Computational Modeling of Human Social Intelligence and Communication

Jeungmin Lee (jeungminlee@kaist.ac.kr)
Department of Bio and Brain Engineering
Program of Brain and Cognitive Engineering
College of Engineering, KAIST, Daejeon, 34141, South Korea

Jerald D. Kralik* (jerald.kralik@gmail.com)
Department of Bio and Brain Engineering
College of Engineering, KAIST, Daejeon, 34141, South Korea

Jaeseung Jeong* (jsjeong@kaist.ac.kr)
Department of Bio and Brain Engineering
Program of Brain and Cognitive Engineering
College of Engineering, KAIST, Daejeon, 34141, South Korea

Abstract

The next level in understanding human social cognition is to model it comprehensively. To this end, we have been developing a framework and model that takes as input an event involving someone (focusing on who it was and what they did), and assesses the event based on whether it should change social accounting among individuals, and whether something should be done, such as communicating with others. Here, we present development of the model computationally and results generated by it as predictions to be tested empirically: e.g., more communication about those socially close to us when their actions are positive, and more about those with higher status (i.e., celebrities) when negative; and the relative merit or egregiousness of a wide range of behavior. Leveraging what is known of the human social mind and brain, our work aims to provide a comprehensive model of human social cognition.

Keywords: social cognition; theory of mind; communication; decision-making; computational model

Introduction

A true understanding of human social cognition should produce a comprehensive cognitive model that successfully explains human social behavior. Among many complexities, this includes successful real-time social interaction involving a great deal of inference such as for mind- and context-reading. Moreover, this inference ranges from more explicit, formal reasoning to what is considered commonsense or social intuition. Additionally, models of social intelligence will need to have not only a much richer understanding of peoples' minds and immediate context, but of sociality more fundamentally. That is, theoretical considerations (including evolutionary ones) and evidence across the social sciences have shown that social interaction can be construed in terms of social economics, with each individual — each agent in the multiagent world — having a certain amount of social value, and each social interaction a transaction, in which individuals spend and accrue social value among themselves (i.e., social accounting) (e.g., Cosmides, & Tooby, 1992; Dunbar, 2004; Foster, 2004; Pinker, 2008; Rosnow, 2001; Lee, Kralik, & Jeong, 2018; Lee, Kralik, & Jeong, 2019).

Thus, successful models will need to track and help maintain this social accounting if they are ever to be fully functional as human-level social agents and society members. Moreover, much of this social interaction resolves to deeper fundamental issues relating to morality, such as treating each other fairly. Finally, because of the primacy of these factors and the complexity of human societies more generally, successful social interactions even between two individuals often require a triad (or larger network) of people to communicate the information: e.g., due to the sensitivity of confronting someone directly or the inability to maintain accurate knowledge about others (e.g., what they have done, have learned) when not present, requiring updating from others (Baumeister, Zhang, & Vohs, 2004; Dunbar, 2004; Foster, 2004). Such indirect communication can be considered "gossip", which has seemed trivial and frivolous, but belies a deeper significance (Dunbar, 2004).

In sum, there is a vast amount of research findings on human social intelligence and sociality more generally; and yet a comprehensive theoretical framework and computational model of human social intelligence and communication has been lacking. Having this will not only organize and integrate what is currently known about human social cognition, but will also help clarify what is yet to be better understood. We have thus been developing such a framework (Lee, Kralik, & Jeong, 2018; Lee, Kralik, & Jeong, 2019). Our work has focused on integrating findings across the social sciences into a general framework and model, and here we present a brief description of the framework and initial development of the model computationally.

Lower-level, network-based architectures, including deep learning, provide a flexibility and generalization power not yet touchable for higher-level symbolic-based architectures; yet the latter reach a level of richness of human social intelligence and communication that, although currently too circumscribed (i.e., generally hand-crafted and brittle), the lower-level models have not yet broached. How the two shall meet we do not know. Thus, at least on the path to a complete understanding of human social cognition in mind and brain, both approaches are necessary as researchers determine what architecture (or combination) can span the entire capacity of human social abilities (and perhaps beyond). Leveraging what is known of the human social mind and brain, we take a top-down theoretical approach beginning at the symbolic level. In what follows, we first briefly describe our overall framework, paradigm and model; we then describe how we have begun to implement the model computationally; and we then describe the results thus far generated, to be considered as a set of predictions for an experimental study on human social information communication that we have recently undertaken in our laboratory. The empirical study intends to test and potentially validate the modeling work with actual human findings; at the same time, the model provides deeper theoretical insight into human sociality, enabling for example, *a priori* predictions of our social behavior.

Framework, Test Paradigm, and Model

As seen in Figure 1, our general framework focuses on a *central problem-solving agent* who receives information about some activity of a *target* person, such as someone going to the movies, helping someone else in need, working well (or not) with others in a group task, beginning a romantic relationship, or cheating on an exam (Lee, Kralik, & Jeong, 2018; Lee, Kralik, & Jeong, 2019). The information is received from an *external source* or observed by the central agent directly, and based on this information she may or may not communicate with a *receiver* about it. The receiver is conceptualized broadly as anyone else, whether an additional person, and thus as "gossip", but also potentially communicating with the *source*, such as in further conversation, or with the *target* person him/herself. We call the event involving the target a "scenario" (e.g., "Kim was caught cheating on the final exam.") and it is to this point always based on someone doing something — and thus social information broadly construed.

In our test paradigm, we focus on three main factors of the event: who did it, i.e., the *target*, what they did, i.e., its *content*, and whether the content was positive or negative, i.e., its *valence*. For target identity, we have chosen to first test *ingroup* versus *outgroup* versus *celebrities*, since they enable examination of two critical social factors: *closeness* (of the target to the central agent), with the differences being ingroup > celebrities > outgroup; as well as *status*, with the differences being celebrities >> ingroup > outgroup (Aronson et al., 2016; Foster, 2004). With three *target* levels, two for *valence*, and eight *content* domains selected to span the space of activities the target may engage in (described below), we produced 48 different scenarios. With this comprehensive set, we sought to generate a set of predictions of how the various combinations are processed by the social mind/brain and drive social behavior and communication.

Model of Social Intelligence

Our model, then, is of the central agent's mind/brain, and how she determines what to do with the scenario information (Aronson, Wilson, Akert, & Sommers, 2016; Gazzaniga, Ivry & Mangun, 2013; Glimcher & Fehr, 2014; Rai, 2012; Kralik, 2017; 2018; Lee, Kralik, & Jeong, 2018; Lee, Kralik, & Jeong, 2019). To make this determination, she must process the scenario across a series of modules (that make up the central agent's social mind/brain).

We focus here on receiving information from an external source (versus observing the event directly). From a neuroscience perspective, the central agent must first sense and perceive the scenario information. Our intention is to ultimately build a system with natural input, such as via language or reading; here, however, we concentrate on more central cognitive components. When the central agent reads or hears a scenario such as "Kim was caught cheating on the final exam.", her mind/brain must first understand the basic concepts, which in our model occurs initially within the *Perception* module via accessing memory for general concept knowledge, generally realizing each word making up the sentence (like Alex, saved, child, etc.), but not the deeper meaning that the scenario is carrying. Further scenario processing is conducted in *Initial Cognitive Process* in which the stimulus takes on a deeper sense of identifying the target and the social domain at issue. This importantly includes the latter's corresponding *affect response* — as an *affect score* — such as 'Kim cheating on the final exam' relating to *fairness* or 'Alex saving a child from a fire' relating to *care* that is particularly self-sacrificing and heroic. To determine the affect score of the particular event or scenario, the *Initial Cognitive Process* accesses an *Affect Knowledge Base*, which represents our main emotion core of the model (Damasio, 1996; Gazzaniga, Ivry & Mangun, 2013; Glimcher & Fehr, 2014; Schachter & Singer, 1962; Kralik, 2017).

More specifically, we organize the *content* of possible events into eight content domains (Table 1). Five were adopted from the well-established moral foundations — prosociality, fairness, community, respect, and purity (see Haidt, 2007) — and the remaining three — competition, social-oriented, and general social affairs — were selected to represent other important social activities (Aronson et al., 2016; Dunbar, Marriott & Duncan, 1997). The affect score of each content domain was assigned via theoretical consideration and empirical evidence for the affective/emotional intensity that the domains carry. For example, events related to *prosociality* — composed of *care* for positive valence, and *harm* for negative valence (e.g., "Alex saved a child from a fire [care]" or "Sam stabbed a person with a knife [harm]") — are expected to be more intense and therefore more emotionally provoking than other content domains such as *fairness* — broken down into *fair* and *cheating* (e.g., "Kim cheated on the final exam [cheating]" — or *social-oriented*, composed of *altruism* and *selfishness* (e.g., "Taylor donated part of his salary to a charity [altruism]") (see Foster, 2004; Haidt, 2007). We discuss more about each domain in Results.

With the *affect score* as output of the *Initial Cognitive Process* module, our model then uses it as a gate for further processing, with the score representing a problem to be

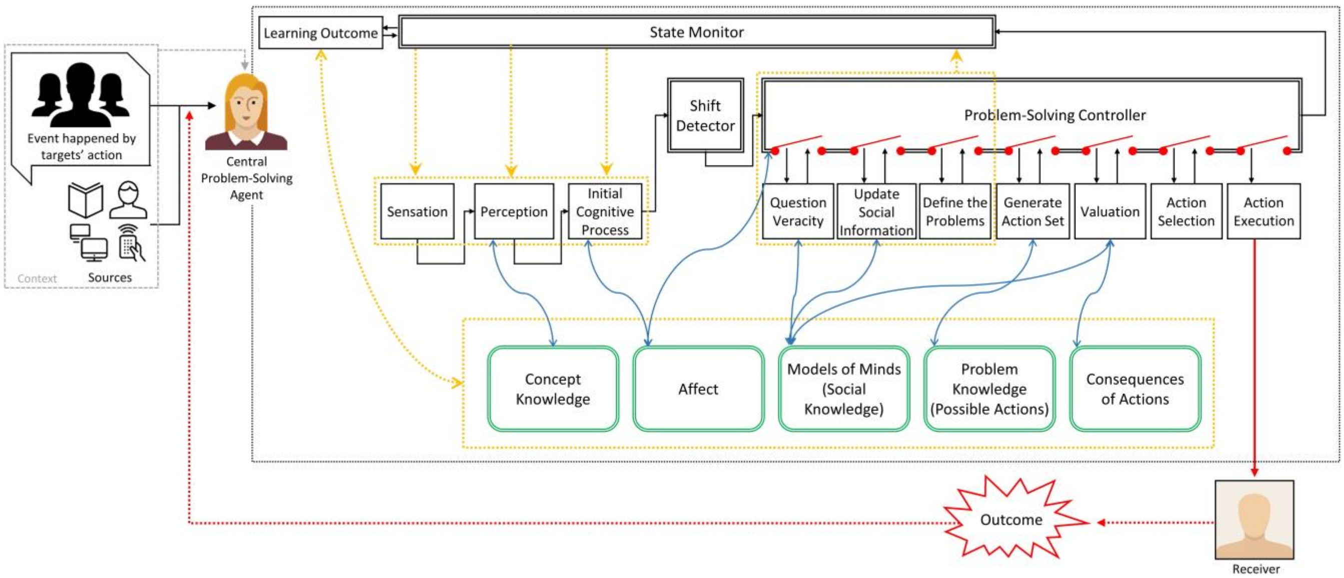

Figure 1: The complete model of the internal processes underlying social information and communication. The central agent goes through a set of internal processes (black rectangles) by accessing her relevant knowledge (green rounded rectangles).

solved or social accounting to be resolved, with the problem extent (its intensity) captured by the score value (Figure 1).

Next, a general *Problem-Solving Controller* module orchestrates problem-solving by activating a sequence of subprocesses. It is the key metacognitive process orchestrating the entire system (Kralik et al., 2018). The first subprocess assesses the likely truth of the information. Here, we assume that the source is trustworthy, and the controller then moves to *Update Social Information* where the main *Social Knowledge Base* is accessed. This knowledge base contains *models of the minds* of the people in the central agent's (multiagent) world. That is, the central agent has her own model of other people's minds (consisting of their beliefs, interests, personal traits, etc.) stored as social knowledge, accessed and used or modified when necessary (Gazzaniga, Ivry & Mangun, 2013; Glimcher & Fehr, 2014; See Lee, Kralik, & Jeong, 2018 and Lee, Kralik, & Jeong, 2019 for details).

Table 1: Eight content domains of possible social events (i.e., someone does something) along with their decomposition into positively and negatively valenced aspects, and their corresponding affect score in descending order (1-7 scale). The five domains adapted from Haidt's moral foundations are marked with superscript "m" (Haidt, 2007).

Content	*Affect Score*
Prosociality (care/harm)^{m}	7
Fairness (fair/cheating)^{m}	6
Competition (positive/negative)	5
Social-oriented (altruism/selfishness)	5
Community (loyalty/betrayal)^{m}	4
Respect (authority/subversion)^{m}	4
Purity (sanctity/degradation)^{m}	3
General social affairs (positive/negative)	1

Although in reality multiple problems are potentially in play or further introduced by the central agent's possible subsequent actions, we focus to this point on the single main problem, such as harm produced by the target's action (e.g., Sam stabbing a person with a knife) that must be resolved or in some way dealt with. Once the problem is defined, the *Problem-Solving Controller* activates the *Generate Action Set* subprocess to determine which actions to consider for the given problem (Figure 1).

Here, in our first computational development of the model, we concentrate on one main action, whether to communicate with an additional person (i.e., not the target or information source, and thus as 'gossip'), in the face of a wide range of possible social scenarios. The controller then moves to the key subprocess of *Valuation*, the central focus of our current computational model development, described comprehensively below. Upon the completion of valuation, the controller then moves to *Action Selection* and then *Action Execution*, which if the action is actual gossip, the central agent communicates with a receiver. An outcome then would occur, such as the receiver directly confronting the target, the receiver telling another receiver — that is, further gossip — or the receiver doing nothing with the information.

Valuation

We now describe the *Valuation* subprocess in detail, the central focus of our current development. *Valuation* evaluates each possible action based on the combination of potential *benefits* and *costs* of taking the action, combined with the scenario *affect score* that provides the original impetus for the problem and possible recourse to take action to resolve it.

Thus, we subtract the potential costs from benefits and multiply it by the significance of the event. More specifically, using the following equation for valuation:

$$Value_{Gossip} = A \cdot (B_{Total} - C_{Total}) \qquad (1)$$

where A is the affect score of the given scenario (target person doing something), and B_{Total} and C_{Total} represent total benefits and costs of taking the given action (in this case communicating with a third-party receiver). B_{Total} and C_{Total} are each composed of multiple potential benefits and costs derived from taking a given action, i.e.,

$$B_{Total} = B_1 + B_2 + \cdots \quad (2)$$

$$C_{Total} = C_1 + C_2 + \cdots \quad (3)$$

Additionally, each individual benefit B_i or cost C_i is composed of weighting factors thus:

$$B_i = t_{Bi} \cdot v_{Bi} \cdot w_{Bi} \quad (4)$$

$$C_i = t_{Ci} \cdot v_{Ci} \cdot w_{Ci} \quad (5)$$

where t represents the relative weighting for *target* (e.g., ingroup, outgroup, or celebrity), v for scenario *valence* (i.e., positive or negative), and w for the relative influence of the individual benefits and costs.

As seen in Table 2, from the social communication literature as well as our own theoretical considerations based on an evolutionary and socio-economic perspective of how the communication about the target could ultimately benefit the central agent, we have identified five main benefits and three main costs for communicating with a third-party receiver (gossiping) (see Aronson et al., 2016; Baumeister, Zhang & Vohs, 2004; Dunbar, 2004; Foster, 2004 for reviews; Rosnow, 2001; Russell & Norvig, 2020; Lee, Kralik, & Jeong, 2018); and in the current development we have added corresponding relative weighting factors, w, t, and v, based again on this literature and our evolutionary and socio-economic theoretical considerations (Aronson et al., 2016; Baumeister, Zhang & Vohs, 2004; Dunbar, 2004; Foster, 2004; Rosnow, 2001; Russell & Norvig, 2020).

Table 2: Benefits and costs of communicating with an additional individual as receiver (i.e., gossiping). Each cell contains the weighting factors for each valuation category, based on the relative impact of each category (w), the target person's identity (t) and the valence (v) of the scenario event. See text for further description. Weights on a 1-4 scale to obtain meaningful relative values distinguishing the critical factors.

Valuation Categories	*w*	*Target (t)*			*Valence (v)*	
		Ingroup	*Outgroup*	*Celebrity*	Positive	Negative
B_1: Avoid direct contact with the target	4	4	1	1	3	4
B_2: Feedback to the gossiper from receiver	3	4	1	2.5	4	4
B_3: Update receiver's knowledge	3.5	4	1.5	2.5	4	4
B_4: Influence target's social status	4	1.5	1	4	1.5	4
B_5: Receiver influences target's behavior	3	4	1	1.5	4	4
C_1: Potential direct contact from the target	4	4	1	1	2	4
C_2: Risk of spreading wrong information	3.5	4	1	2.5	4	4
C_3: Earn bad reputation as a gossiper	3.5	4	1.5	3	2.5	4

More specifically, one potentially powerful benefit is that the central agent does not have to face the target directly (i.e., indirect communication) (B1). The advantage of such indirectness is much greater when the target person is within one's ingroup, and much less so otherwise (as celebrity or stranger), reflected in the corresponding target weights (t). The advantage is also clearer when the information is something negative about the target (e.g., Kim cheating on the final exam), although it can also be relevant for positive information (e.g., it can be uncomfortable and awkward to speak highly of someone directly to them), also reflected in the valence weights (v).

In addition, the central agent may be able to obtain more information about the scenario event circumstances or confirm the information source veracity/truthfulness by checking with a receiver (B2) (i.e., another individual in the central agent's purview), since the receiver may have more information about the target than the central agent does; and the corresponding t and v weights reflect this. At the same time, the central agent can also importantly provide the receiver with new information to update the receiver's mental model of the target (B3) (with again the relative weights reflecting this benefit based on target closeness, which also reflects the relative detail of the mental models of target individuals in the receiver's mind).

Moreover, social communication also plays an important role in society by promoting fairness in terms of social order. That is, there is an inevitable hierarchy where some individuals have higher status in terms of power, wealth, fame, etc. than others. Although *status* can refer to both macroscale hierarchy (such as nationwide or worldwide celebrities and public figures) and microscale (within a smaller social group like school, workplace, or neighborhood), we focus here on the macroscale. Social communication can potentially influence this status based on disseminating relevant information about individuals (B4) (Aronson et al., 2016; Baumeister, Zhang & Vohs, 2004; Dunbar, 2004; Foster, 2004). Furthermore, because higher status requires justification, the general public is expected to be extra judgmental and critical with those of higher status, reflected in the w, t, and v weightings.

The last benefit is also important: the possible influence of the receiver on the target to reward or "punish" them appropriately (B5). This is particularly effective (a) if the receiver is in better position to influence the target and the central agent (e.g., closer, more respected), (b) as a means to reduce a possible defensive response by the target if confronted directly, or (c) as general social pressure (i.e., reputation).

Although there are advantages to disseminating social information, there are also significant disadvantages. First, there is the possibility that the target hears of the 'gossiping' and confronts the central agent directly (C1), reflected in the weightings accordingly. The central agent also runs the risk of being wrong about the information (C2). Spreading wrong information may not only influence the target's reputation, but also actively damages the model of the target's mind in the mind of the receiver(s). The importance of having accurate models of others is discussed in detail in our previous work (Lee, Kralik, & Jeong, 2018; Lee, Kralik, & Jeong, 2019). The last main disadvantage of information spreading is related to the traditional view of gossip as malicious behavior. By divulging information about a person (i.e., the target) absent during the conversation, the central agent may earn a bad reputation as a gossiper (C3); influenced both by closeness and valence, and thus reflected in the (t, v) weights.

We next examine our model results as a set of predictions about how the target, content, and valence of a given social act would compel someone to act on it, and in particular, to communicate with others about it.

Results

Based on the 48 scenarios produced by the combinations of *target* (ingroup, outgroup or celebrity), *content* (eight domains), and *valence* (positive or negative), and using the factors described in Tables 1 and 2 and Equations (1-5), we calculated the model's action values that reflect the likelihood of the central agent communicating to a third-party receiver (i.e., gossiping). The results are then predictions about how the human social mind processes and responds to key social information (target, content, and valence).

Figure 2 shows the model action values according to the independent effects of *valence* (Figure 2A) and *target* (Figure 2B). (Note that "total benefits > total costs" does not mean that the action will necessarily be executed, only that it increases its likelihood; and thus the key findings are the comparative values of the bar graphs.)

For *valence*, our model finds (a) a fairly comparable degree of communication (i.e., gossiping) about positive and negative events; and at the same time (b) a slightly higher amount for negative events. These predictions, especially the first of comparable amounts, are partially at odds with the prevalent view and some evidence for gossiping, in which it is believed to be predominantly negative. The model suggests that studies thus far have perhaps inordinately focused on events of negative valence (see Foster, 2004).

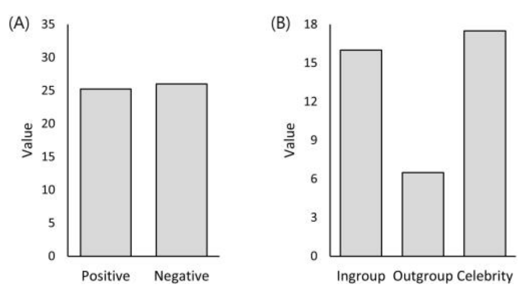

Figure 2. Predicted values of communicating social information based on its (A) valence and (B) target group.

For *target*, i.e., the person involved in the scenario event, the model predicts that celebrities will be discussed more, even more than those of one's ingroup — although again the difference between the two is not extreme (Figure 2B). Outgroup gossiping, however, is indeed predicted to be much less than the other two. The major factors underlying the target effect are *closeness* (i.e., the social distance between the target and the gossiper) and *status* of the target in terms of the larger societal hierarchy. Since outgroup is low in both status and closeness, the combination leads to the lowest amount of information spreading.

The difference between ingroup and celebrity, in contrast, is not as simple. In short, spreading ingroup and celebrity information both can have relatively high benefits (for ingroup, due to higher closeness; for celebrity, modest closeness and greater status effects; see Table 2); whereas spreading ingroup information can also lead to relatively higher costs (due to closeness, such as possible confrontation of the central agent by the target), leading to the Figure 2B result with celebrities more likely discussed even over ingroup members. Empirical evidence thus far is mixed, attesting to the need to computationally delineate the underlying factors (such as closeness, status, and the specific benefits and costs of information spreading), and to generate *a priori* predictions based on it (Foster, 2004).

Considering the potential interaction of the *target* and content *valence* of the scenario, we see a related but different predicted pattern (Figure 3A). For positively valenced scenarios (e.g., "Alex saved a child trapped in a burning building."), information involving ingroup members is predicted to be spread more than about either celebrities or outgroup members. This is due to the higher benefits yielded from ingroup information spreading, as well as the cost of ingroup information spreading decreasing dramatically and more so than in the other two target groups (e.g., no concerns about target retaliation).

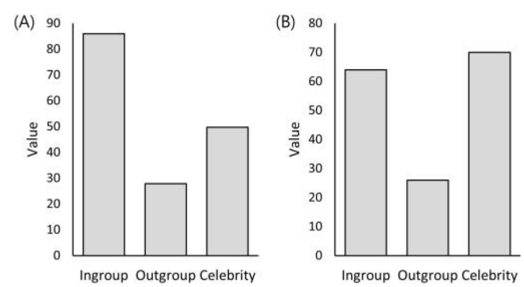

Figure 3: Predicted values of communicating information about the three target groups for scenarios of (A) positive or (B) negative valence.

For scenarios of negative valence (e.g., "Sam stabbed a person with a knife."), in contrast, the *cost* of ingroup information spreading is high; whereas the *benefit* of spreading celebrity information is high (especially due to status influence) and the risk relatively lower (Figure 3B). Therefore, with scenarios of negative valence, the model predicts that celebrity information dissemination will again be higher than for ingroup targets. Figure 4 shows more clearly the opposite patterns predicted for ingroup and celebrity targets based on the content valence.

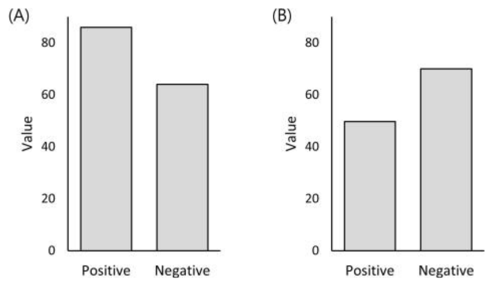

Figure 4: Predicted values of communicating information based on its valence for (A) ingroup and (B) celebrity.

Again, the empirical evidence is thus far mixed (Foster, 2004): for example, one study found that people spread more positive information about allies (friends and family) than for non-allies, including both strangers and those with high status (professors); and yet they also found more negative information spreading for *both* strangers and high-status people (McAndrew & Milenkovic, 2002). Moreover, another study found no differences in information spreading between ingroup and celebrity targets for either positive or negatively valenced events (Peng et al., 2015). Our model shows how specific factors need to be isolated to clarify the true nature of the current findings. Studies have yet to tease apart these factors sufficiently.

In sum, our computational model of social intelligence and communication has generated a number of predictions about how people evaluate the behavior of others, regarding who did it, what they did, and its general valence. These predictions include a higher degree of communication about ingroup targets when positive, and of celebrities when negative; as well as the relative merit or egregiousness of a wide range of behavior, enabling effective social interactions, generating stronger bonds among individuals, and preserving society more generally.

For model validation, we have also conducted a laboratory experiment using the same scenarios and asking participants whether they would communicate this information (e.g., Person X cheating on the final exam) to others. Thus far, our preliminary examination suggest that the model predictions are supported by the empirical results.

Discussion and Conclusions

For models of social cognition to capture human sociality they will need to include not only mind- and context-reading, but also a deeper understanding of social transactions and moral behavior. Indeed, people spend nearly 65% of their time discussing social events, suggesting that such events are especially cared about and processed (Dunbar et al., 1997). However, although a great deal is known about human social cognition, a comprehensive theoretical framework and computational model has been lacking.

To meet this challenge, we have first focused on the nuanced act of maintaining stable social behavior and societal structure by potentially communicating to others about another individual's behavior. To determine whether to do so, an individual (the central agent) must weigh its relative potential benefits and costs, which we have enumerated and quantified here. The model then makes an important theoretical contribution by producing a series of specific predictions about social communication that we are currently testing empirically (e.g., more communication about those socially close to us when their actions are positive, and more about those with higher status when negative; as well as the relative merit or egregiousness of a wide range of behavior). Indeed, in our preliminary examination of the data, the model predictions are supported.

With respect to generalizability and scale, we believe the model is poised to readily generalize and scale to larger amounts of social scenarios. For example, the critical features of all social events would be expected to resolve to *who did what*, and thus to the event's main individual(s) and the content of what happened. For target, people on first order are defined based on our relationships to them, which we have captured via ingroup, outgroup, and those of higher status. However, human understanding of others obviously goes beyond this, which in fact underscores what we believe will be the major contribution of our model: the prominence of *models of others' minds* within each agent. This component is poised to be developed substantially in the future.

Regarding content of the event, i.e., what the target did, we have organized social behavior into a comprehensive set of basic categories, ranging from deeply moral (such as *fairness*) to more everyday social activities (e.g., going to the movies). Most other types of social events are expected to fall into these categories, and thus should be readily added to the model with more detail and a mapping structure added.

Although *learning* is not yet built into the model, limiting generalizability, we have begun by adhering to evidence that suggests not only learning and cultural influences on human social cognition, but a significant underpinning of relatively hard-wired components. These include abilities such as mind-reading as well as having basic moral dimensions that appear to be universally shared and thus likely evolved, such as for *fairness* (Dunbar, Marriott, & Duncan, 1997; Dunbar, 2004; Haidt, 2007; Pinker, 2008). Learning and culture effects would then be expected to influence their relative weightings (e.g., Pinker, 2008).

In general, then, we believe a model that ultimately captures the richness of human sociality will entail both significant hard-wired components (reflecting evolutionary via genetics influences) and learned ones (reflecting culture and other environmental effects). To build such a model, we believe it is best to start with the basic more hard-wired foundations, and extend from there.

Beyond learning capabilities, an additional avenue of future development will be to determine whether the richness of human social intelligence is best captured by a symbolic-level model or ultimately resolved to a network-based one or some combination of both. In any event, we believe our framework and model help point the way forward.

References

Aronson, E., Wilson, T. D., Akert, R. M., & Sommers, S. R. (2016). *Social Psychology*. Pearson: NYC.

Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as cultural learning. *Review of General Psychology, 8(2),* 111–121.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In: J. H. Barkow, L. Cosmides, J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. Philosophical *Transactions of the Royal Society B: Biological Sciences, 351(1346)*, 1413–1420.

Dunbar, R. I. M., Marriott, A., & Duncan, N. D. C. (1997). Human conversational behavior. *Human Nature, 8(3),* 231–246.

Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology, 8(2),* 100–110.

Fernandes, S., Kapoor, H., & Karandikar, S. (2017). Do We Gossip for Moral Reasons? The Intersection of Moral Foundations and Gossip. *Basic and Applied Social Psychology, 39(4),* 218-230.

Foster, E.K., (2004). Research on gossip: taxonomy, methods, and future directions. *Review of General Psychology, 8,* 78–99.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2013). *Cognitive neuroscience: the biology of the mind.* WW Norton & Company.

Glimcher, P. W., & Fehr, E. (2014). *Neuroeconomics: Decision making and the brain*. Oxford: Academic Press.

Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316 (5827),* 998–1002.

Kralik, J. D. (2017). Architectural design of mind & brain from an evolutionary perspective. Proc. AAAI 2017 Fall Symposium: A Standard Model of the Mind.

Kralik, J. D. (2018). Core High-Level Cognitive Abilities Derived from Hunter-Gatherer Shelter Building. In I. Juvina, J. Houpt, & C. Myers (Eds.), Proc. of the 16th International Conference on Cognitive Modeling (pp. 49-54). Madison, WI: University of Wisconsin.

Kralik, J. D., Lee, J. H., Rosenbloom, P. S., Jackson, Jr., P. C., Epstein, S. L., Romero, O. J., Sanz, R., Larue, O., Schmidtke, H. R., Lee, S. W., McGreggor, K. (2018). Metacognition for a Common Model of Cognition. *Procedia Computer Science*: 145, 730–739.

Lee, J., Kralik, J. D., & Jeong, J. (2018). A Sociocognitive-Neuroeconomic Model of Social Information Communication: To Speak Directly or To Gossip. CogSci 2018.

Lee, J., Kralik, J. D., & Jeong, J. (2019). A General Architecture for Social Intelligence in the Human Mind & Brain. AAAI Fall Symposium 2018.

Rai, T. S. Thinking in Societies and Cultures. (2012). In: Holyoak, K. J., & Morrison, R. G. (Eds.). *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press.

McAndrew, F. T. & Milenkovic, M. A. (2002). Of tabloids and family secrets: *The evolutionary psychology of gossip, 32(5),* 1064-1082.

Peng, X., Li, Y., Wang, P., Mo, L., & Chen, Q. (2015) The ugly truth: negative gossip about celebrities and positive gossip about self entertain people in different ways. *Social Neuro, 10,* 320-336.

Peters, K., & Kashima, Y. (2015). Bad habit or social good? How perceptions of gossiper morality are related to gossip content. *European Journal of Social Psychology, 45,* 784-798.

Pinker, S. (2008). The Moral Instinct. *New York Times Magazine*.

Rosnow, R. L. (2001). Rumor and gossip in interpersonal interaction and beyond: A social exchange perspective. In R. M. Kowalski (Ed.), *Behaving badly: Aversive behaviors in interpersonal relationships* (pp. 203-232). Washington, DC: APA.

Russell, S. J., & Norvig, P. (2020). Artificial Intelligence: A Modern Approach. Pearson: NJ.

Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69(5),* 379–399.