviors in a given context—are fundamental building blocks for organizations and societies. Strangers shake hands on their first meeting in western societies. Sending unsolicited bulk emails is called "spamming." Hedge fund managers have long been paid "two and twenty." The person who is the driving force behind an academic paper will be the first listed author in many disciplines. The defining feature of a convention is that there are multiple behaviors that would leave the individuals using it about equally well off. Non-western societies have a variety of non-handshake greetings; what we now call spamming was also referred to as "trashing" or "flooding" in the early days of the internet (Templeton, 2001). Group-wide agreement on the appropriate behavior for a particular scenario is essential to the success of a variety of social processes.[1]

How do groups establish conventions? Some have argued that public rituals are an essential means. For example, a royal procession through an audience creates a convention around who has authority through the creation of common knowledge about the willingness of other denizens to submit to that authority (Chwe, 2001). This particular ritual, in turn, can allow rulers to solve other coordination problems by fiat. However, the conventions listed above were not solved by public rituals or fiat. Schelling instead argued that certain choices have a naturalness to them (Schelling, 1960). The first author's name is most visible on an academic paper, which is fitting for the one contributing the most. "Two and twenty" is tidy and alliterative. These focal points have a conspicuousness or salience that makes them easier to coordinate around. Although plausible, this argument often ignores counterexamples—alphabetical ordering of authorship is used in massive scientific collaborations, for example—and ignores the salience of alternatives—"one or thirty" is an investment management fee structure that is being put forward as an alternative to the "two and twenty" compensation convention (Elliot Davis, 2024).

The primary alternative to the public ritual and salience mechanisms is the *force of precedent* mechanism. Hume argued that a convention "acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it" (Hume, 1888, p. 490). Using this framework, Lewis (1969) defines a convention as a Nash equilibrium of a repeated coordination game. He argues that coordinating is not trivial for even dyads because

[1]This article does not address social norms, which also establish behavioral expectations, but in a more prescriptive and proscriptive fashion (Hechter & Opp, 2001). As discussed below, conventions are conceptualized as solving a *coordination* problem, whereas norms solve a *cooperation* problem. Coordination is about *which* option of many to select and cooperation is about *whether* to select the riskier cooperative option. Although the mechanics of the emergence of norms are beyond the scope of this article, the insights might well be applicable.

higher-order anticipatory beliefs quickly become too unwieldy. Instead, coordination must arise through trial and error; when individuals manage to coordinate successfully with one partner, they are more likely to try that same behavior with others. Although individual interests are best served by the existence of a group-wide convention,[2] past success with interaction partners determines future behavior. Accordingly, the core question becomes whether this trial-and-error process can lead to a single convention for a group.

It turns out that the force of precedence mechanism is insufficient when actors are embedded in clustered social networks. In the well-mixed populations that are standard in formal game theory, conventions can emerge through trial-and-error (Garrod & Anderson, 1987; Santos et al., 2006; Skyrms, 2004; Young, 2015). It may take many rounds of play, but eventually a single option wins out over others. This is encouraging for the plausibility of the force of precedent mechanism, but most observed social networks exhibit substantial clustering (Newman, 2003b) and coordination successes *within* clusters tend to increase competition *between* clusters, a tension well known in the network literature.[3] When an individual node's behavior or state depends on more than one other node in the network, clusters become "obstacles" (Easley & Kleinberg, 2010, p. 507) to the cascades that are necessary for the force of precedent mechanism to work.

This tension is present for many types of network-based processes, but especially so for coordination problems because short-term incentives to remain aligned within one's cluster can prevent the exploration that can lead to group-wide alignment. Consider again the now conventional use of "spamming" to refer to the sending of unsolicited bulk messages. The term originated in a small chat community in the early days of the Internet, but different groups used other terms. The introduction of "spamming" to a message board currently using "trashing" would disrupt the equilibrium, a negative outcome in the short term.[4] How can clusters overcome such disruption and short-term losses on the way to a single, network-wide

[2]Chwe 2001 refers to some rituals as rational because this alignment makes it rational to work toward a stable global equilibrium.

[3]The challenge clustering presents to group processes was a core premise of Granovetter's *Strength of Weak Ties* article (1973); subgroup structure impedes alignment because within-subgroup order is hard to perturb and reorient toward the broader group. This is why Burt (1992) labeled his structural holes re-description of the problem as a "theory of the social structure of competition."

[4]It took about 15 years for this term to win out over "trashing", "flooding" and other alternatives used in other communities (Templeton, 2001).

convention?

In a series of large-group experiments (24 or more participants) on the emergence of conventions, Centola & Baronchelli (2015) provide evidence that they cannot. Lattice networks, which have a high degree of clustering, produced stable competing alternatives. Stochastic regular graphs, which have a much lower degree of clustering, also did not successfully organize. Only in totally randomized interactions did a convention emerge. These results seem to imply that the force of precedent mechanism is insufficient for the emergence of conventions in the heterogeneous networks that we actually observe. Yet, we observe conventions in the real world and the question of what mechanism can enable their emergence remains.

In what follows, I argue that conventions are more likely to emerge when individuals are able to engage in *spectating* behaviors in which they observe social activities and practices of others they do not have a tie but are nonetheless still socially proximate. Following Adut (2018), I use the term spectating to refer to a range of social behaviors that are acknowledged in social theory as being distinct from direct interactions but are very much the "essence of the public sphere" (p. 18). Although in common usage "spectacle" often refers to fully public event, here *spectating behaviors* involve the "negligibly mundane" (*ibid*, p. 151), such as inadvertent eavesdropping on colleagues' conversations, seeing the code pull request of a distant colleague, or systematically noting the sartorial regularities at one's place of employment. Sometimes we are aware of doing this, but cognitive psychologists have identified several different modes of "spontaneous social inference" (Uleman et al., 2008) that allow us to spectate using subconscious processes.

Drawing on the research on network coordination processes, I show that spectating upon the activities and practices of proximate others with whom we are not directly interacting has the potential to transform the nature of competition between alternatives; broader knowledge about the existence of alternatives allows the group to unwittingly balance local (i.e., dyadic) coordination success with a steady reduction in the number of alternatives in circulation. The resulting patterns of information flow enhance the potential for the whole group to converge on a single shared convention, even within networks with high degrees of clustering. I test this potential by modifying the "Name Game" experimental paradigm (Centola & Baronchelli, 2015) to include spectating-linked information. Trials of groups of 24 participants in a variety of network topologies confirm that the availability of spectating-linked information increases the probability that the group will establish a convention in clustered networks. 50% all trials on heterogenous networks resulted in a single convention; the informational value of spectating is high and can make up for the information-for-coordination gap created by the existence of clusters.

These results demonstrate the potentially transformative role that spectating plays in social dynamics. Network-based thinking has dominated studies of social dynamics, including problems of coordination and cooperation without demonstrating that network channels are the only important non-broadcast sources of information. Spectating behaviors are a rich part of our lived experiences and have the potential to change the outcomes of group processes; more research is needed to better understand the prevalence of spectating behaviors and the conditions under which they actually induce different group outcomes.

The Challenge for Self-organizing Conventions

Our ability to establish social conventions—consensus expectations about behaviors in a given setting—is arguably one of the defining traits of our species (Elster, 1989). A simple example is the understanding that greetings in an American business context should include a handshake with only one's right hand. But the use of language is also a defining trait (Tomasello, 1999), and Saussurian and pragmatist approaches to linguistics argue that language is conventional (Grice, 1989; Garrod & Doherty, 1994). Many aspects of a culture are also conventional. The *culture as consensus* framework (Strauss & Quinn, 1997; Zerubavel, 1999; Patterson, 2014) is built on the observations that i) there are a range of plausible actions for the given situation, and yet ii) groups create consensus expectations about actions and meaning, but in the end iii) different groups come to have different, durable consensus expectations. This means the term convention can subsume social patterns otherwise labeled habits, customs, routines and practices (Biggart & Beamish, 2003). Anything that encompasses " 'habituated normativity,' where largely tacit routinized views direct actor preferences, decision-making, and hence behavior repertoires concerning 'what ought to be' " (*ibid*, p. 454) can be thought of as conventions—"institutional theory writ small" (*ibid*, p. 457).

Dispositionally oriented accounts of culture (Lahire, 2011; Tavory, 2018; Rawlings & Childress, 2019) look to the life course and common human-mediated experiences to account for such sharedness. Individuals who never interact can have the same expectations and meanings because their experiences are similar in virtue of the macroscopic patterns of their societies. This orientation is essential to, for example, Bourdieu's account of the class-based origins of habitus (Lizardo, 2004). Situational approaches contrast with dispositional ones by highlighting the relational pragmatism (Emirbayer & Mische, 1998) found in the actual interactions of shared situations. Exemplars include White's *Identity and Control* (2008) and the work of Gary Alan Fine (c.f. Fine, 2012), but the perspective is a facet of the broader relational and interactionist framework that emerged with the rise of social network analysis (Emirbayer & Goodwin, 1994). In this approach, processes such as social influence and consensus building are each "a micro-level *social process* that unfolds in a macro-level *social structure*" (Friedkin & Johnsen, 2011, p. 13, italics in original). Thus, for example, "very strong, dense, and relatively isolated social networks facilitate the development of uniform 'subcultures' and of strong collective identities" (Emirbayer & Goodwin, 1994, p. 1419).

Although dispositional and contextual antecedents are largely inseparable in practice (Lahire, 2011; Rawlings & Childress, 2019), this article focuses on the synchronic and *situational* emergence of conventions. Conventions have been studied this way since Hume argued that conventions emerge not through explicit agreements, but through the tacit understanding that parties have shared interests and that their prior success should ensure the "future regularity of their conduct" (Hume, 1888, p. 409). That is, if we have a need to coordinate behavior somehow and stumbled into a solution once, we have every reason to behave the same way in similar situations again.

The philosopher David Lewis elaborated this perspective within the framework of coordination games (Lewis, 1969). Many are familiar with non-cooperative games like the prisoner's dilemma, in which both players would do well to agree to an action, but each is incentivized to defect if they know the other is planning to cooperate. In contrast, in coordination games, there is no incentive to defect. If a player is confident about what the other will do, they maximize their reward by coordinating. This structure of coordination games, Lewis argues, is the essence of conventions. It does not matter much what we choose to do, but we all benefit when we agree on the choice. This is a stable equilibrium qua convention.

When there are few parties and few possible actions, coordination problems are clearly solvable. Families can quickly, maybe even tacitly, agree on a seating arrangement at the dinner table. However, as the number of possible options and parties increases, successfully coordinating becomes more challenging. Can Hume's *force of precedent* argument hold for broadly used *social* conventions? Thomas Schelling argued that maybe the problem is not as bad as we might fear (Schelling, 1960). Certain behaviors or options might naturally draw attention or have salience. For example, the emergence of the practice of sharing items on social media begged for a name. Perhaps the term *meme* won out despite being a fairly obscure academic term for a tangentially related concept (Dawkins, 1976) because it is short and unique.

Although salience can shrink the space of possibilities, it is unlikely to solve the coordination problem completely. We might now be inclined to think that the choice of *meme* as a naturalness to it, but it required a change in the meaning of the existing term and there were viable alternatives (e.g., "forwards" or "ebaums"). This is the case for spam and many other neologisms too. As the analysis below shows, having a few competing alternatives does not reduce competition so much as entrench it.

One sure way to circumvent competition is to make expected behavior visible to mass audiences simultaneously. This top-down means of establishing conventions is the subject of Chwe's work on "rational rituals" (2001). A ritual creates common knowledge about the beliefs and behaviors of others in order to establish an equilibrium social order. Rituals are rational because any coordinated equilibrium is better for each individual than a lack of coordination. The public nature of rituals broadcasts information widely, but is rarely a means available for the establishment of contemporary social conventions. Broadcast media more generally can be a viable mechanism—see the now-consensus use of the term "nepo baby" for public figures benefiting from their parents' success after a single article in *Vulture* (Jones, 2022)—but are more likely to act as the final act of entrenchment for a long-running, bottom-up process.

For these reasons, the literature on the origins of conventions has focused on whether the force of precedent mechanism is viable. The game-theoretic approach pioneered by Lewis and H. Peyton Young (1993) has concluded that even when players do not play their optimal response (i.e., stochastic evolutionary game theory), large *well-mixed* populations can converge to a global convention (Young, 2015).[5] The assumption that the population is well-mixed—the probability of any given pair interacting is uniform—is common in the analysis of dynamical systems because it greatly improves the mathematical tractability of the model. However, it is antithetical to research on social networks. The main premise of the social network perspective is that networks are fundamental social structures through which social patterns emerge (White, 2008; Friedkin & Johnsen, 2011). "It is through these networks that small-scale interaction becomes translated into large-scale patterns, and that these, in turn, feed back into small groups" (Granovetter, 1973, p. 1360). Interpersonal interactions influence behavior, beliefs, and emotional states. Even when not mutual, this leads to an increase in the alignment of behaviors among those interacting. As groups of individuals become more alike, this local patterning becomes causally implicated in social life as identities, audiences, and discourses take hold and themselves structure the nature of influence (Flache & Macy, 2011; DellaPosta et al., 2015; Centola, 2015).

Heterogeneous network structures (contra well-mixed interactions) have important functional implications through the ways in which they structure competition and the flow of information. One of the most robust features of observed social networks is the high degree of *clustering*, a measure of the density of the ties within the group relative to the ties with others outside the group. At the lowest level, within-group density is high because of the principle of transitivity; if A and B have a tie and A and C have a tie, it is likely that B and C will also have a tie. The social causes of transitivity might be rooted in balance theory (Harary, 1959) or the

[5]For these results convergence is only loosely defined; there is no threshold fraction of the population required to be using a convention, but rather that the dynamics have tipped toward one of the possible equilibria ("a broken symmetry" in the terminology of statistical physics). Because individual incentives are aligned with the existence of a global convention, once the system tips toward one equilibrium, full adoption is ensured. Note that, given enough time, stochastic play leads to a change in conventions (Young, 2015). The possibility of conventions changing (e.g., a nascent movement away from men proposing marriage in heterosexual couples) is an interesting feature of this class of models, but a topic outside the scope of this paper.

social capital tighter groups accrue (Coleman, 1986; Portes, 1998). Regardless of the origins of clustering, if one takes social networks to be the primary vehicle for the social process of consensus formation, mathematical and experimental evidence indicate it is a significant impediment to the emergence of consensus.

Mathematical treatments of diffusion and social contagion—types of influence processes with similarities to coordination processes—on large networks clearly highlight the challenge clustering presents for unidirectional influence. Clustering has an inverse relationship with the degree to which a diffusion process is able to penetrate the broader social network (Pastor-Satorras et al., 2015). On random networks, the size of outbreaks in the SIR model of epidemics decreases as a function of clustering in the network (Keeling, 2011; Newman, 2003a). The Watts-Strogatz's model of small-world networks also exhibits this same pattern (Moore & Newman, 2000). More generally, when degree assortativity, a positive correlation in the number of total ties neighbors have, is held constant, more clustering leads to smaller outbreaks (Miller, 2009; Serrano & Boguñá, 2006).

Contagion and diffusion have some commonalities with multidirectional social influence processes, but the latter involve feedback (Flache et al., 2017) and are likely to exhibit more complex dynamics. The DeGroot model of multidirectional influence (Degroot, 1974) assumes that individuals conform to others in the group by revising their estimates of a continuous parameter based on a convex combination of all other estimates in the group. In this framework, consensus is assured given sufficient rounds of updating. However, clustering is still an impediment. Golub & Jackson (2010) show in a DeGroot-like model that convergence is possible, but the rate is inversely related to the size of the second eigenvalue of the adjacency matrix. This value is called the coefficient of algebraic connectivity or the Fiedler value, and the larger it is, the longer it takes for a network to reach a consensus value (Olfati-Saber & Murray, 2004; Olfati-Saber et al., 2007; Newman, 2010). Although networks with high clustering tend to have a high coefficient of algebraic connectivity, the two are not the same. The latter is related to the number of node- or tie-removals that would separate a network into multiple components. This robustness to separability turns out to be the topological feature that has the closest relationship to the success of coordination processes. In this sense, connectivity is the functional basic of coordination potential that originates in the behavioral tendency for social networks to be clustered.

Thus, in their review of social influence models, Mason, Conrey & Smith (2007) conclude that formal models indicate that diversity is generally maintained if there is a community structure (i.e., clustering) in the network, and influence is nonlinear, meaning actors do not incrementally and evenly incorporate information from all network neighbors. This claim has been supported in a variety of laboratory experiments (Judd et al., 2010; Moussaïd et al., 2013; Centola, 2015; Chandrasekhar et al., 2015).

Note that sometimes results that show the maintenance of competition are treated as *explanations* for social conventions (Axelrod, 1997; Bednar & Page, 2007; Bednar et al., 2010; Derex & Boyd, 2016). Such claims generally suffer from an incommensurability of scales, however; the term culture has been applied to aggregates ranging from nations with populations of hundreds of millions to fewer than a dozen individuals (like Fine's Little League teams 1979), and while the concept at either of these extremes may be useful for the question at hand, they are clearly not the same. Processes that create subgroup consensus but maintain whole-group diversity within small populations do not explain cultural heterogeneity between larger aggregates because the interaction patterns do not scale up. Conventions in particular are useful when they are held by large populations, and therefore the maintenance of alternatives within small-scale networks is inconsistent with the notion of broader social conventions. Thus, while one could use clustering within networks to account for one community using the term *spam*, another using *trashing* and a third using *flooding*, that mechanism cannot account for cross-cluster alignment, i.e., the now near-global use of the term *spam*.

In summary, a great number of social conventions must be established by force of precedent. Formal and experimental studies have shown that a force of precedent mechanism is sufficient to explain the emergence of a single convention in populations with well-mixed interactions. However, well-mixed interactions are antithetical to a network-based perspective. The clustering of network ties is an essential feature of observed social networks and ensures that interaction patterns deviate significantly from being well-mixed. Once heterogeneously structured networks are used to model the interaction patterns, the force of precedent mechanism is no longer a viable account for the emergence of conventions. Another explanation is needed if we are to take the existence of network clustering seriously. Staying within the network and force of precedent frameworks, in the next section, I argue that spectating behaviors are a realistic means of social learning that facilitates better circulation of information in parallel to the usual network pathways. As such, the addition of spectating-based information has the potential to enable the emergence of conventions, a proposition tested experimentally in the subsequent section.

The Psychology and Sociology of Spectating Behaviors

I use the term spectating as a general label for a variety of behaviors in which individuals observe a social practice, or evidence of it, without direct participation in it. Individuals engaged in a spectating behavior might be doing so intentionally because of perceived relevance, but as I discuss below, that need not be the case. This usage can encompass the behavior associated with the word *spectacle*, which is generally taken to be an event put on, or an object put out, for the public's gaze (Adut, 2018). A spectacle, however, is much closer to Chwe's notion of rational rituals; it is organized for the public, often with the intention of defining or conveying some social reality. Spectating can involve much more mundane activities (Adut, 2018).

Consider a Ph.D. student going to her first large academic conference. Her expectations of how to, for example, effectively construct a short presentation and interact with others might initially be informed by the advice from those in her social network (her home department). The conference offers many opportunities to engage in spectating behaviors and update these expectations. Attending numerous presentation sessions will reveal more about the conventions around how to structure presentations, how and when to ask questions, and how to introduce oneself to speakers. Walking the halls will provide a broad sample of acceptable conference attire by rank and gender. Overhearing conversations in lobbies or adjacent restaurants can reveal the degree to which people use this conference to start new collaborations versus catch up with friends.

Crucially, any two people who have this first conference experience will not observe the exact same people or spaces. Spectating behaviors involve individuals having different experiences and only rarely does spectating involve widely broadcast signals and shared experiences (e.g., a presidential address at the conference). Instead, spectating behaviors exist between the intensive, private interactions codified in networks and the fully public broadcasts intended to create common knowledge.

How common are spectating behaviors? We might occasionally intentionally read a crowd or allow ourselves to eavesdrop on a conversation, but most try not to make a habit of it. Does that mean that it is an infrequent type of behavior? Research on social cognition in light of dual process models of cognition indicates that we might be spectating much more than we realize because intentionality and explicit attention are not necessary for making robust social inferences. The dual process model of cognition categorizes cognitive processes under binary labels–implicit versus explicit, unconscious versus conscious, fast versus slow, automatic versus deliberative, system 1 versus system 2, and others. Although human cognition is not as simple as the dual process model of the public imagination (Evans, 2008; Fiske et al., 1999; Melnikoff & Bargh, 2018), each binary captures an aspect of the variation in cognitive processes. Processes of the implicit, unconscious, fast, and automatic sort operate in the background, often without our awareness. Research has shown that these types of processes are abundant in social cognition.

The factor structure of social cognition is an active area of research (Happé et al., 2017), but humans clearly employ a wide range of processes to make sense of the social world (see Fiske & Taylor, 2017 for an overview). Of primary interest here is *attribution*, the processes through which we make inferences about the causes and meanings of the behavior or perceived mental states of others. Tomasello and colleagues (Tomasello, 1999; Tomasello, Carpenter, Call, Behne & Moll, 2005) argue this attributional ability, our *theory of mind*, is the fundamental evolutionary building block that makes cumulative cultural evolution and our evolutionary success possible. The study of attribution originated in the commonsense psychology of Heider (Heider, 1958) and its emphasis on controlled perception of others' mental states through natural language and explicit causal reasoning, but the field has since identified a variety of *spontaneous social inference* processes (Uleman et al., 1996; Krull &

Dill, 1996; Duff & Newman, 1997; Ham & Vonk, 2003; Uleman et al., 2008).

Spontaneous social inference is when we “routinely and habitually impose meaning on the events around us, even when we have no immediate purpose in doing so” (Uleman et al., 1996:211). These inferences are wide-ranging in nature, but most relevant here is evidence that we make spontaneous inferences about the goals, beliefs, and values of others (Uleman et al., 2008). See, for example, Hassin, Aarts & Ferguson (2005) for goals, Onishi (2005) for beliefs, and Ham & van den Bos (2008) for values. These inferences are made with cognitively efficient implicit or automatic processes, even when we are engaged in explicit comprehension of others’ mental states. For example, while conversing with others and making explicit and controlled inferences, we simultaneously make inferences based on gestures (Willems & Hagoort, 2007; Kelly et al., 2010) and prosody (Sander et al., 2005).

Some of these spontaneous inferences originate in visual field processing that feeds into higher-order, more semantically oriented cognition. For example, ensemble perception allows us to rapidly gauge summary statistics of similar objects (Whitney & Yamanashi Leib, 2018). This includes average gaze direction, resemblance, and affective states (Chen & Whitney, 2019). When combined with existing mental schema, these faculties allow a near instantaneous comprehension of the core elements of a social context.

Spontaneous inferences are also found in language comprehension. Although words themselves can be ambiguous, the larger problem of human communication is pragmatically linking utterances to the mental states *cum* meaning in the speaker’s mind (Grice, 1989), a problem that has been described as impossible on first principles (Levinson, 2000). Yet we have a remarkable ability to rapidly infer meaning and intent from utterances (Garrod & Pickering, 2004; Wilson & Sperber, 2002; Sperber & Wilson, 2002). This “mind-reading” capability is likely enabled by a combination of dedicated and shared cognitive submodules, including some of the above visually-oriented ones (*ibid*). Key to the success of mind-reading are the subattentive (i.e. implicit) processes that do the vast majority of processing of the welter of visual and auditory stimuli to extract the most relevant ones. These processes manifest themselves in phenomena such as our tendency to imitate ambient speech (Delvaux & Soquet, 2007) and the particularly distracting nature of half-heard conversations (Emberson et al., 2018). The automaticity of much of language processing ensures that any audible speech is processed. When the relevance of automatically processed language is high, explicit attention can be triggered (Wilson & Sperber, 2002), but lack of explicit processing does not mean the stimuli flow by without consequence.

The relative infrequency with which we explicitly spectate can limit our appreciation of the potential importance of implicit modes of spectating. We are rarely aware of our engagement in implicit or automatic modes of cognition and, accordingly, can be unappreciative of spectating behavior in those modes. Spontaneous social inferences sometimes serve as cognitive antecedents to more meaningful moments of spectating behavior. For example, an implicit categorization of an overheard conversation as having an argumentative tenor could lead to explicit eavesdropping and attempts to identify the interlocutors. In that case, the informative element of the complex of spectating behavior is the eavesdropping that was initialized after making spontaneous tonal inferences. Like many behaviors, spectating can shift between cognitive modes as we allocate explicit attention when our spontaneous inferences identify characteristics of interest (Fiske et al., 1999).

Spontaneous social inferences have long been acknowledged in social theory. George Herbert Mead, as a key theorist of pragmatism and microsociology, argued that "we are reading the meaning of the conduct of other people when, perhaps, they are not aware of it. There is something that reveals to us what the purpose is–just the glance of an eye, the attitude of the body which leads to [our response]" (Mead, 1934, p. 14). He also argues that "consciousness is emergent from such behavior; that so far from being a precondition of the social act, the social act is the precondition of it" (Mead, 1934, p. 18); that is, the reading of conduct of others happens before we even understand it, and therefore it becomes a social act. Furthermore, our self and social consciousness is defined in relation to a "generalized other", the attitude of the whole reference group that we have in mind as we process social acts. Thus, we have an interest in observing the whole reference group, not just those with which we interact.

Work that builds on Mead, including the symbolic interactionism of Herbert Blumer (1969) and the dramaturgical approach of Ervin Goffman (1959), is focused on how "group or collective action consists of the aligning of individual actions, brought about by the individual's interpreting or taking into account each other's actions" (Blumer, 1969:82). Goffman makes the spectating audience present in all acts of interpretation clearly through his extended quotation of William Sansom's novel *A Contest of Ladies* (1956) in which the logics of impression management of the subject, Preedy, is relayed in great detail (Goffman, 1959:p. 4-5). Preedy is concerned with the interpretive affordances his physical movements and facial expressions give to any would-be observer of his solo trip to a beach. The other occupants of the beach, none of whom he knows, are Preedy's audience. Goffman does not think that this sustained dramaturgy requires intentionality. His approach is concerned with communication of "the non-verbal, presumably unintentional kind, whether this communication is purposely engineered or not" (Goffman, 1959, p. 4).

Veblen's notion of conspicuous consumption also highlights the audience but assumes more intentionality. Consumption is conspicuous to the extent that it is visible to a broad audience to signal the class status of the individual. Consumption becomes a social act via the implied spectatorship, and through this act, the class structure is reified. Another important realm for spectatorship through which society is defined is the *public sphere*. *The Spectator* was the name of an influential eighteenth-century publication in London[6] that Habermas credits as being a mirror through which a public viewed itself, an important turn by which the notion of the public sphere was born (Habermas, 1991:43).[7] The literary device that made *The Spectator* popular among period readers and later scholars is its first-person perspective rendered in rich sensory detail (Powell, 2012). The narrator, Mr. Spectator, is a chronicler of people and place; the sights, sounds, smells, and even tastes fall within his gaze. Although this gaze can be objectifying, it is innocent because it "is often a natural, noninstrumental curiosity that makes us interested in other people's business" (Adut, 2018, p. 149).

General visibility (or audibility) and the spectatorship it allows is central to a coherent notion of the public sphere (Adut, 2012). Once we enter spaces with visibility, we "cannot

[6]The current periodical *The Spectator* pays homage to the original through its name and moralizing tone, but not its literary forms.

[7]Habermas's history of the public sphere as presented in his 1962 monograph has been criticized as misrepresenting many trends and details (Adut, 2018). Literary scholars have noted in particular that the editorial project of *The Spectator* was "the reform and the discipline of public sociability," not the encouragement of an inclusive public discourse within a public sphere ((Cowan, 2004, p: 346),(Newman, 2005, chap. 1)).

help giving off signs" that are "seen and heard by unspecified others and submitted to their judgment" (Adut, 2012, p. 242-43). Arendt bemoans this visibility as a "'withering away' of the private realm" (Arendt, 1998, p. 72), but others view it as essential for healthy communities. As a core of her critique of city planning, Jane Jacobs argued that spectating behaviors on the street are "ostensibly utterly trivial, but the sum is not trivial at all" because engaging in them provides "a feeling for the public identity of people, a web of public respect and trust" (Jacobs, 1992, p. 61) that makes communities viable as such. This view has been supported through empirical work (Browning et al., 2017) and extended to include the "reserved sociality" of semi-private spaces adjacent to the street, namely balconies (Zacka, 2020).

Innocent curiosity in the public sphere is distinct from intentional observation in competitive arenas. The framework of structural equivalence eschews "the interpersonal synapse" (Burt, 1987, p. 1288) of network ties as the medium of influence and observation, arguing that in many strategic settings "the process responsible for social influence shifts from communication within a primary group to competition and relative deprivation within a status" (Burt, 1987, p. 1294) order shared by actors who "have no relations with each other and so do not socialize each other directly" (Burt, 1987, p. 1293). Effective competition requires that one observes similarly embedded actors, even if they are socially distant. Similar assumptions are made in various *field* theories that situate purposive action in meso-level orders constituted by—as is the case with fields in physical systems—diffuse forces operating simultaneously on all elements in it (Martin, 2003). Bourdieu's agents have a *habitus*, a cognitive construct that helps to organize experiences of fields into practices and representations (Lizardo, 2004). Much of this revolves around cultural distinctions observable at the embodied level, including bodily dispositions, speech patterns, and style of dress. Effective agents have a spectating imperative if they are to successfully learn the allocation of cultural capital in the broader field. In their version of field theory, Fligstein and McAdam's label effectiveness *social skill*, a "highly developed cognitive capacity for reading people and environments"(Fligstein & McAdams, 2012, p. 17).

The above is not an exhaustive survey of spectating behaviors covered in sociology, but these different theoretical lineages point to both how common spectating behaviors are and their importance to human sociality. By using a common label for a range of different behaviors, I aim to accomplish two things. First, I highlight a basket of behaviors that circulate social information in ways that network-based theories currently omit. There is variation in content and goals, but these behaviors have the common feature of informing our understanding of social situations through pathways that would not be coded as ties in the framework of social networks. By identifying this commonality, I am able to create a functional and testable account of the potential of such behaviors to facilitate the coordination necessary for the bottom-up emergence of conventions.

Before turning to the experimental test, I discuss a key experimental result in the emergence of conventions, which reveals how the properties of different network topologies influence the dynamics and outcome. Reviewing these formal properties of networks shows why spectating-based information has the potential to facilitate the emergence of conventions.

The Network-based Study of Coordination

The central question in this paper is whether spectating behaviors can change individual beliefs enough to facilitate the group-wide coordination central to the emergence of conventions. Doing this necessarily builds on the Name Game introduced by Centola and Baronchelli (2015)—henceforth *CB*—to explore the relationship between network topology and the emergence of conventions. The game is derived from theoretical (Lewis, 1969; Young, 1993; Garrod & Doherty, 1994) and formal (Baronchelli, Felici, Loreto, Caglioti & Steels, 2006; Baronchelli, Loreto & Steels, 2008) models that define conventions as Nash equilibria in repeated coordination games.

The Name Game incentivizes participants to coordinate on a name for a pictured individual. Participants are embedded in a network ($N \geq 24$) and are randomly paired with their network connections in rounds. In each round, participants have a timed period to independently propose a name for a pictured individual. Focusing on a picture of an individual restricts the possible names to a meaningful, yet plentiful subset of valid strings (e.g., any string of characters). Upon submission, both participants are shown the names they and their partner played. Partners who match receive a small monetary reward for the round ($0.10). Partners who do not match are penalized a smaller amount ($0.05), as long as their current cumulative rewards are sufficient.

The partners do not know each other's identities before or after the interaction, as this would introduce second-order social processes and skills that would obfuscate the dynamics of the core coordination task and undermine the value of the experiment; participants would likely quickly solve the unique two-person coordination games by memorizing identities and the corresponding names, leading to pairwise solutions to the dyadic coordination problem. The social value of conventions, however, is in the fact that they facilitate successful interactions across large groups of known alters and strangers alike by removing the cognitive load of having to memorize highly specific settings for particular behaviors. Withholding identities allows the experiment to model the more generalized group coordination setting. Although participants do not have direct incentives to create a single convention for the whole group, their long-term interests are best served by one. This feature of the setup is not directly articulated to the participants.

In *CB* the independent variable is the topology of the network in which the participants are embedded. *CB* considers a random graph of constant degree four, a circular lattice of degree four, and a fully-connected graph. Across several trials with groups of different sizes, group-wide conventions emerged only in fully-connected networks. The groups embedded in the other two types of networks exhibited entrenched alternatives at the end of game play. As discussed above, this failure to coordinate is predictable (Mason et al., 2007); clustering induces competition that results in competing alternatives.

What is it about the fully-connected networks that allow participants to successfully create global conventions? The label fully-connected is sensible insofar as all participants have a positive chance of interacting with each other, but if participants play for approximately as many rounds as there are participants—as they do in *CB*—they have a vanishingly small chance of realizing all possible interactions. That is, a *complete* network is never realized. Instead, the pairing process results in a variant of the $G(n, p)$ Erdős-Rényi model of random graphs.[8] The $G(n, p)$ model realizes each edge between n nodes with a probability of p. This model produces a very similar network ensemble as the *CB* pairing process when $p = .65$; the expected degree distribution has the same mean and similar variance as the graphs created by the pairing process.[9]

This connection is relevant because a long tradition in network control theory addresses coordination and consensus in random networks. The formal results show that autonomous decision-makers create consensus when an alignment process plays out on an Erdős-Rényi graph (Hatano & Mesbahi, 2005; Bertsekas & Tsitsiklis, 2007; Tahbaz-salehi & Member, 2008). More generally, the higher the coefficient of *algebraic connectivity* (i.e., the second smallest eigenvalue of a network's Graph Laplacian), the faster convergence to a consensus value occurs (Olfati-Saber & Murray, 2004; Olfati-Saber et al., 2007; Newman, 2010). Algebraic connectivity quantifies how robustly a network is connected; if it would take the removal of many ties to separate the network into separate components, the coefficient is high. While inversely correlated with community structure in some instances, this measure is different for the same reason small-world networks are possible; the addition of a small number of ties linking clusters can radically change the length of shortest paths between any two nodes and, therefore, introduce redundant connections between them. The above formal results show that this robustness translates into efficient information circulation.

The lattice and random regular graphs studied in *CB* have much smaller coefficients (0.33 and approximately 1 on average, respectively) than the fully-connected network (approximately 11 on average).[10] By this fact alone, fully-connected networks are far more likely to facilitate convergence to a single global convention than the other networks. Humans do not update their beliefs indefinitely with mathematical precision, so the applicability of model dynamics over long time frames is questionable, but the results in *CB* cleanly correspond to what algebraic connectivity would predict.

What does the addition of spectating behaviors do to the algebraic connectivity of a

[8]The graph with the label *random* in *CB* is a random regular graph in which every node has 4 edges. This is very different from the $G(n, p)$ graph, which has a Poisson degree distribution.

[9]For $N = 24$. The underlying analyses by the author are available upon request.

[10]The coefficient was determined simulating pairing in a fully connected network with a *perfect matching* constraint in which all nodes get exactly one edge each round. A *complete* network would have a coefficient of 24.

network? That depends on the relationship between the structure of social networks and the sites of the spectating behaviors. In most contemporary societies, social networks overlap in physical space. Neighborhoods “are a series of overlapping social networks” (Forrest & Kearns, 2001, p. 2130). Cities contain“heterogeneous and densely populated public spaces” where people congregate, what Anderson (2011) calls cosmopolitan canopies. Household-place *ecological* networks are more *extensive* than *intensive*, meaning households tend not to share multiple place ties so much as connect the broad network of places (Browning et al., 2017). Administrators adopted “social distancing” policies as a tool to fit COVID-19 because network non-alters frequently share physical space (Glass et al., 2006). This research shows social networks overlay themselves when projected into physical space and if additional ties are formed on the basis of physical proximity, it implies an increase in (algebraic) connectivity and the attendant possibility of more relevant information being obtained.

Imagine wanting to know the candidate preferences for a congressional primary within your commuter community. Assuming there is consolidation among traits, social network alters are likely to be within clusters and have similar preferences and knowledge (Centola, 2015). Thus, while the community might exist in a single connected network, denizens asking their network alters what their beliefs about the distribution is will not allow them to converge an accurate estimate. A few bridging ties can yield more insight, but unless there are many such ties across social space, it will not be enough.[11] By contrast, seeing yard signs for candidates on the way to a grocery store or gym will reveal crucial information about the distribution of preferences. That route is not a random sampling of streets, but it travels through multiple clusters within the social network, and in that sense increases the algebraic connectivity of the network of observations. Overhearing people talk about candidates at the grocery store or gym would also increase connectivity because these establishments draw people from across the network and community, although again not a random sampling of them.

Despite individuals’ observations not being representative, the increase in algebraic

[11] Weak ties provide valuable information (Granovetter, 1973; Burt, 1992) from other regions of the network and their addition would also increase the algebraic connectivity. Intuitively, however, a few weak ties added to a highly clustered network are unlikely to allow individuals to understand the broader picture. Formally, this turns out to be true, as adding a few weak ties is not sufficient to induce a phase transition towards consensus(Olfati-Saber, 2005).

connectivity implies that information will be better circulated. Could this increase be enough to tip the group toward consensus on a single convention? The additional ties increase the network density, but they are directed, which increases algebraic connectivity comparatively less than bidirectional ties. The mutual observation facilitated by bidirectional ties can directly create alignment between the relevant parties, but the coordination value of unidirectional observations is necessarily indirect and diffuse. Furthermore, because actors could be aware of this fact, participants' perceived salience of that information might be less than information from direct interactions with alters. Both of these factors could undermine the theoretical utility of the increased connectivity. The assess the value of spectating-linked information, I modify the Name Game to include opportunities to spectate. In light of the fact that spectating behaviors could rely on implicit or subconscious information processing and therefore be underappreciated by actors themselves, the binary outcome of a group coordination problem is a stark test of the potential for spectating behaviors to be consequential.

Spectating Behaviors in the Name Game

To assess the effect of spectating behaviors on the emergence of a single convention in large groups I extend the Name Game introduced in *CB*. 24 participants are randomly placed in a predetermined network structure. For each of 25 rounds of game play, the individuals are paired with one of their network "neighbors". They do not know in advance of submitting a name which partner it might be, nor do they know anything about the structure of the network, from the number of neighbors they have to the overall size of the network. They know an approximate amount of time the game should last, but not the number of rounds to be played. They just know that they are trying to submit the same name for an individual shown on the interface (i.e., a head shot of a woman). The image is the same for all participants throughout all rounds.

Participants are compensated for completing the task, for any time spent waiting for the game to start, and for matching names with their partner for a given round. They receive a bonus reward of \$0.10 for each match and a penalty of \$0.05 for each time they fail to match, as long as they have a positive reward balance. After each partner has submitted a valid name (see the Supplemental Information for a detailed explanation of what names are valid), they see the names and the associated bonus or penalty. The interface displays each player's cumulative bonus and their round-by-round history of matching successes or failures. This does not include the names played, so the participants are on their own to recall the names they and their partners have played and whether they matched.

Whereas in the original game, the only source of information about the names in circulation was through those interactions with partners, my extension adds spectating-based observations by revealing to participants one or two additional names played in a given round. There are a number of ways to structure the pattern of names participants see that could correspond to spectating behaviors, but to test the general proposition that spectating-linked information *can* alter the outcome of the coordination process, I use three simple treatments: no additional names, one per participant per round, and two per participant per round. The names shown to each individual are drawn at random with replacement from the population, excluding the participant herself. This can include "None" if the corresponding player failed to submit a valid name in the 15 seconds that they have to do so.[12] This means that, in a given round, it is likely that some participants' behaviors are not observed by anyone other than their partner, while others' behavior will be viewed by a handful of others; while spectating is not a broadcast mechanism (i.e., from a single source to all possible receivers), it is also not a strictly localized mechanism either.

To enforce the desired non-localized nature of this design, no information is given regarding who played these spectating-linked names, or whether they matched that of the respective partner. The names are labeled as coming from "random players", while the name played by the participant's partner is identified as such. Crucially, there is no direct benefit in matching these names and, without knowing the behavior of the corresponding partner, the value of that name for the sake of coordination is not known. It is merely an observation of behavior that participants can process as they see fit. This design does not diminish the

[12]To facilitate the round-by-round matching of players and ensure that there are not segments of the network playing faster than others, the next round does not start until all participants in the network have submitted or timed out and given at least 10 seconds to view the names. In practice, participants take about five seconds to submit and about four seconds to view the results. See the Supplemental Materials for more details about the management of game play.

immediate, direct incentive to match with one's current partner or the indirect, long-term incentive for the whole network to have a common convention; if participants ignore the names from the "random players", the incentive structure is identical to the original game.

There are other plausible designs for the addition of spectating-linked information. For example, the independent treatment variable could be an *expected* number of names per round, and the count that a participant sees is drawn from a binomial distribution with the corresponding mean. Or, names could be sampled with a bias toward a particular region of the network for each participant. Participants could also observe the names of both partners in an interaction to learn if they coordinated successfully. However, there is no research on the specific patterns of observations, and I opt for a straightforward and relatively unstructured design. Ultimately, the goal of the experiment is to assess whether spectating-based information has the potential to change outcomes, not to reproduce spectating behavior in the lab. The design used here is a useful baseline case. If it fails to induce different outcomes, the specific statistical patterns of spectating observations are likely not relevant for consensus-building processes. If it does, those patterns merit more research.

The experiment includes the same three network treatment factors in *CB*; a lattice network of degree four, a random network with constant degree four, and the fully-connected network. I also include a small-world network seeded with a circular lattice of degree four. Small-world networks share the important characteristics of local clustering and short characteristic path lengths with real-world social networks (Watts & Strogatz, 1998; Watts, 1999). Because the random networks in *CB* did not converge to a single convention, a small-world network, which is an interpolation between a lattice and a random network, would not have added anything to the analysis. The potential for spectating-linked information to overcome the challenges of clustering makes its inclusion more important.[13,14]

[13]Because realizations of stochastic network models (i.e., random and small world) yield a range of basic network characteristics (i.e., clustering coefficient, diameter, and characteristic path length), the experiment uses the four realizations closest to the mean from among 100 realizations of the model. These four networks were used once for each level of the additional names factor. Similarly, the same random pairings were reused for each level of the fully connected network trials. The networks used have the following average clustering coefficients, characteristic path lengths, and diameter, respectively: random (0.14, 2.34, 4.25), small world (0.28, 2.51, 4.75), lattice (0.5, 3.39, 6). Empirical clustering coefficients usually fall between those of these random and small world graphs (Newman, 2003b).

[14]All graph generation and measurement was done using the NetworkX python package (Hagberg et al., 2008).

Before articulating hypotheses, it is useful to consider the algebraic connectivity of the observations in the game. The nature of the spectating observations in the game is unidirectional, so I use Wu's definition of algebraic connectivity for directed networks 2005 that assigns each directed edge half the weight of an undirected edge. The directed edge for each spectating observation is added to the network of mutual partner observations, which are coded as reciprocal directed ties.

Table 1 displays algebraic connectivity by condition. In *CB*, the fully-connected networks with no spectating successfully coordinated on a single convention with an expected coefficient of connectivity of approximately 11. The fact that all conditions with spectating have higher coefficients implies that they should also be able to converge within the 25 rounds of game play, if not much faster. Interestingly, the addition of spectating ties for both one and two names is almost wholly determinative of the coefficient for random regular, small-world, and lattice networks. Although the underlying networks are topologically quite distinct, the additional ties from spectating quickly mask over those differences to produce equally connected graphs.

[Table 1 about here.]

This analysis suggests that the availability of even one spectating-based name can allow a group to converge to a single convention in all network topologies. However, because of its incentive structure, the Name Game is different from the formal network control problems in which algebraic connectivity is generally used. Therefore, it is not obvious that the presence of spectating-linked names will change the outcome of the games. The motivating thesis of this paper is that it will, but there is a gap between the informational value of the spectating-linked names and the presence of behaviors that would give it coordination value. Participants may simply ignore information that is not immediately relevant to their incentives. It is clear in the instructions and the interface which names are those of their partners and that only those matter for incentives. If a participant's partners are all playing "Amy" but the spectating names are mostly "Amanda", the participant would be myopically rational to ignore the Amandas to only monitor the names played by partners. Such behavior would mean that the informational value of spectating does not translate into coordination value.

Alternatively, even if participants do monitor the spectating names, the volume of names seen may be too much for individuals to process into actionable information in the small amount of time they have. Or, the way participants make use of the spectating-linked names could induce the circulation of so many names that it impedes the emergence of locally consistent behavior. The literature discussed above argues that we spectate a great deal as we move about the social world, but actualizing coordination potential requires the right set of cognitions and adaptive behaviors. Given the power of It is possible that the laboratory setting may fail to induce them.

Following *CB*, I conservatively define the successful emergence of a global convention as all players using the same name in the 25th round. It is possible that within the 25 rounds, the competition has broken in favor of a single name, and eventually all participants would adopt it. However, these cases can be hard to distinguish from a metastable state of competing but evolving conventions. Alternatives might simultaneously gain purchase in new clusters while losing others, leaving the competition very much intact. Such shifting could look like the beginning of a convergence dynamic, but not in fact be that.[15] For this reason, each games will be coded as having a single convention only if all 24 participants played the same name in the last round.

The analysis of algebraic connectivity suggests that if participants treat spectating-linked names the same way they treat names from partners, all trials in the spectating treatments will result in a single convention. Undercutting this argument is the possibility that participants treat spectating-linked names differently in different networks. I argue that the higher the clustering coefficient, the more likely participants are to ignore spectating-linked names to focus on matching within their cluster. The effect of this behavior is to reduce the chances that a single convention will emerge. These points result in the following specific claims:

Claim 1. Increasing the number of spectating-linked names seen each round increases the number of conventions that emerge.

Claim 2. The number of spectating-linked names seen in each round necessary for a

[15] *CB* argues that the metastable states will resolve to a global convention given enough time, but it is not clear whether human cognition supports that conclusion.

convention to emerge increases with the clustering coefficient of the underlining network topology.

I use the term claim, not hypothesis, because conducting a sufficient number of trials to make statistical hypothesis testing possible is not practical.

Experiment Results and Analysis

To test the above claims, I conducted an IRB-approved study (institution and protocol number withheld for review) with participants recruited from Amazon's Mechanical Turk platform. The pool of participants was limited to those identified as residing in the United States. The game interface was created using *oTree*, an Python platform for the implementation of experiments (Chen et al., 2016).

Each network requires 24 concurrent participants, so I use a two-stage panel recruitment method (Mason & Watts, 2012). I first post a compensated recruitment task that introduced participants to the game and let them experience the interface. At the end of the task, they have the option to opt in to receiving notifications about games in the future. This allows the recruitment of a suitable number of participants to play the game at the same time and establish control over who is joining the game. Details on the creation and management of this panel, including any attempts to coordinate through means outside the game, are detailed in the Supplemental Information.

Upon arriving to play the game, participants read a brief description of the game and compensation rules. After reading a risk statement, they either consent to be included in the study or leave it. Those who consent then read the complete game instructions and answer two comprehension questions. They are placed in a "waiting room" until a sufficient number of participants are ready. A chime and a message alert them that the game is about to begin.

I first replicate the results in *CB* to confirm the commensurability of the game interface despite minor differences in visual appearance. The first row of Table 2 reveals a full replication of the results reported in the original study; groups embedded in fully-connected networks succeeded in finding a global convention, but those in the other topologies did not. I also include small-world networks, which were not tested in the original experiment, to further validate the results.

Rows 2 and 3 in Table 2 represent the results of the spectating treatments (i.e., one and two spectating-linked names). The evidence supports the first claim—spectating-linked names are beneficial for coordination—and the second claim—the effect is attenuated as the degree of clustering in the underlying network increases. The addition of spectating-linked names did not disrupt the consensus building process in the fully-connected networks. The availability of a single spectating-linked name leads two random-regular and one small-world networks to a single convention. One more trial for both random-regular and small-world networks would likely have reached a global convention if the game had lasted another several rounds, as competition had likely broken in favor of a single name. However, the game ended before all participants adopted it and therefore these trials did not meet the criteria established above for having a single convention. For two additional names, all random-regular and small-world networks converged, and a single lattice network converged as well.

[Table 2 about here.]

The results in 2 make clear that the coefficient of algebraic connectivity was not wholly determinative of the outcome; the higher the clustering coefficient for the network, the more spectating-linked names needed to tip the group toward convergence on a single convention. Although adding spectating-based ties significantly increases the connectivity, participants are still only incentivized to coordinate with their network neighbors. This lack of incentive can be seen in Table 3, which classifies the source of a name that a participant played for the first time. A significant number of names are categorized as *self-introduced*. The initial name every participant plays is necessarily self-introduced, but many participants play another name they have not seen. Names first seen and played because the focal participant's partner played it are classified as *alters only*. If the name was only seen via spectating, it is classified as *spectating only*. If it was visible through spectating and an alter before being played, it is *alters & spectating*. Classification is determined by whether that name appeared via the source in any period before being played by the participant for the first time. In practice, it was played within the last three rounds.

[Table 3 about here.]

For each condition, Table 3 reports the average number of names of a given type played by participants. Also reported (within parentheses) is the percentage of times the play of those names lead to a match with the respective partner, in the present round and the subsequent rounds it is played. Participants do play names they have only seen via spectating, and the probability that those result in a match is generally greater than names from the *alters only* source. The exception is for the small-world and lattice networks for one additional name. There, the observed behavior suggests that the participants are more intent on mirroring what their partners play and are more successful for it; the increased clustering in those networks creates a self-reinforcing dynamic in which it is better to focus on alters' names because that leads to high rates of success within that cluster.

The best strategy is to play names seen both via alters and spectating sources (i.e., *Alters & Spectating*). Participants do so regularly and often match. Part of this success is because that behavior comes later in the game, when the relevant name emerges as the winning convention. However, the matching rate is high when a single convention does not emerge, suggesting that those trials might have converged if participants had been more responsive to the successes originating from observation of spectating-linked names.

This can be seen in Figure 1, which details the dynamics of a selection of trials. The plot shows the number of different names in circulation, the percentage of interactions that successfully matched, and the concentration of the use of names. Concentration is defined by the normalized Herfindahl-Hirschmann index (HHI), which captures how consolidated or concentrated a "market" is and reaches its maximum of one when there is a single type left. Because participants sometimes alternate names between rounds, the name counts are smoothed over the previous and subsequent rounds. In all trials, the number of different names decreases from slightly more than one per participant to between 12 and one. This indicates at least a moderate amount of local coordination and an increase in the concentration of names, meaning the trials that did not result in a global convention nevertheless did succeed in creating competing local conventions.

[Figure 1 about here.]

The dynamics of change in HHI reveal a tight correspondence between greater concentration in the early rounds and the establishment of a single convention. This is true for the trials not shown as well. It indicates the means by which exposure to spectating-linked names aids in successful coordination; behavior that is responsive to the wider range of names seen through spectating tends to eliminate possible alternatives from circulation. This does not affect the rate of successful matching because there is hardly any in the early rounds. However, once there is enough regularity in choices that participants start to match, there are far fewer names in circulation than there would be in the absence of the spectating-linked names. So, while there are incentives for local coordination, the rapid reduction in practically available alternatives before those incentives can be realized can allow groups to circumvent much of the competition that comes from local successes.

This mechanism does not fit the explanation given in *CB* for the success of fully-connected trials–early coordination failure leads to later success–because failure is neither necessary nor sufficient. Network treatments with no additional names exhibit early coordination failure and do not converge to a single convention, and trials with additional names can experience local success early on and still converge to a global convention as long as the concentration of names is also rapidly increasing. See the center and right panels in the second row for an example. This change in concentration is a better indicator of eventual success than the local success rate or the total number of names in circulation. Local success rates can be quite variable and the count of names can decrease substantially without much change in the concentration because a few entrenched alternatives with equal frequencies also corresponds to a low concentration score.

Together, these observations strongly suggest that the availability of spectating-linked names can affect coordination by increasing the rate at which alternatives are eliminated from the system. By creating unsystematic exposures to the diversity of alternatives within the population, spectating-linked names can tip the system toward convergence. In this experimental design, however, the additional names are also simply additional information. In order to understand whether the spectating-linked information is qualitatively distinct, the next section considers the value of each additional name and compares them to other potential source types.

Informational Content of Spectating-based Information in the Name Game

To understand if the observed results are driven by the unique properties of a spectating-linked information rather than simply the fact that participants see more names, I first use simulations to quantitatively compare the informational content of names from alternative sources. If those alternatives contain the same amount of information as the spectating-linked source, it would imply that spectating is not a distinct mechanism, albeit a cognitively plausible one. If the alternatives are less informative, spectating behaviors constitute a distinct mechanism of information diffusion and influence.

I consider several alternative sources from which participants could see additional names: existing neighbors (alters), a non-alter selected at random and viewed for the entire 25 rounds, the non-alter at the end of the longest possible bridging tie, and the maximally informative non-alter. The actual names played in the games are used to construct synthetic sources for comparison. This facilitates the quantification of the counterfactual amount of information would have been gained by seeing the names from the given source at any time step.

The *longest bridge* source captures the fact that bridging ties lead to important information gains. The theory of weak ties maintains that such ties are most likely to yield valuable information, but as a primarily qualitative assessment of interactions, tie weakness is not a suitable basis for constructing comparisons. The framework of structural holes and bridging ties (Burt, 1992) instead provides a clear criterion for thinking about potential sources of valuable information. Within the networks used in this experiment, there are only local bridging ties and many of them. To make the comparison as competitive as possible, I identify the bridge with the longest shortest path to the focal participant, the structurally most distant person in the network. A direct tie to that most distant person becomes the *longest bridge*. When two longest bridges are added, the second is identified within the network with the first longest bridge included.

The alternative sources have a range of potential information amounts. Alters with whom the participant is not actively partnered with (i.e., *Random Alters*) are likely to provide little additional information because the names are likely to have already been seen. Exposure to names played by randomly selected (and static) non-alters (*Random Non-alters*) should initially yield more information, but become more predictable and less informative as play progresses. The names played by the *longest bridge* alter should have a high information content. Similarly to the random non-alters, however, that amount will decline over time. Finally, to put an upper bound on information content of additional names, the final comparison is with the most informative names, measured *post hoc*, each round (*Max. Info. Others)*.

To compare the informational content of these alternative sources with the spectating-linked information, I calculate the Kullback-Leibler divergence measure (Kullback & Leibler, 1951; Lin, 1991) from the hypothetical distribution of names each participant would have seen if exposed to the alternative source to the true distribution of names at each round of the game. I also calculate the divergence from the observed distributions that include spectating-linked names to the true distribution. The supporting information details the divergence measure and construction of the alternative sources.

Figure 2 plots the average and standard errors of the differences in participants' divergences between the spectating-based information treatment and the respective comparison sources. Where the difference is positive, the spectating-linked names yield more information than the alternative. The mean of the difference in divergences is greater than zero for nearly all rounds for the comparisons to the existing alters, new non-alters, and longest bridges. The spectating source gains only less than the maximally informative source. Differences are minimal initially because the low concentration among the alternatives means almost any additional name is equally informative. From there, the comparative informativeness of spectating builds. In trials in which a convention does emerge, this advantage appears to decline over time as the increasing concentration of names reduces the total amount of information necessary to describe the distribution. In cases with no emergent convention, the advantage is maintained, although the information is not enough to overcome the now-entrenched alternatives.

[Figure 2 about here.]

These differences provide another angle on the uniqueness of spectating behaviors. As constructed in this experiment, it yields more information than a range of other possible sources, and often yields close to the maximum amount for a sampling-based method. However, this simulation approach is only illustrative because in a dynamic process the true distribution of names changes based on participants' exposure to names. It is likely that the distribution of available names would evolve differently if participants actually saw and used them. To ensure a more rigorous comparison, I run an additional set of trials. Given the theoretical importance of the weak ties/structural holes argument, I focus on the longest bridges source. It is likely that real bridging ties do not span structural holes to the most structural distant node, but focusing on the most distant makes for the strongest comparison. The maximum information comparison is substantively interesting, but, as a dynamic post-hoc measure, challenging to implement within a live experiment.

Table 4 displays the results of trials with a setup identical to the main experiment except that the additional names shown are those played by the structurally most distant non-alter, in other words the *longest bridge* source. Note that the ties need not be reciprocal and that the distant non-alter can be observing the name played by someone other than the focal participant. The simulation results in Figure 2 show that the *longest bridge* source contains less information than the spectating-linked source. For a single name, the difference is small for random-regular networks and larger for the small-world networks. For two names, the differences are positive but minimal for all network types. The simulations imply that small-world network trials are unlikely to reach a global convention for a single additional name but should do so for two names. Random-regular networks should mirror the results for the spectating-linked source in Table 2.

[Table 4 about here.]

The *longest bridge* source trials result in the same counts of conventions as the random regular networks with the spectating-linked names; two of four trials with one additional name converged, while all four trials with two names converged. Only a single small-world trial resulted in a convention however, and a detailed analysis suggests that outcome is the result of good fortune; two structurally distant participants started with the same name and were paired with others in a way that led to quick local agreement and a single name dominating early. Furthermore, the fact that none of the trials with two names converged highlights the degree to which the one that did converge is an aberration.

These results provide further evidence of the special properties of spectating behaviors, but also of the limitation of the simulation approach. In particular, the simulation suggests that small-world networks with two additional longest bridge names should converge, but they do not. This gap is likely due to the fact that all four trials with the spectating source converged, which means the system becomes easier to describe and any source of names reflects the true state of the system. In the trials, small-world networks have 4 to 6 visually discernible clusters. Thus, even two additional longest bridges will fail to provide information about the behavior in one or more clusters. Spectating samples at random from the whole population, gains information about behavior in all clusters, and therefore is more likely to overcome the competition induced by clustering.

Discussion

The primary experiment detailed above tested whether spectating-linked information has the potential to facilitate group-wide coordination on a single convention when the network structure is otherwise an impediment. The results demonstrate that this potential is very real; many groups converge to a single convention within the time frame provided, although additional names are needed as clustering in the network increases.

The proximate cause of convergence is the rapid reduction in the number of alternatives in circulation within the first 10 rounds of game play, see Figure 1. By individually seeing more of the names in circulation, participants collectively reduce the number of alternatives in competition, but not evenly. In the early stages of those trials, one name starts to win out. Graph theory would predict this dynamic on the basis of the coefficients of algebraic connectivity associated with the respective networks, see Table 1. However, that approach does not consider individuals' incentives and, in the case of conventions, actors must balance a direct, local incentive to match with partners against the indirect, global incentive to match with everyone. Analyzing how participants thought about incentives is beyond the scope of this paper, but it seems likely that until there is regularity in the names they see from their partners, the local incentives are indistinguishable from the diffuse and underspecified incentive to coordinate more broadly. Many participants do attend to the spectating-linked names and make choices in response to those exposures. Doing so is an important driver of the chances of matching for the entire group (see Table 3).

The strength and timing of the emergence of the local incentives is related to the clustering coefficient of the network. When clustering is high, there is more triadic closure–groups of three for which all three possible pairs are connected. The upshot is an increase in the probability that a participant will be partnered with a neighbor who recently partnered with a common connection and will see a familiar name. Participants do not know this, but they see more consistent usage of names. This gives them the information to act on the explicit incentive structure earlier in the game. This is why algebraic connectivity is not fully predictive of outcomes and why participants play spectating-linked names less in the networks with higher clustering coefficients. However, their playing of spectating-linked names can make a difference in whether the entire group converges to a single convention. In the earliest rounds, seeing two spectating-linked names can give glimpses of regularities that participants then act on. This does not change the local incentives, but, once there are regularities with partners, it is more likely that different clusters are using the same names.

This dynamic raises the question of whether being exposed to any additional names could have the same effect. I explored this question first by running counterfactual simulations using the observed distributions of names. The information gain from observing the names linked to the spectating is greater than that from observing the names of even the most structurally distant participant (i.e., *longest bridge* source). Observing spectating-linked names is often competitive with observing the names that would be the most informative names given what happens later (i.e., *max info*).

To validate the insights from the simulation approach, I ran trials of the same experiment but with the additional name(s) coming from the *longest bridge* source. The outcomes are identical for random-regular networks, but spectating maintains an advantage over the *longest bridge* source in small-world networks, which have a higher degree of clustering. This is consistent with the differences in information gain found in the simulations.

Together, these results indicate that spectating behaviors have the informational potential to tip competition toward coordination, and furthermore that participants are willing and able to attend to spectating-linked sources. There are two facets of the methodological approach taken here that might limit the external validity of the conclusions. First, the pattern of spectating in the experiment is minimally structured. If people are able to spectate on a more limited segment of the underlying network, doing so will be less informative than the design implemented here. Whether adding such biases reduces information enough to alter the outcomes is likely conditional on the particular biases introduced to the patterns of spectating. More research is needed on the effects and realism of different patterns of spectating.

The second limitation is the necessity of explicit incentives within the game. This choice is in accordance with the literature treating the emergence of conventions as the successful outcome of a coordination game, but it is not clear that individuals feel incentives as modeled in the game. They likely feel the inconvenience of a lack of a convention, but it is not clear that they feel positively motivated to solve this problem instead just of avoiding similar scenarios in the future. That said, there is good reason to believe that a great deal of conventions must come about through the force of precedent mechanism and that any sort of re-enforcement learning framework that rewards the precedent of successful outcomes will look something like the incentives as structured in the Name Game.

Conclusion

Conventions are a constitutive feature of human societies, yet how exactly they come about is not well understood. Starting with Hume, there is an assumption that a *force of precedent* mechanism plays a crucial role; when a behavior has worked in the past, it is reasonable to behave the same way in the future. Prior work has shown the plausibility of this mechanism to drive convergence on a single widely-held convention, but has neglected a crucial feature of the structure of interactions. The inclination for acquaintances to share acquaintances, combined with a natural limit to the number of acquaintances an individual has, leads to clustering in social networks.

It is well known in the network dynamics literature that when individuals' knowledge is formed exclusively from information received through interactions with network alters, clustering inhibits convergence or alignment. The straightforward implication of this fact is that network-based processes are not alone sufficient for groups to rely on the force of precedent mechanism to create conventions. Clustering induces competition between different alternatives because individuals are incentivized to heed the dominant behavior in their primary cluster. Rather than abandon the force of precedent mechanism, I have argued that *spectating* is a plausible set of behaviors that can facilitate the emergence of widely held conventions. Other mechanisms could yield the same result, but spectating corresponds to known cognitive processes, sociological theory, and intuitions.

Various social theories have pointed to the importance of spectating–observation without direct engagement–but there has not been an attempt to assess whether such behaviors have the potential to transform processes of self-organization. This paper does not offer a systematic review of these theories. Instead, it focuses on their cognitive plausibility and what they might offer in terms of information. There is evidence that we engage in a variety of spontaneous social inferences using stimuli received via spectating, and the evolution of such faculties suggests that those stimuli provide valuable information.

To test whether that information can allow the force of precedent mechanism to overcome the frictions introduced by clustering, the paper extends the *Name Game*, a large group human-subject game previously used for modeling the formation of conventions (Centola & Baronchelli, 2015), to capture the effects of availability of a spectating-linked names. Large groups, each with 24 participants, play the game for 25 rounds, and when they are shown names from individuals from all across the network, they are often able to converge on a single convention even when there is clustering in the network.

This result suggests that, while direct interaction is the bread and meat of social life and rightfully the subject of research, spectating might be the condiment necessary to make social life cohere. Conventions are a meaningful domain in which to consider this possibility, but spectating behaviors could play an important role in any number of social influence processes. It sits between the extremes of the dominant models of information sharing: network-ties-only or a global aggregation and broadcast. The reach of spectating behaviors is greater than that of social networks because social networks overlap in physical and digital space, but also does not require the same level of assumed infrastructure as broadcast mechanisms. Given this demonstration of its potential, the particulars of spectating behaviors merit more attention.

References

Adut, A. (2012). A Theory of the Public Sphere. *Sociological Theory*, *30*(4), 238–262.

Adut, A. (2018). *Reign of Appearances: The Misery and Splendor of the Public Sphere* (1 ed.). Cambridge University Press.

Anderson, E. (2011). *The cosmopolitan canopy: race and civility in everyday life*. New York, NY: W.W. Norton & Co.

Arendt, H. (1998). *The Human Condition* (2nd ed ed.). Chicago: University of Chicago Press.

Axelrod, R. (1997). The Dissemination of Culture: A Model with Local Convergence and Global Polarization. *Journal of Conflict Resolution*, *41*(2), 203–226. ISBN: 0893-3200 _eprint: 0803973233.

Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, *2006*(06), P06014–P06014.

Baronchelli, A., Loreto, V., & Steels, L. (2008). In-Depth Analysis of the Naming Game Dynamics: the Homogeneous Mixing Case. *International Journal of Modern Physics C*, *19*(05), 785–812.

Bednar, J., Bramson, A., Jones-Rooy, A., & Page, S. (2010). Emergent cultural signatures and persistent diversity: A model of conformity and consistency. *Rationality and Society*, *22*(4), 407–444.

Bednar, J. & Page, S. E. (2007). Can Game(s) Theory Explain Culture? *Rationality and Society*, *19*(1), 65–97. ISBN: 1043463104044 _eprint: 1112.1699.

Bertsekas, D. P. & Tsitsiklis, J. N. (2007). Comments on "Coordination of groups of mobile autonomous agents using nearest neighbor rules". *IEEE Transactions on Automatic Control*, *52*(5), 968–969.

Biggart, N. W. & Beamish, T. D. (2003). The economic sociology of conventions: Habit, Custom, Practice, and Routine in Market Order. *Annual Review of Sociology*, *29*, 443–464.

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method* (1998, Univ ed.).

Engelwood Cliffs, NJ: Prentice-Hall.

Browning, C. R., Calder, C. A., Soller, B., Jackson, A. L., & Dirlam, J. (2017). Ecological Networks and Neighborhood Social Organization. *American Journal of Sociology*, *122*(6), 1939–1988.

Burt, R. S. (1987). Social Contagion and Innovation: Cohesion versus Structural Equivalence. *American Journal of Sociology*, *92*(6), 1287. ISBN: 00029602 _eprint: /www.jstor.org/stable/2779839.

Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, Massachusetts: Harvard University Press.

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156–2160.

Centola, D. (2015). The Social Origins of Networks and Diffusion. *American Journal of Sociology*, *120*(5), 1295–1338. ISBN: 00029602.

Centola, D. & Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, *112*(7), 1989–1994.

Chandrasekhar, A. G., Larreguy, H., & Xandri, J. P. (2015). Testing Models of Social Learning on Networks: Evidence From a Lab Experiment in the Field. *National Bureau of Economic Research*.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. Publisher: Elsevier B.V.

Chen, Z. & Whitney, D. (2019). Tracking the affective state of unseen persons. *Proceedings of the National Academy of Sciences*, 201812250.

Chwe, M. S.-Y. (2001). *Rational Ritual: Cultural, Coordination and Common Knowledge*. Princeton: Princeton University Press.

Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American journal* *of Sociology*, *91*(6), 1309–1335. ISBN: 0002-9602.

Cowan, B. (2004). Mr. Spectator and the Coffeehouse Public Sphere. *Eighteenth-Century Studies*, *37*(3), 345–366.

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PloS one*, *8*(3), e57410. ISBN: 1932-6203.

Dawkins, R. (1976). *The Selfish Gene*. Oxford, UK: Oxford University Press.

Degroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, *69*(345), 118–121.

DellaPosta, D., Shi, Y., & Macy, M. (2015). Why Do Liberals Drink Lattes? *American Journal of Sociology*, *120*(5), 1473–1511.

Delvaux, V. & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, *64*(2-3), 145–173. ISBN: 3265373142.

Derex, M. & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proc. Natl. Acad. Sci. USA*, *113*(11).

Duff, K. J. & Newman, L. S. (1997). Individual Differences in the Spontaneous Construal of Behavior: Idiocentrism and the Automatization of the Trait Inference Process. *Social Cognition*, *15*(3), 217–241.

Easley, D. & Kleinberg, J. (2010). *Networks, crowds, and markets: reasoning about a highly connected world*. New York: Cambridge University Press. OCLC: ocn495616815.

Elliot Davis, L. (2024). Current Trends in Fee Structures.

Elster, J. (1989). *The Cement of Society: a Study of Social Order*. Cambridge University Press.

Emberson, L. L., Lupyan, G., Goldstein, M. H., & Spivey, M. J. (2018). Overheard Cell-Phone Conversations: When Less Speech Is More Distracting. *Psychological Science*, *21*(10), 1383–1388.

Emirbayer, M. & Goodwin, J. (1994). Network Analysis, Culture, and the Problem of Agency. *American Journal of Sociology*, *99*(6), 1411–1454.

Emirbayer, M. & Mische, A. (1998). What Is Agency? *American Journal of Sociology*, *103*(4), 962–1023.

Evans, J. S. B. T. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology*, *59*(1), 255–278.

Fine, G. A. (1979). Small Groups and Culture Creation: The Idioculture of Little League Baseball Teams. *American Sociological Review*, *44*(5), 733.

Fine, G. A. (2012). Group culture and the interaction order: Local sociology on the meso-level. *Annual Review of Sociology*, *38*, 159–179.

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten Years later. In S. Chaiken & Y. Trope (Eds.), *Dual Process Theories in Social Psychology*. New York: Guilford Press. Section: 11.

Fiske, S. T. & Taylor, S. (2017). *Social Cognition: From Brains to Culture* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Flache, A. & Macy, M. W. (2011). Local Convergence and Global Diversity: From Interpersonal to Social Influence. *The Journal of Conflict Resolution*, *55*(6), 970–995.

Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., & Lorenz, J. (2017). Models of social influence: Towards the next frontiers. *Jasss*, *20*(4).

Fligstein, N. & McAdams, D. (2012). *A Theory of Fields*. Oxford, UK: Oxford University Press.

Forrest, R. & Kearns, A. (2001). Social Cohesion, Social Capital and the Neighbourhood. *Urban Studies*, *38*(12), 2125–2143.

Friedkin, N. E. & Johnsen, E. C. (2011). *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. New York: Cambridge University Press.

Garrod, S. & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218.

Garrod, S. & Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, *53*(3), 181–215.

Garrod, S. & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive*

Sciences, *8*(1), 8–11.

Glass, R. J., Glass, L. M., Beyeler, W. E., & Min, H. J. (2006). Targeted Social Distancing Design for Pandemic Influenza. *Emerging Infectious Diseases*, *12*(11).

Goffman, E. (1959). *Presentation of Self in Everyday Life*. Anchor.

Golub, B. & Jackson, M. O. (2010). Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*, *2*(1), 112–149.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, Massachusetts: Harvard University Press.

Habermas, J. (1991). *Habermas, Junger, The Structural Transformation of the Public Sphere, Cambridge,The MIT Press, 1991, pp. 301.*, volume 68. Cambridge, Massachusetts: MIT Press. ISSN: 00308684 Publication Title: an Inquiry Into a Category of Bourgeois Society _eprint: arXiv:1202.3162v2.

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In *Proceedings of the 7th Python in Science Conference*, (pp.5̃).

Ham, J. & van den Bos, K. (2008). Not fair for me! The influence of personal relevance on social justice inferences. *Journal of Experimental Social Psychology*, *44*(3), 699–705.

Ham, J. & Vonk, R. (2003). Smart and easy: Co-occurring activation of spontaneous trait inferences and spontaneous situational inferences. *Journal of Experimental Social Psychology*, *39*(5), 434–447.

Happé, F., Cook, J. L., & Bird, G. (2017). The Structure of Social Cognition: In(ter)dependence of Sociocognitive Processes. *Annual Review of Psychology*, *68*(1), 243–267.

Harary, F. (1959). On the measurement of structural balance. *Behavioral Science*, *4*(4), 316–323.

Hassin, R. R., Aarts, H., & Ferguson, M. J. (2005). Automatic goal inferences. *Journal of*

Experimental Social Psychology, *41*(2), 129–140.

Hatano, Y. & Mesbahi, M. (2005). Agreement over random networks. *IEEE Transactions on Automatic Control*, *50*(11), 1867–1872.

Hauser, D. J. & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*(1), 400–407. ISBN: 9783642134371 _eprint: arXiv:1011.1669v3.

Hechter, M. & Opp, K.-D. (2001). Introduction. In K.-D. Opp & M. Hechter (Eds.), *Social Norms*. Russell Sage Foundation.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.

Hume, D. (1888). *A Treatise on Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. Oxford University Press.

Ipeirotis, P. (2010). Demographics of Mechanical Turk. *Working Paper CeDER-10-01*. ISBN: NYU Working Paper No. CEDER-10-01.

Jacobs, J. (1992). *The death and life of great American cities* (Vintage Books ed ed.). New York: Vintage Books.

Jones, N. (2022). How a Nepo Baby Is Born Hollywood has always loved the children of famous people. In 2022, the internet reduced them to two little words. *Vulture*.

Judd, S., Kearns, M., & Vorobeychik, Y. (2010). Behavioral dynamics and influence in networked coloring and consensus. *Proc. Natl. Acad. Sci. USA*, *107*(34), 14978–14982. ISBN: 1104407108.

Keeling, M. J. (2011). The effects of local spatial structure on epidemiological invasions. *The Structure and Dynamics of Networks*, *9781400841*, 480–488. ISBN: 9781400841356.

Kelly, S. D., Creigh, P., & Bartolotti, J. (2010). Integrating speech and iconic gestures in a stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience*, *22*(4), 683–694.

Krull, D. S. & Dill, J. C. (1996). On Thinking First and Responding Fast: Flexibility in Social Inference Processes. *Personality and Social Psychology Bulletin*, *22*(9), 949–959.

Kullback, S. & Leibler, R. (1951). On information and sufficiency. *Annals of Mathematical*

Statistics, *22*, 79–86.

Lahire, B. (2011). *The Plural Actor*. Cambridge, UK: Polity Press.

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature, Speech, and Communication*. Cambridge, Massachusetts: MIT Press. Publication Title: Book.

Lewis, D. (1969). *Convention: A philosophical study*. Oxford, UK: Blackwell.

Lin, J. (1991). Divergence Measures Based on the Shannon Entropy. *IEEE Transactions on Information Theory*, *37*(I), 145–151.

Lizardo, O. (2004). The cognitive origins of Bourdieu's Habitus. *Journal for the Theory of Social Behaviour*, *34*(4), 375–401.

Martin, J. (2003). What Is Field Theory? *American Journal of Sociology*, *109*(1), 1–49.

Mason, W. & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1–23.

Mason, W. & Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, *109*(3), 764–769.

Mason, W. A., Conrey, F. R., & Smith, E. R. (2007). Situating Social Influence Processes: Dynamic, Multidirectional Flows of Influence Within Social Networks. *Personality and Social Psychology Review*, *11*(3), 279–300.

Mead, G. H. (1934). *Mind, Self and Society*. Chicago: University of Chicago Press.

Melnikoff, D. E. & Bargh, J. A. (2018). The Mythical Number Two. *Trends in Cognitive Sciences*, *22*(4), 280–293.

Miller, J. C. (2009). Percolation and epidemics in random clustered networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *80*(2), 1–4.

Moore, C. & Newman, M. E. J. (2000). Epidemics and percolation in small-world networks. *Physical Review E*, *61*(5), 5678–5682.

Moussaïd, M., Kämmer, J. E., Analytis, P. P., & Neth, H. (2013). Social influence and the collective dynamics of opinion formation. *PLoS ONE*, *8*(11). _eprint: 1311.3475.

Newman, D. (2005). Introduction. In D. J. Newman (Ed.), *The Spectator: Emerging Discourses*. Cranbury, NJ: Rosemont Publishing & Printing Group. Section: 1.

Newman, M. E. J. (2003a). Mixing patterns in networks. *Physical Review E*, *67*(2), 026126.

Newman, M. E. J. (2003b). Properties of highly clustered networks. *Physical Review E*, *68*(2), 026121.

Newman, M. E. J. (2010). *Networks: an introduction*. Oxford ; New York: Oxford University Press. OCLC: ocn456837194.

Olfati-Saber, R. (2005). Ultrafast consensus in small-world networks. In *Proceedings of the 2005, American Control Conference, 2005.*, (pp. 2371–2378)., Portland, OR, USA. IEEE.

Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and Cooperation in Networked Multi-Agent Systems. *Proceedings of the IEEE*, *95*(1), 215–233.

Olfati-Saber, R. & Murray, R. (2004). Consensus Problems in Networks of Agents With Switching Topology and Time-Delays. *IEEE Transactions on Automatic Control*, *49*(9), 1520–1533.

Onishi, K. H. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, *308*(5719), 255–258.

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision making*, *5*(5), 411–419. ISBN: 1930-2975.

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, *87*(3), 925–979. _eprint: 1408.2701.

Patterson, O. (2014). Making Sense of Culture. *Annual Review of Sociology*, *40*(1), 1–30.

Portes, A. (1998). Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*.

Powell, M. N. (2012). See No Evil, Hear No Evil, Speak No Evil: Spectation and the Eighteenth-Century Public Sphere. *Eighteenth-Century Studies*, *45*(2), 255–276.

Rand, D. G. (2012). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of Theoretical Biology*, *299*, 172–179. ISBN:

0022-5193 Publisher: Elsevier.

Rawlings, C. M. & Childress, C. (2019). Emergent meanings: Reconciling dispositional and situational accounts of meaning-making from cultural objects. *American Journal of Sociology*, *124*(6), 1763–1809.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *NeuroImage*, *28*(4), 848–858.

Santos, F., Rodrigues, J., & Pacheco, J. (2006). Graph topology plays a determinant role in the evolution of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1582), 51–55.

Schelling, T. C. (1960). *The Strategy of Conflict*. Harvard Univ. Press.

Serrano, M. & Boguñá, M. (2006). Percolation and epidemic thresholds in clustered networks. *Physical Review Letters*, *97*(8), 1–4. _eprint: 0603353.

Shank, D. B. (2016). Using Crowdsourcing Websites for Sociological Research: The Case of Amazon Mechanical Turk. *American Sociologist*, *47*(1), 47–55. ISBN: 1210801592669.

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press.

Smith, V. L. (1976). Experimental Economics: Induced Value Theory. *The American Economic Review*, *66*(2), 274–279.

Sperber, D. & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind and Language*, *17*(April), 3–23.

Strauss, C. & Quinn, N. (1997). *A Cognitive Theory of Cultural Meaning*. Cambridge, UK: Cambridge University Press.

Tahbaz-salehi, A. & Member, S. (2008). A Necessary and Sufficient Condition for Consensus Over. *Nonlinear Dynamics*, *53*(3), 791–795.

Tavory, I. (2018). Between Situations: Anticipation, Rhythms, and the Theory of Interaction. *Sociological Theory*, *36*(2), 117–133. ISBN: 0735275118777.

Templeton, B. (2001). *Origin of the term "spam" to mean net abuse*. Publication Title:

www.templetons.com.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, Massachusetts: Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691. Publisher: Robert Crown Law Library.

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. ISBN: 9780824302597.

Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as Flexible Interpreters: Evidence and Issues from Spontaneous Trait Inference. In *Advances in Experimental Social Psychology*, volume 28 (pp. 211–279). Elsevier.

Watts, D. J. (1999). Networks, Dynamics, and the Small-World Phenomenon. *American Journal of Sociology*, *105*(2), 493–527.

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*, 3.

White, H. C. (2008). *Identity and Control*. Princeton: Princeton University Press.

Whitney, D. & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology*, *69*, 105–129. ISBN: 978-0-19-973433-7.

Willems, R. M. & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: A review. *Brain and Language*, *101*(3), 278–289.

Wilson, D. & Sperber, D. (2002). Relevance Theory. In G. Ward & L. Horn (Eds.), *Handbook of Pragmatics*. Blackwell.

Wu, C. W. (2005). Algebraic connectivity of directed graphs. *Linear and Multilinear Algebra*, *53*(3), 203–223.

Young, H. P. (1993). The Evolution of Conventions. *Econometrica*, *61*(1), 57–84.

Young, H. P. (2015). The Evolution of Social Norms. *Annual Review of Economics*, *7*, 359–387. ISBN: 1941-1383 1941-1391.

Zacka, B. (2020). Chapter 4: What's in a Balcony? The In-Between as Public Good. In D. Bell & B. Zacka (Eds.), *Political Theory and Architecture* (1 ed.). Bloomsbury Publishing Plc.

Zerubavel, E. (1999). *Social Mindscapes: An Invitation to Cognitive Sociology*. Cambridge, Massachusetts: Harvard University Press.

Appendix

Administration of Experiment

All subjects were recruited through Amazon's Mechanical Turk marketplace (Mturk). This platform enables the recruitment of workers to complete online tasks and has been validated as a source of subjects for a wide range of behavioral experiments (Paolacci et al., 2010; Mason & Suri, 2012; Rand, 2012; Casler et al., 2013; Crump et al., 2013; Hauser & Schwarz, 2016; Shank, 2016). To build the subject pool, any worker with a reasonably healthy history of work on Mturk was offered compensation to complete a short training module and take a comprehension test. The task paid well for the amount of time it took, so it was easy to recruit workers into the subject pool. This initial interaction familiarized subjects with the interface and game structure, but also satisfied Mturk's requirement that posters of tasks only directly communicate with workers they have interacted with in the past. This allowed the creation of a well-defined subject pool and control over which workers were permitted to join live sessions of the experiment. The day before a scheduled trial, a number of subject-pool members were invited to play. Following the recommendation of Mason and Suri (2012), approximately four times as many members of the pool were invited than necessary to carry out the experiment. Those invitees were contacted shortly before the scheduled game as a reminder and again at the start of each game to provide a link to the interface.

The subject pool included slightly more than 300 members at any given time, as members were periodically removed for not responding to invites and then replaced. Pool members were allowed to complete up to five sessions; the average number of sessions completed was 2.95, with an average of 20 days between sessions. For each game a list of invitees was created from this pool using the following criteria; no participants from the last session can be included; those who were recently invited but did not start a session receive the highest priority; those who have yet to be invited get the next highest priority; those who have been invited but have not played get the next highest priority; for the remaining pool members, priority is less than the previous categories and inversely related to how recently they completed a session. Because the set of members of equal priority was larger than was necessary to create the invitee list, the correct number of members was added by sampling uniformly from that set. The above procedure was repeated until no more than four participants who completed the same previous session were included.

Game Play Details

The game interface was created using oTree, an open source Python platform for experiments (Chen et al., 2016). As participants arrive for a session, a brief description of the task and the IRB risk statement are shown (Exempt status, [DETAILS OMITTED FOR REVIEW]). If they accept the task, they are again shown the game instructions, which include a small comprehension test. All subject pool members have previously read the directions and have had a chance to interact with the interface, but this step serves as an important refresher. Next, they must agree not to use means outside the interface to attempt to coordinate. This establishes grounds for dismissal, but the real controls against such attempts are built into the software (see Software and Subject Management Details). Once they agree, they are taken to a waiting page until enough participants arrive. Subjects are paid one cent for every five seconds of waiting. They are shown how long they have been waiting and their compensation for that wait time. This greatly increases the retention of subjects and their satisfaction with the task and compensation, which is crucial to maintaining a responsive pool.

Once enough participants have arrived, the game begins. Participants are shown a headshot of an individual and asked to submit a name for her. Participants have 15 seconds to submit it, and failure to do so in that time results in their submission being recorded as "None". Figure 3 is an example of this submission page. After everyone has submitted, the results are displayed. This differs from the design of the original experiment, which allowed for asynchronous progression through rounds. Although it seems unlikely to make a difference, synchronizing the rounds was necessary in this design because the second treatment factor required sampling from all names played in a round. To maintain commensurability, rounds were synchronized for all treatment levels, a choice that ultimately did not have an effect on coordination outcomes.

Partners who successfully coordinated earn $0.10. Partners who did not coordinate are penalized \$0.05, unless their current cumulative total is already \$0.00. The results page always shows the names each partner played, whether they matched, the amount of the reward or penalty, their current cumulative rewards, and their history of successes or failures through the previous twelve rounds (but not the names associated with those successes or failures). In the treatments where it was applicable, additional names were displayed. Figure 4 is an example of the results page. This page times out after 10 seconds. The progression through the submission, waiting, and results pages constitutes a round of play. There are 25 such rounds in each session. After the 25 rounds, the final page reported the participant's total earnings, broken down into base pay, waiting bonus, and coordination-based bonuses. Participants were also offered the opportunity to provide feedback.

Software and Subject Management Details

In order to ensure the experiment is internally valid, there are a number of important details in the design of the software and subject management. The features described below were added after reviewing data from trial runs of the experiment and additional research on the experience from the perspective of the workers done through direct (and compensated) correspondence with some, and by reviewing forums that host robust communities related to working on Mturk (e.g., MturkNation, TurkerHub, Reddit).

In general, Mturk workers are interested in participating in academic research, but not at rates lower than they typically earn. In fact, because a new academic researcher lacks a good reputation[16] and presents a risk to workers, a premium on the typical rate may be necessary to ensure data quality, especially if the task takes more than a couple of minutes. The modal worker does not perform tasks for entertainment, but rather to supplement income (Paolacci et al., 2010; Ipeirotis, 2010). This makes them sensitive to the effective hourly wage and, although there is variability in their reservation wages, most workers seem to have the federal minimum wage in mind. Because most choose to accept a task based on a rough calculation of the effective hourly wage, any misrepresentations of estimated earnings or time to completion are likely to provoke ire.

[16] Turkopticon is an example of a system to create and disseminate reputation scores for all job posters

Importantly, because workers think in terms of an effective hourly wage, the compensation structure of the experiment should correspond to the workers' overall incentive structure in the Mturk marketplace. In particular, a significant portion of the overall payment must be guaranteed for the completion of the task. Workers always have the opportunity to quit the current task and start a different one and, therefore, are often aware of opportunity costs. If the task progresses slowly or the bonus earnings appear lower than expected or advertised, a small guaranteed payment might lead them to exit the task despite already having committed time to it. It is important then that the subject's current expected payment is roughly at the subject's reservation wage for that time. In practice, this means that the guaranteed payment should be large relative to potential bonus earnings and idle time should be additionally compensated. The former incentive structure generally corresponds to rational behavior with respect to opportunity costs, but furthermore does not appear to undermine the ability of within-game incentives to induce preferences in the sense of Smith (Smith, 1976); once the guaranteed payment meets the reservation wage, in-game bonuses become an exciting opportunity to exceed the worker's earnings goal, and workers try in earnest to maximize bonus earnings.

Compensation of idle time should, in principle, not be necessary if that time is included in the estimate of the time necessary to complete the whole task. However, workers generally view idle time as distinct from time spent on the task, likely because the experience of waiting makes them more aware of opportunity costs. Experimentation with the compensation design resulted in an idle-time compensation structure that delivered an effective hourly wage close to the prevailing reservation wage, broken into small increments. Paying one cent for every five seconds of waiting ensured a wage of $7.20/hour, and that any amount of "unpaid" time was very small. Furthermore, live tracking of the elapsed time and accrued bonus on the wait page created a gratifying experience for workers of being able to watch earnings grow.

A complication of this overall compensation structure is that it is more challenging to implement a "show-up" payment often used with traditional subject pools. Ensuring that the required 24 subjects are available requires over-recruiting, carefully monitoring arrivals, and immediately removing the task from the Mturk listings once enough accept. This often still results in subjects in excess of 24. The Mturk platform does not allow workers to be turned away once the task has been accepted, but one can ask workers to“return” the task. Unfortunately, this precludes being able to compensate them through the normal means, requires them to trust the task poster to follow up, and generally risks upsetting them, possibly causing reputational damage. Furthermore, given that workers have been personally invited and have been watching for an email to start the task, sending them off with a smaller show-up fee could harm their responsiveness to future invitations. Although it increased costs, the long-term solution used here was for subjects in excess of 24 to play a version of the game against bots designed to make moderately intelligent decisions. None of those subjects ever gave any indication that they knew that they were not doing the real task (although, interestingly, some playing live participants commented that they thought they were playing with bots). Subjects who played against bots were paid what they earned and later dropped from the analysis.

Once subjects joined a session (after having read a description of the task and the IRB consent statement), they were shown the game instructions. This page had a comprehension task built in; rather than clicking a standard button to advance to the next page, subjects were directed at the end of the instructions to click a hidden button in the form of the text “match names with playing partners” where it first appeared in the instructions. Knowing to click on these words and scanning for them reinforced the basic task of the game and ensured active engagement. Approximately 10% of the prospective subjects never made it past this stage, although it was not possible to tell if they left the game for other reasons. Any subject who did not pass this comprehension test was unable to join the game and was forced to return the task.

The subjects who passed the instructions and comprehension task were then asked to agree not to attempt to use means outside the game interface to coordinate. This is a real concern for a population that has a robust set of forums and chat rooms dedicated to its community. I describe the primary software and management mechanisms used to defeat such attempts below. By asking subjects to agree to not use external means of coordination—a very weak form of control—I hoped to make subjects think twice about engaging in such behavior, but more importantly, create defensible grounds for removal from the subject pool if there was evidence of attempts to coordinate.

Once subjects agreed to not use external means of coordination (all did), they were taken to a page to wait until enough subjects arrived. This page showed them the number of minutes and seconds they have been waiting and their compensation for that wait. As soon as 24 subjects made it to the wait page, the game began. The first page of each round displayed a head shot of an individual and a field to enter a name. An example of this page appears in Figure 3. The picture was always a younger woman because pretesting showed that other demographics had some natural focal points; photographs of middle-aged and older men without distinctive characteristics often quickly elicited a handful of competing alternatives (e.g. Bob, Mike, Bill). The same was true for older women (e.g., Janet, Susan, Patty). Pictures of younger women (approximately 25-45 years old) generated a range of first submissions. Participants had 15 seconds to submit a name and were taken to a wait page until all participants submitted names. On the results page, participants were shown the name they submitted, the name their current partner played, and, when applicable, the spectating-linked names. They were also explicitly told whether the names matched, their payoff for the round, their cumulative earnings, and their history of matches and failures (but not the associated names). The results page timed out after 10 seconds, although participants could advance past it to another wait page. An example of this page appears in Figure 4.

I took several measures to ensure that there was no external collusion between the subjects. Together, these measures address multiple potential means of collusion and work together to all but ensure that it did not happen. As subjects first arrive, the related IP addresses are screened to make sure there is only one subject per IP-address. The first subject to arrive from any address is allowed to remain, but all others are blocked from continuing, informed why, and asked to return the task. It is not uncommon for Mturk workers to work in the same household or workspace and, while they might honestly avoid collusion, I erred on the side of caution and permitted only one. The more problematic case is a worker with multiple Mturk accounts. This is a violation of the Mturk Terms of Service agreement, but by their own admission on forums, some workers use multiple accounts. Those willing to ignore such rules might also try to use virtual private networks to use different IP addresses for each account, but they would have had to do that before arriving for my task. Furthermore, in virtue of how invitation groups are chosen, the probability that two or more of their worker IDs are present in the same group is very low. This is because the experimenter controls which worker accounts receive an invitation to a game and have the necessary "qualification" for it. (A qualification is a virtual token within the Mturk platform that can be used to control which workers can do which tasks.) Without an invitation and the necessary qualification, workers cannot join a session. I used qualifications to ensure that no more than 4 workers ever played the same session together. The final precaution I took to protect against collusion was to make sure there was no relationship between arrival time and the location in the network; while small-scale efforts at collusion could be successful if those colluding are network neighbors, the nature of the game renders such efforts ineffective or counterproductive if would-be colluders are not network neighbors.

Within the game itself, there are several features to prevent collusion. The first is just a basic design of the game; participants have only 15 seconds to submit a name. Failure to submit costs a participant rewards and can lead to expulsion from the game, a fact participants are reminded of every time they fail to submit. Given this time constraint, participants typically submit immediately ($\tilde{5}$ seconds). Even those who do not submit immediately have very little time to attempt to communicate with other participants about emerging patterns in the names. Timed submissions do nothing to protect against premeditated attempts at collusion, however. Given that workers know the date and time of the game in advance, they might seek each other out on forums in advance and agree to use the same name. I have found no evidence of this on the forums to which I was able to gain access (some have very high bars for admission), but I nonetheless added more comprehensive features to defeat such efforts. The primary means is the screening and blocking of names that exhibit surges in frequency. A name that appears for the first time with more than two instances is blacklisted for that round. It is not far-fetched for the name "Sarah" to be the first submission for multiple participants, but, erring on the side of caution, my software barred any name that first appeared with three or more instances. When a name was barred, those submitting the name were told the name is blacklisted, but not why. Partners of those submitting the barred name were informed that their partner had not submitted a name, not that it had been barred. Multiple names could be barred in any given round, but the barred status is confined to that round only. If one or two instances of a previously barred name are submitted in a subsequent round, it is accepted without comment. This is because participants frequently exhibit the suboptimal behavior of introducing new names well into the game (this behavior was also observed in the original *CB* experiment) and there is no reason for barring these submissions. This barring mechanism was triggered a total of 20 times in the 825 total rounds of play. All instances occurred in the first round of play. 12 of these instances were three participants submitting the same name. Five instances included multiple names being barred. The largest incident was the submission of the same name by 9 participants. The barring mechanism effectively removed the name from circulation, as participants immediately abandoned it, meaning that a barred name never became a global convention. Interestingly, only 5 of the 20 trials for which the barring mechanism was trigger ultimately resulted in a global convention, suggesting something about the barring mechanism or the behavior it targets might have harmed the prospects for future success. It is clear that, whether premeditated or coincidence, these instances did not contribute to the emergence of conventions, although they might have inhibited it in some instances.

The screening of names for the purpose of barring is applied only to names that have not yet appeared in the session. Once a name has been successfully introduced, it might still be implicated in collusion efforts that occur outside the game interface. If the name "Sarah" is submitted by a single participant in the first round, but then by 10 participants in the second round, collusion could be the reason. However, that is not guaranteed to be the case. After a participant submits the name, their partner sees it and is very likely to play it. In treatments where spectating names are also shown, any number of other players may have seen the name on their results page and accordingly played it. Given this structure, preventing collusion requires judging whether a name is spreading around the network too quickly. To do this, I track the exposure of each participant to names, that is, simply any name they have seen or successfully played (i.e., was not barred). Once a participant has been exposed to a name, I assume it is completely reasonable for them to play it, and therefore the spread of that name to this participant did not happen outside the confines of the game interface. One could have a more demanding definition of exposure that considers the likelihood that participants will forget names seen many rounds ago, but implementing a suitable rule for the game would require making specific assumptions about how participants process the information that might not be justified.

To reconcile the idea of proper exposure with the fact that participants contribute names to which they have not been exposed, only one player per round is allowed to play a name to which they were not properly exposed. If two or more participants submit a name they were not properly exposed to, the name is barred for all such participants, and their submission is recorded as a null. Others playing the same name, but who have been properly exposed, do not have their submission barred. Although not a foolproof system to guard against impermissible coordination, it would take an impressive amount of organizing by the participants to defeat the system, and it still grants participants significant latitude in name choice. This type of screening occurred only six times in the 792 applicable rounds of play, and none of the affected names went on to become a global convention.

The two mechanisms of barring name described above offer significant protection against any organized effort on the part of participants to game the system. Another possible way to game the system would be to submit natural focal point words, such as “woman” or “name”. Such efforts would have the virtue of not needing to be centralized. This potential problem is addressed simply by creating a list of words the input field will not accept, thereby preventing the participant from submitting and advancing to the next page. Whereas barring surges in names requires the names to have been submitted and tallied, focal point words can be screened out before submission, and participants can still submit a valid name. Roughly twenty focal point words, mostly nouns, including a number of obscenities, were screened out. Furthermore, any names that appeared in the training module were added for the obvious reason that participants could be primed to focus on them.

A few additional design features are used to ensure the integrity of the experiments. The pictures of the individuals to be named by the participants were changed every two games. No worker in the subject pool can play two games in a row, so changing the picture every two games ensured that no worker ever saw the same picture. The usual means of downloading images was also disabled for the game, just in case workers considered attempting to use the image as a means of coordination on forums.

As the above description of the software suggests, using Mturk as a subject pool presents some unique challenges and still has some of the issues that any subject pool can have. These issues were addressed through subject pool management and software design. The attempts to game the system were successfully defeated and, if anything, led to the underestimation of the effect of spectating-linked information sources. In principle a very sophisticated effort built around knowledge of these designs could go undetected, but even in the highly improbable event such an effort succeeded regularly, one would still need to account for the consistent pattern in the results; the information dynamics play a crucial role in determining group-level outcomes. Finally, the fact that the members of the subject pool could participate in multiple games could in principle allow them to learn from their experience and possibly develop a strategy to encourage the emergence of a group convention. They seem to intuit the value of a single convention, but it is not clear how their individual behavior could bring that about, especially because there was no systematic relationship between treatment variables and when the game was played. It is quite possible that what works in one network topology and spectating-name treatment would not work in another (e.g., high commitment to a single name could be good in a fully connected network, but not a lattice network).

[Figure 3 about here.]

[Figure 4 about here.]

Network Design Details

All networks were generated using the Python package NetworkX (Hagberg et al., 2008). The small world networks had a rewiring probability of 0.2 and were guaranteed to be connected. These graphs were used to create pairs of neighbors for the 25 rounds of each trial. Participants need a partner for each round, so the set of pairings for each round was chosen randomly from among complete pairings. Because the sequence of pairings could have its own effect, the sequences were reused for the different treatment levels of spectator-linked information. For example, the 12 trials with small world networks used only four network-pairing sequence combinations, reusing each of the four once in each information treatment. There was no obvious effect of pairing sequences, so no further analysis was pursued.

Divergence Difference Measure

The divergence difference reported in 2 is constructed with a memory parameter. The distribution of names under consideration in a given round likely includes names seen in previous rounds, but not all previous rounds. Consequently, a number of rounds to include in the construction of the distribution of names in circulation is required. All results reported here include at most the eight previous rounds. The results are not qualitatively different for 4, 12 or 16 rounds. Thus, the divergence difference for the t^{th} round of a game is defined as follows:

$$DD_t(\Phi_t, \Theta_t, \Lambda_t) = D_{KL}(\Theta_t||\Phi_t) - D_{KL}(\Lambda_t||\Phi_t) \quad (1)$$

where $D_{KL}(P||Q)$ is the Kullback-Leibler divergence from Q to P, Φ_t is the true distribution of the names played in the t^{th} round and the seven rounds before it. Θ_t and Λ_t are the comparison distributions. Θ_t is participant i's distribution with names from the synthetic alternative source, and Λ_t is the observed distribution with names from the spectating-linked source. Because the KL divergence increases the less similar two distributions are, DD_t is positive when the spectating-linked source yields more information. I use the Kullback-Leibler divergence instead of the symmetric Jensen-Shannon measure because there is a clear reference distribution that encodes the state of the system. Because the true distribution and the distribution of names seen by each participant do not have a common support, the latter requires adjusting the participants' "seen" distributions to have non-zero weights (here .001) on each missing name and then renormalization of the weights. The results reported in 2 are averages across all participants for all games in the condition.

Simulation of Alternative Name Sources

To clarify the information content of the spectating-linked names, I compare that source to other source types by simulating exposure to names from the alternative sources. I start with the list of names that a given participant had seen through direct interactions with partners. Names from the simulated source are added to that list. The four simulated sources are constructed as follows. *Random alters* selects names each round at random without replacement from among those played by the participant's alters with whom she is not currently paired. *Random non-alters* selects one or two non-neighbors at the start of the game and shares the name they played each round. The *longest bridge* source identifies the other participants in the network to which the focal participant has the longest shortest path and reports the names that other participant played each round. When there are multiple paths of the longest length, the target is chosen randomly. Because the addition of a bridge changes the shortest paths, when two longest bridges are needed, the path lengths are recalculated after the addition of the first bridge. Finally, the *maximum information* source quantifies the information gained by the exposure of a participant to i) each name in the single-name case, or ii) each possible pair of names in the two-name case. The name(s) that produce(s) the most information become the exposure for that participant that round. This calculation factors in what the participant has seen in previous rounds, including the hypothetical exposures.

List of Figures

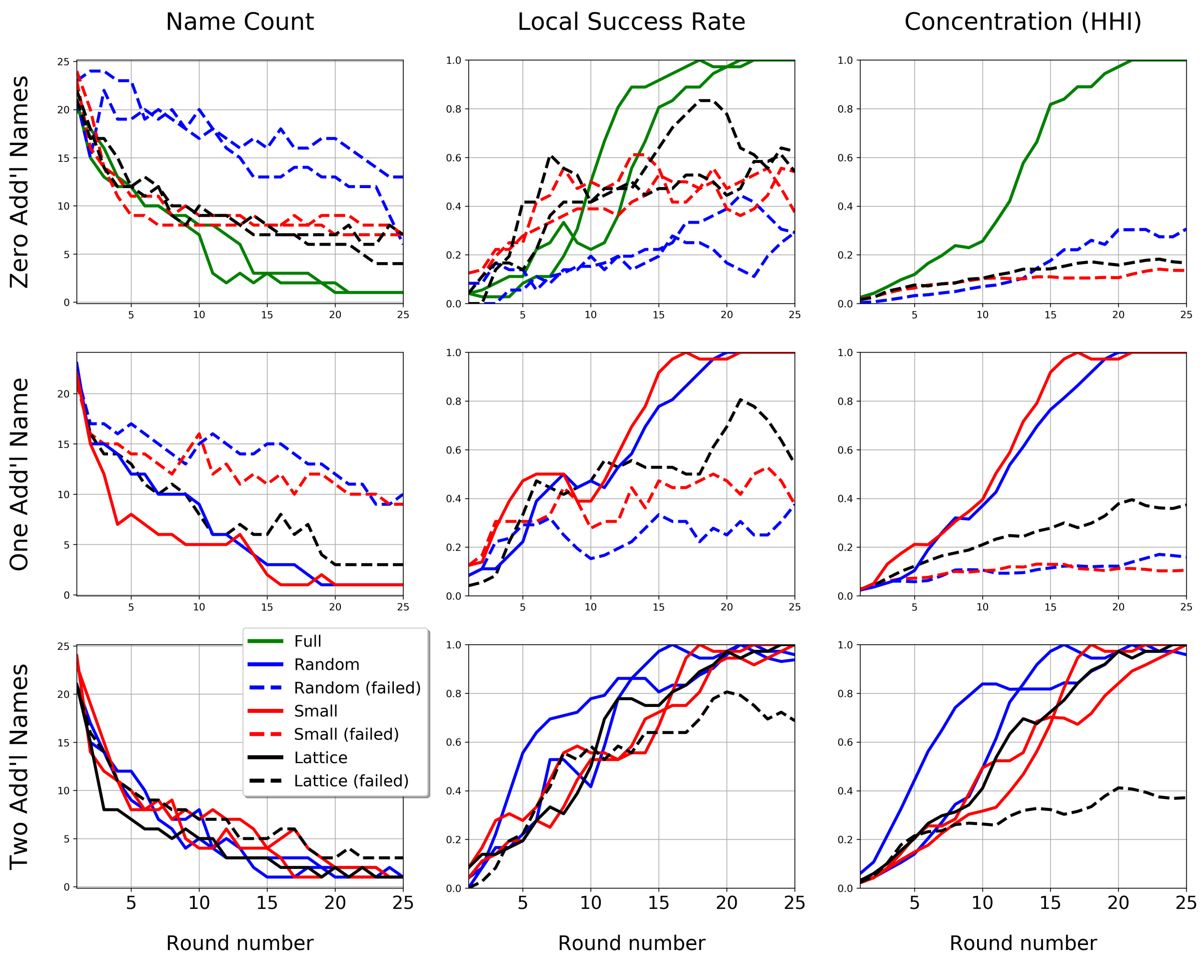

Figure 1: Measures of name counts, local coordination success rate, and concentration (HHI) of names for a sample of runs for visual clarity. Rapid increases in concentration early in the run is associated with the emergence of a single convention.

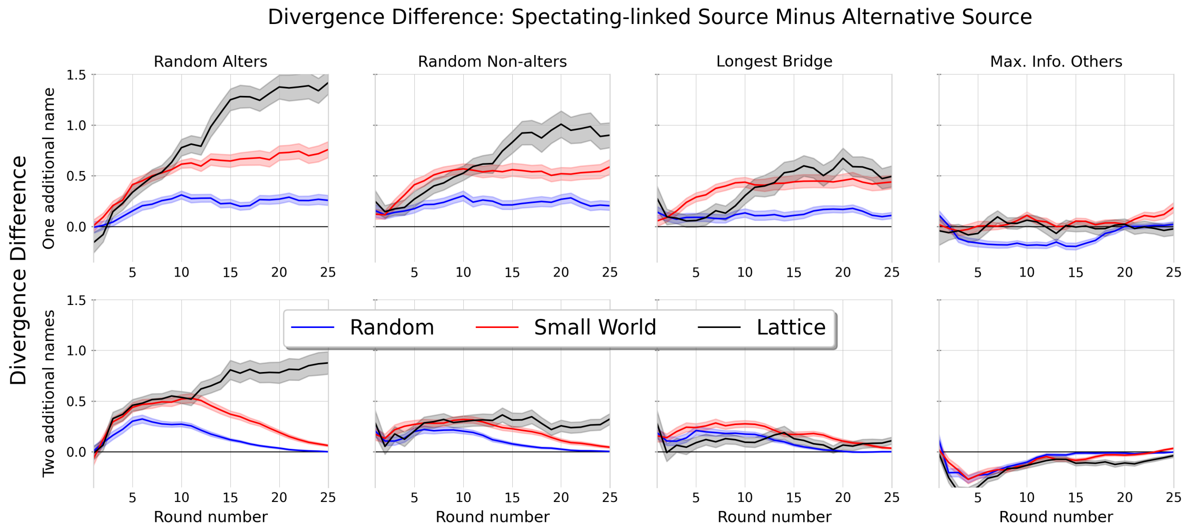

Figure 2: The average of differences in source divergences from the true distribution of names, with standard error bands. The difference is positive when the spectating-linked source reveals more information that the alternative source.

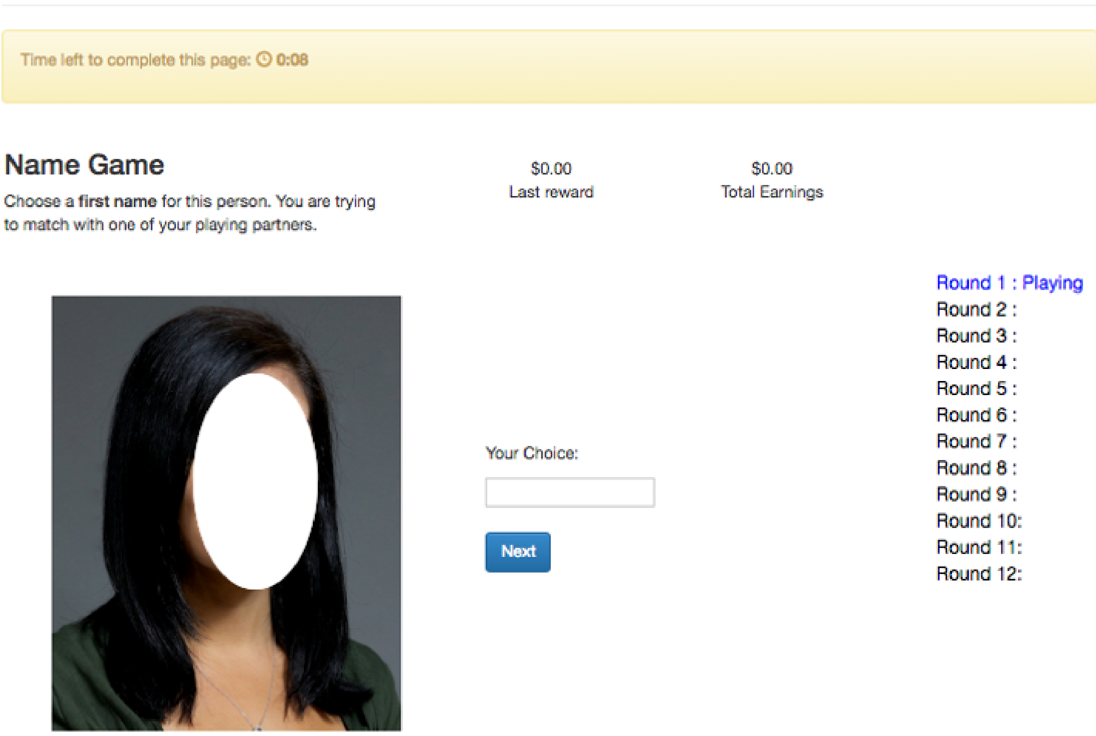

Figure 3: Name Game submission page: The page displays the time remaining to submit, the total earnings, the last reward (or penalty), the current round and the history of the last 12 rounds of play. The face is obscured only for publication.

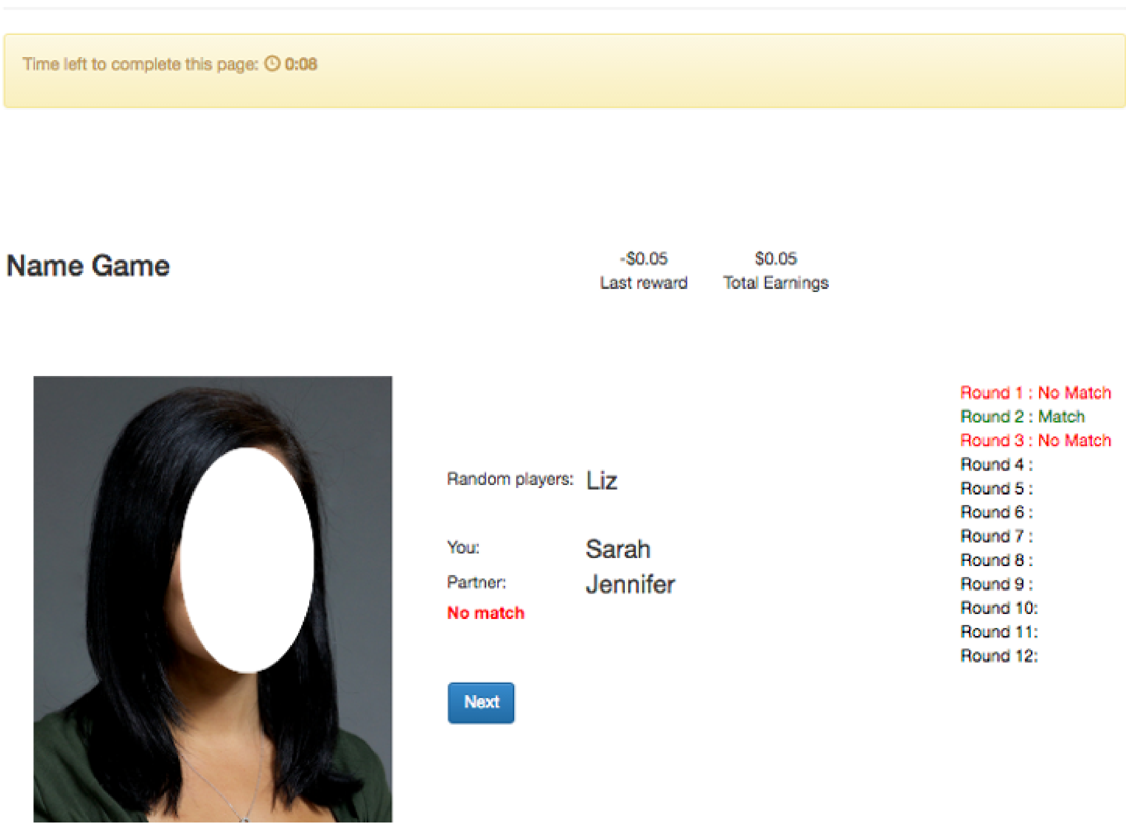

Figure 4: Name Game results page. The display includes the submission of the participant, partner and a random player and the history of rounds. The face is obscured only for publication.

List of Tables

	Network Topology			
Name Count	Fully Connected	Random Regular	Small World	Ring Lattice
0 names	11.0	1.0	0.56	0.34
1 name	18.1	13.6	13.5	13.4
2 names	21.0	18.9	18.9	18.9

Table 1: Coefficient of algebraic connectivity by condition. The maximum value of the coefficient for a network of 24 nodes (i.e., a complete network) is 24. Values for the stochastic networks (i.e., all but the ring lattice) are the average of 500 individual networks within the ensemble.

Add'l Name Count	**Network Topology** Fully Connected	Random Regular	Small World	Ring Lattice
0 names	Convention, Convention, Convention, Convention	No Convention, No Convention, No Convention, No Convention	No Convention, No Convention, No Convention, No Convention	No Convention, No Convention, No Convention, No Convention
1 name	Convention, Convention, Convention, Convention	Convention, Convention, No Convention, No Convention	Convention, No Convention, No Convention, No Convention	No Convention, No Convention, No Convention, No Convention
2 names	Convention, Convention	Convention, Convention, Convention, Convention	Convention, Convention, Convention, Convention	Convention, No Convention, No Convention, No Convention

■ (green) = Convention ■ (red) = No Convention

Table 2: Each trial represented as a box within its condition and color-coded by whether the group established a single convention within the 25 rounds of the game. Two additional trials, here coded as *no convention*, would likely have reached a global convention within another five rounds of game play.

Spectating Name Count	Name source	Network Topology: Fully Connected	Random Regular	Small World	Ring Lattice
1 name	Self Introduced	2.15 (0%)	2.73 (0%)	1.70 (0%)	1.77 (0%)
	Alters Only	1.11 (27%)	1.57 (33%)	1.18 (42%)	1.00 (43%)
	Spectating Only	0.63 (39%)	0.68 (37%)	0.39 (27%)	0.50 (20%)
	Alters & Spectating	1.05 (68%)	1.26 (51%)	0.84 (56%)	0.92 (64%)
2 names	Self Introduced	1.94 (0%)	2.08 (0%)	1.67 (0%)	1.62 (0%)
	Alters Only	0.88 (23%)	0.82 (24%)	0.98 (38%)	0.75 (41%)
	Spectating Only	0.73 (37%)	0.66 (40%)	0.61 (45%)	0.43 (49%)
	Alters & Spectating	1.68 (53%)	0.97 (66%)	1.17 (67%)	1.00 (66%)

Table 3: Average number of names played, sorted by the source it was previously seen by, and the average individual success rate associated with playing that name.

Longest Bridges	Network Topology: Random Network	Network Topology: Small World
1 bridge	Conv., Conv., No Conv., No Conv.	Conv., No Conv., No Conv., No Conv.
2 bridges	Conv., Conv., Conv., Conv.	No Conv., No Conv., No Conv., No Conv.

(green) = Conv. (red) = No Conv.

Table 4: The Name Game with the additional longest bridge name exposures. A detailed analysis of the single small-world trial resulting in a convention appears to be the result of good fortune; two structurally distant participants started with the same name and were paired with others in a way that lead to a single name dominating early.