Social Network Analysis and Mining (2020) 10:11
https://doi.org/10.1007/s13278-019-0620-8

ORIGINAL ARTICLE

Networked experiments and modeling for producing collective identity in a group of human subjects using an iterative abduction framework

Vanessa Cedeno-Mieles[2,4,6] · Zhihao Hu[5] · Yihui Ren[7] · Xinwei Deng[5] · Abhijin Adiga[1] · Christopher Barrett[1,3] · Noshir Contractor[9] · Saliya Ekanayake[12] · Joshua M. Epstein[10] · Brian J. Goode[2] · Gizem Korkmaz[1] · Chris J. Kuhlman[1] · Dustin Machi[1] · Michael W. Macy[11] · Madhav V. Marathe[1,3] · Naren Ramakrishnan[4,8] · S. S. Ravi[1] · Parang Saraf[8] · Nathan Self[8]

Received: 6 March 2019 / Revised: 9 December 2019 / Accepted: 16 December 2019

Abstract

Group or collective identity is an individual's cognitive, moral, and emotional connection with a broader community, category, practice, or institution. There are many different contexts in which collective identity operates, and a host of application domains where collective identity is important. Collective identity is studied across myriad academic disciplines. Consequently, there is interest in understanding the collective identity formation process. In laboratory and other settings, collective identity is fostered through priming a group of human subjects. However, there have been no works in developing agent-based models for simulating collective identity formation processes. Our focus is understanding a game that is designed to produce collective identity within a group. To study this process, we build an online game platform; perform and analyze controlled laboratory experiments involving teams; build, exercise, and evaluate network-based agent-based models; and form and evaluate hypotheses about collective identity. We conduct these steps in multiple abductive iterations of experiments and modeling to improve our understanding of collective identity as this looping process unfolds. Our work serves as an exemplar of using abductive looping in the social sciences. Findings on collective identity include the observation that increased team performance in the game, resulting in increased monetary earnings for all players, did not produce a measured increase in collective identity among them.

Keywords Online social experiments · Agent-based models · Abduction · Abductive loop · Collective identity

1 Introduction

1.1 Background and motivation

1.1.1 Collective identity: types, contexts, and applications

Group or collective identity (CI) is an individual's cognitive, moral, and emotional connection with a broader community, category, practice, or institution (Polletta and Jasper 2001).[1] There are several themes of, and implications for, CI, including: (1) an individual's willingness to place the needs of the group above personal needs [e.g., contributions to Public Goods Games (PGGs) (Charness et al. 2014; Brewer and Gardner 1996)]; (2) a person's susceptibility to positive social influence from group members [e.g., sensitivity to evaluations from a collective group (Charness et al. 2014; Brewer and Gardner 1996)]; (3) one's desire to differentiate from others not in the collective [e.g., allocation between in-group/out-group (Bornstein and Yaniv 1998)]; (4) an individual's willingness to enforce conformity to group norms established by the collective identity (McAuliffe and Dunham 2015; Brewer and Gardner 1996; Kozlowski and

[1] There are other definitions for collective identity (CI). For example, McFarland et al. (2014) state that collective identity means that members become more familiar and equal. Wendt (1994) defines CI as the positive identification with the welfare of another, such that the other is seen as a cognitive extension of the self, rather than independent. See Owens (2006) for a discussion of various definitions of CI.

✉ Vanessa Cedeno-Mieles
vcedeno@vt.edu

Extended author information available on the last page of the article

Published online: 07 January 2020

Ilgen 2006; DeChurch and Mesmer-Magnus 2010); and (5) a person deriving self-esteem from the group (Tajfel 1974; Kozegi 2006). Hence, there are many behavioral and attitudinal manifestations as consequences of CI.

There are many types of, and contexts for, collective identity, including: (1) religious identity (Peek 2005; Benjamin et al. 2016), (2) philosophical identity (Muller 1996; Greene 2000), (3) gender identity (Cameron 1997; Butler 1988), (4) (sports) fan identity (Snow 2001), (5) labor movements (Goldberg 2003), (6) social movements such as African-American civil rights, women's suffrage, gay rights (Taylor and Whittier 1992; Polletta and Jasper 2001; Snow and McAdams 2000), (7) political identities (Plutzer and Zipp 1996; Juergensmeyer 2003), (8) racial and ethnic identities (Tatum 2003; Alexander et al. 2004; Nagel 1996; Eriksen 2010), (9) national and cultural identities (Manchester 1993; Alexander et al. 2004; MacGregor 2018), and (10) ideologies (van Dijk 2000).

CI is a widely studied concept across academic disciplines. Extensive experimental research in social science, political science, psychology, biology, geography, anthropology, religion, criminology, philosophy, and economics shows that CI influences human decision making (Kahn and Ryen 1972; Paris et al. 1972; Goldman et al. 1977; Worchel et al. 1977; Erikson 1980; Oldenquist 1982; Brewer and Gardner 1996; Perdue et al. 1990; Brody 2000; Rousseau and van der Veen 2005; Currarini et al. 2009; van Zomeren et al. 2008; Lustick 2000; Silke 2008; Eriksen 2010; Pilny et al. 2017; Suri and Watts 2011; Shank et al. 2015; Brewer 1991; Benjamin et al. 2016).

There is a host of applications for which CI is important, including team formation, maintenance, and behavior in organizations and communities (Kozlowski and Ilgen 2006; DeChurch and Mesmer-Magnus 2010). The ability to generate identity within (marginalized) groups, e.g., through sacred values, is an important aspect of violent group formation (Silke 2008; Atran et al. 2014a, b). These are compounded by effects of culture and ethnicity (Gilwhite 2001; Atran et al. 2007; Ginges and Atran 2013). International relations are affected by CI among independent states (Wendt 1994). Political leaders of minority or marginalized groups may control identity narratives to persuade their constituents of posturing with governments (Choup 2008). Relatedly, CI is a cohesive force for groups fighting governments to secure rights and indigenous lands (Snow and McAdams 2000; Brody 2000). Religious identity can be a source of stability for immigrants assimilating into a new country (Peek 2005). Language and preservation of culture are intimately tied to collective or group identity (Brody 2000). Ramifications of a lack of identity are studied in Stout et al. (2017).

Individuals may possess several group identities, with different degrees of salience (strength of affinity and association), such that multiple identities may be simultaneously operative (Snow 2001; Peek 2005; Benjamin et al. 2016). There may be a hierarchy of identities, with different identities coming to the fore in different situations (Stryker 1980). [The ability to use different identities in different situations has been referred to as *freedom* in a philosophical context (Heller 2019)]. Multiple identities may also be negatively correlated, e.g., religious and national identities (Verkuyten and Yildiz 2007). Furthermore, identities and their saliences may be transient over short time scales, and may ebb and flow over longer time scales (Butler 1988; Snow 2001; Vryan et al. 2003; Owens 2006; Benjamin et al. 2016). Consequently, a person's identity may include a combination of dynamically changing, hierarchical collective identities.

Relationships between CI and other phenomena can be intricate. We take collective action (CA), for which there is a massive literature (e.g., Olson 1965; Granovetter 1978; Schelling 2006; Tarrow 2010), as an example. Causal relations between CI and CA are very complicated, with the causal direction between the two changing for different circumstances (Wendt 1994; Polletta and Jasper 2001; Snow 2001; Choup 2008; Fominaya 2010).

1.1.2 Dimensions of collective identity to study

These issues make the study of CI both interesting and challenging. It is the generality of the concept of CI, its application in a wide range of contexts, its many types and its ramifications for humans and their behaviors that have led to myriad CI studies since the term *collective identity* was first coined by Durkheim some 65-plus years ago (Durkheim 1951). *Our focus here is the CI formation process: how CI is formed among a group of people.* We now overview work on the CI formation process to set up what is entailed in our "focus on the CI formation process".

CI *formation* is studied in several works (Wendt 1994; Brewer and Gardner 1996; Peek 2005; Choup 2008; Greenhill 2008; Chen and Li 2009; Ackland and O'Neil 2011; Charness et al. 2014; Swanson 2015; Brunsdon 2017; Pilny et al. 2017). See Related Work, Sect. 3.3. All of these works, except one, are empirical, examining events in the field, under varying circumstances. Surveys, questionnaires, and interviews with human subjects are used to establish, through expert judgment, whether CI has formed within a group.

The work by Charness et al. (2014) also studies CI formation, but is quite different from these other works. They use controlled experiments to produce CI among human subjects through priming using team anagram games, wherein players work cooperatively to form words from a collection of letters that they are given. For example, letters *t*, *c*, *a*, and *s* can be used to form words such as *cat* and *cats*. There are many other aspects to their game. Group identity was then measured after the anagram game using a public goods game (PGG). The greater the PGG contributions of individuals to the team, the greater the collective identity of these individuals. It was found that the priming activity increased PGG contributions. To the best of our knowledge, these are the only controlled experiments that seek to produce CI through priming (in an anagram game) and measure CI quantitatively (through the proxy of PGG contributions). Charness et al. (2014) influence our work herein.

We note in passing that priming tasks are central in social and economic experiments (e.g., Drouvelis et al. 2010; Feher et al. 2016; Smith et al. 2017) and are therefore worthy of study for this reason alone.

In addition to the references above, CI formation is discussed and theorized about in Melucci (1995), Melucci (1989), Snow and McAdams (2000), Snow (2001), Polletta and Jasper (2001), Tarrow (2010), Fominaya (2010). We note that these theoretical works are descriptive and qualitative in nature and are *not* concerned with computational modeling.

1.1.3 Dimensions of our work for collective identity

From the foregoing, we specify what is entailed in “focusing on the CI formation process” for this work. First, there is no agreed-upon method for generating CI. So specifying a priming or CI-producing activity is non-trivial. Our priming game is motivated by the one in Charness et al. (2014), but our game is significantly different, and we describe how, and the technical challenges that exist, in Sect. 1.4. Since we want to study CI formation, we must measure it quantitatively. This, also, is non-trivial and is discussed in Sect. 1.4. Only Charness et al. (2014) measures CI quantitatively in a game; other CI-producing works resort to qualitative methods, which is surely an indication of its difficulty. We want not only to experimentally study CI, but also to model the process. Despite all of the work on CI (described here and in Related Work, Sect. 3), we know of no works that quantitatively model any CI formation process.[2] In summary, a focus on the CI formation process includes specifying experiments; conducting and analyzing experimental data; quantitatively measuring CI; and modeling both the CI formation and measurement processes. We also use CI as an exemplar for using abductive iterations in the CI formation process. *This work makes contributions in all of these areas.*

Fig. 1 The main components of the experimental setup: (i) players are recruited through Amazon Mechanical Turk; (ii) players collectively participate as a team in a *priming activity* in the form of a *group anagram game* with the goal of *producing* collective identity among the players; and (iii) players individually state their affinity for (i.e., identity with) the team through a *dynamic identity fusion index* (DIFI) task (Swann et al. 2009) that we take as quantifying or *measuring* the level of CI formed in the anagram game. The modeling effort is to model the second and third steps of this process. Note that because the group anagram game is more complex than the DIFI task, more work is required to produce the experimental platform, to conduct experiments and to analyze the experimental data, and to build models of player actions in this priming activity. Furthermore, it is in the priming activity that game conditions are altered, and it is these conditions that we seek to correlate with CI

After our work is overviewed in the next two subsections, we address, in turn, technical challenges, the novelty of our work, and our contributions, and these address all of the topics just itemized. In the next section, we translate these topics into work tasks.

1.2 Overview of work scope

Our work has three broad elements. First, we develop an online experiment, motivated by the work of Charness et al. (2014), that is designed to produce CI within a group of participants, through priming, and then measure the CI produced. See Fig. 1. Specifically, the priming activity consists of players cooperating in a group anagram game, where participants share letters with their neighbors in order to help all players form more words. Specifically, the main player actions are: (1) requesting letters from neighbors, (2) replying to letter requests of neighbors, and (3) forming words. This activity is intended to *produce* CI among a group of players. This priming activity is accompanied by a dynamic identity fusion index (DIFI) task that measures how much a person associates with the team or group as a result of playing the group anagram game (it is a proxy for CI). Unlike the group anagram game, the DIFI task is done individually, in isolation. This DIFI task is to *measure* the CI formed in the group anagram game.

For the second element of our work, we construct models of the CI priming process (the group anagram game) and of the DIFI task, and compare predictions of player behavior against experiments in the group anagram game. We develop and evaluate three agent-based models for the group anagram game, and a statistical model for the DIFI task. Because the group anagram game is much more intricate than the DIFI task, more effort is devoted to the former activity.

As the third element of our work, we use abduction as our framework for this study where both experimental work and modeling work take place within an abductive

[2] We use the term *model* to mean a representation of equations and algorithms to compute some result. In contrast, in the social and some other sciences, *model* often refers to a qualitative (textual) description of some process that is much more conceptual and not computational. Our models that we present herein are of the first type. We use the term *model* in the former (quantitative) sense, unless otherwise specified.

loop framework. We now address abductive iterations, since they tie together experiments and modeling.

1.3 Overview of our experiment and modeling approach: abductive iterations

Abduction is an inference approach that uses data and observations to identify plausible (preferably, best) explanations for phenomena (Pierce 1931; Flach and Kakas 2010). That is, abduction is reasoning from effects to causes (Chamiak and Santos 1992). Effects are often generated by results from (laboratory) experiments or in situ observations of systems. One then constructs hypotheses and identifies or develops theories that explain these observations.

Much of the work on abduction has focused on topics such as producing explanations for different logic settings (e.g., Echenim et al. 2013); determining the computational complexity of abduction problems (e.g., Wei-Kleiner et al. 2014); and generating solutions for special problems or transformations that are useful in obtaining solutions (e.g., Pfandler et al. 2013). Abduction has broad application in robotics, genetics, automated systems, and image understanding (Shanahan 2005; Andrews and Bonner 2011; Vanderhaegen and Caulier 2011; Juba 2016).

However, in contrast to the above notion of abduction, our focus is the specification and implementation of an abductive *looping* process, wherein abduction is executed in successive iterations. Every iteration builds off of all previous ones, so that explanations may evolve from accumulated data from experiments and observations. As a differentiator from previous work, our interests are behaviors and human interactions within networked groups in the social sciences (Contractor 2019). In particular, our exemplar is to understand whether a cooperative game can produce collective identity (CI) within a group.

The abductive loop (AL) process that we employ is described in Sect. 2, but among its components are experiments and modeling, and we make note of works on coupling experiments and modeling here. There have been several controlled experimental studies of comparable size to our experiments (e.g., Kearns et al. 2009; Judd et al. 2010; Kearns et al. 2012). Also, empirically grounded, data-driven modeling of human behavior is done (Mason and Watts 2012; Li et al. 2014; Nguyen et al. 2017; Zhang et al. 2016). We combine these two ideas, in a particular way that is guided by abduction, and perform them iteratively. The proposed abductive analysis is to form hypotheses to evaluate theories as part of the looping process, and develop new insights about CI. Our approach provides an exemplary case of coupling theory development/evaluation with real problems.

1.4 Technical challenges

There are several technical challenges in our work. The first two, and foremost, challenges are generating and measuring CI in a group setting. CI can be a transient phenomena, so it can be generated in a group that is newly formed. A group anagram game has been successfully used to produce CI among people that do not know each other (see Sect. 3.6 of Related Work). This is the only previous work in a group setting where CI was produced over a single game or encounter of relatively short duration and then quantitatively measured. To our knowledge, ours is the first work to attempt CI formation under similar conditions, but in addition, in an online setting. We also use a group anagram game (our game is different from the one referred to above, but is motivated by that work). In our setting, however, a team's or group's members do not have the benefit of observing facial expressions and body language as was done in Charness et al. (2014). By comparison, other works seek to form CI within groups that meet regularly over periods weeks or months [see Sect. 3.3, e.g., Swanson (2015)]. Still other works [see Sect. 3.3, e.g., Peek (2005), Choup (2008)] rely on CI being formed over years, based on cultural and other factors. The point is that there are many works that seek to produce CI in different ways, so studying methods of producing CI is important. A second point is that our requirement to produce CI quickly in an online game, with no use of visual cues or body language, is novel and extremely challenging.

Also, our method of measuring CI quantitatively has never been used in relatively short-duration settings. Specifically, we use the DIFI score (Jiménez et al. 2016) to measure CI formed in the group anagram game. Previously, DIFI score has been used to *pre-select* people from a larger population that have *already formed* CI, often based on cultural or other long-lived attributes that germinate and grow over years, and that was formed in situ [see Sect. 3.4, e.g., Swann et al. (2010a)]. In Charness et al. (2014), a public goods game was used to quantify CI among group members. Several other works (e.g., Swanson 2015) use a binary scale of determining the existence of CI (i.e., "no CI formed," or "CI formed") based on opinions of experts that, for example, base their subjective determinations on language used by group members. Our challenge here is to determine whether we can use DIFI score to characterize CI; we are using the DIFI method for a different class of CI, for which it has not been used. Quantitative methods of measuring CI, are few, and it has very rarely been done.

From this description thus far, it should be clear that we have unique and challenging setups and goals for producing and measuring CI. There are other challenges.

Our goal is to study quantitatively CI within an abductive loop setting. Looping over abductive analyses is relatively rare (see the robotics work Shanahan (2005) as an exception), and the use of abduction and abductive iterations in the social sciences is very rare.

Furthermore, our abductive analyses involve both experiments and modeling. Most abductive work focuses solely on using experiments within loops. Here, we make modeling a first-class element of the abductive process. In fact, we know of no work whatsoever in modeling a priming activity designed to produce CI among interacting members of a group. Developing models is all the more challenging because our game consists of multiple player actions (versus binary choice games) that can be repeated over time (versus one-shot games) in any user-defined sequence of repeated actions, where neighboring player actions (or inaction) can influence action choices among players.

Also, there are finite resources and a finite pool of candidate players for conducting experiments, using the constraint that we want participants to play the game only one time. (Amazon Mechanical Turk does not provide a gateway into an infinite pool of players). A challenge is to specify game conditions that produce measurable CI while also providing data over a range of input conditions to enable development of models that are more general than the experimental conditions.

We believe that this work pushes the state of the art in all of these dimensions. Moreover, we attempt to interpret our CI results in a way that focuses on player interactions, and not specifically on our group anagram game. Consequently, while we believe that this work achieves several firsts (see the novelty section immediately below), this work is not the final word on CI, abductive analyses, and conducting and modeling group CI experiments. Rather, we view it as a beginning: a beginning that opens new avenues for producing and quantitatively measuring CI.

1.5 Novelty of our work

The novelty of our work is in the areas of research process (through abductive iterations), experiments, modeling, and social science (through CI). Specific novelties include:

1. Performing the first online group anagram game (for producing CI) and using a dynamic fusion index (to measure CI);
2. Conducting experiments where players can choose actions from a candidate set; these actions can be repeated any number of times over a specified time duration; players interact, cooperate, and can affect others' subsequent actions (versus binary choice, one-shot games);
3. First-of-their-kind experimental results and implications for CI from online experiments;
4. First modeling of this temporal, multi-action, interacting group anagram game;

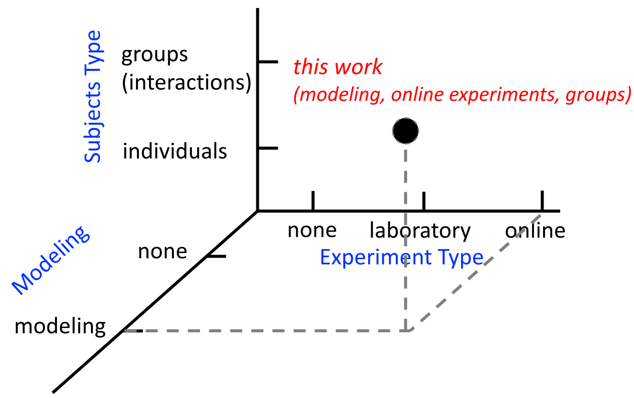

Fig. 2 Conceptual view of three dimensions of this work, illustrating how our group anagram game study (for producing CI) and our approach for measuring the CI produced is situated. Our experiments consist of online web-based human subjects experiments. Our modeling component consists of model and algorithm development, and agent-based modeling. We study groups of interacting individuals. To our knowledge, this combination of study features is unique. These dimensions are used in Table 1 to compare our work with those of others

5. First evaluation of these models by comparing model predictions with experimental results;
6. One of the first uses of abductive looping in a social science context, and the first use in the study CI.
7. One of the first demonstrations of abductive looping to test (social) theories.

Figure 2 shows the unique context of our work along three dimensions of experiments, modeling, and types of experimental subjects. Table 1 makes this more concrete by presenting representative works along various combinations of values along the three dimensions of Fig. 2. It is clear that our work—identified at the bottom right of the table—is unique.

1.6 Contributions

Our major contributions follow.

1. Insights on the collaborative anagram game (Sect. 4) We present novel experimental data that illustrate how players interact in group anagram games played through an online game/experimental platform. We focus on experimental data that are useful in modeling. We find that letter requests and letter replies are made throughout the game, rather than solely at the outset. However, if there are few neighbors ($k = 2$) and consequently fewer available letters (3 letters per neighbor), there are fewer letter requests and letter replies near the end of the game. Also, players generally respond relatively quickly to their neighbors' letter requests: replies are typically made within 30 s of the request. In the same

Table 1 A hierarchy of different collective identity (CI) experiments in the literature, which puts the uniqueness of our work on anagram game experiments and modeling into the context of the works of others. For each of in-laboratory and online environments, there are categories of works on individual subjects and groups of subjects. With groups of subjects, we are interested in interactions among these subjects. Then, for each of these four categories, we break those works that are experiments only (i.e., Exp. Only), and those works that combine experiments and modeling (i.e., Exp. & Modeling). The last row has one representative work within each category; some categories have no work. Our work (labeled **This Work**) studies online experiments of collective identity with modeling and interaction among subjects. It is the *only* work that combines experiments and modeling of human subjects in a group setting. (These references are not exhaustive; more detail is included in the related work of Sect. 3. However, for the categories labeled with *No Work*, there is no work in the literature, to the best of our knowledge)

In-laboratory				Online			
Individual subjects		Group of subjects		Individual subjects		Group of subjects	
Exp. Only	Exp. & Modeling	Exp. Only	Exp. & Modeling	Exp. Only	Exp. & Modeling	Exp. Only	Exp. & Modeling
Brewer and Silver (1978)	Rousseau and van der Veen (2005)	Worchel et al. (1977)	*No work*	Pilny et al. (2017)	Ackland and O'Neil (2011)	*No work*	**This work**

way as letter requests and letter replies, word submissions are made throughout the 5-min game, but the numbers of neighbors and available letters do not affect this type of action.

2. Data-driven networked agent-based models (ABMs) of experiments: design, construction, and evaluation (Sect. 5) We design, construct, and evaluate three data-driven ABMs of the group anagram game experiment. We adapt a conditional random fields (CRF) (Sutton and McCallum 2011) modeling approach with four parameters to flexibly incorporate history effects on agent actions that evolve in time. That is, our models predict time histories of player actions in the group anagram game. These actions are: (1) requesting letters from neighbors, (2) replying to letter requests of neighbors, (3) forming words, and (4) thinking (or idling). We capture these activities through a state transition matrix approach, where, in our most sophisticated model, the action at time $(t + 1)$ is based on the action at time t and on a feature vector that captures an individual's state. Our approach can alleviate the overfitting problem that would arise with, e.g., a static Markov model that would require capturing many more state transitions.

ABM is used as our simulation modeling approach because of its fine granularity and for its generative properties (Epstein 2007). That is, local interactions produce population-level dynamics. We use inductive and deductive inference in three ways, use KL divergence to compare model predictions with experimental data, and compare results across multiple ABMs. For example, our KL-divergence evaluations are broken down by ABM, player action, and number of neighbors in a game. For each combination, we use overall data at the end of the 5-min group anagram game and at 1-min intervals during the game to evaluate temporal effects. All of these are used to demonstrate that the ABMs successively improve with the process of incorporating more data that enables greater modeling sophistication.

Our three successive ABMs are named M0, M1, and M2. Our work in evaluating the ABMs shows that ABM M1 reduces KL-divergence values by 4× or more, over those for ABM M0, in many cases. (Smaller KL-divergence values are better; they indicate that model predictions are in better agreement with experimental data. A KL-divergence value of zero means that a distribution from a model and from data are interchangeable.) Our work also shows that in many cases, ABM M2 has KL-divergence values that are 4× or more reduced from those of ABM M1. Interestingly, ABM M1 does slightly better than our most sophisticated model (ABM M2) for a small range of parameters that were used in generating M1, but M2 does much better over the entire input parameter space.

3. Specification and demonstration of iterative abductive analysis process (Sects. 2 and 7) We perform experiments (Contribution 1), and modeling and evaluation (Contribution 2), within an iterative abductive process. Using Haig (2005), Timmermans and Tavory (2012) as a starting point, we explicitly incorporate modeling and iterations into the abductive process. The latter necessitates specifying what is to be done in the next iteration. The iterative process is successfully demonstrated through the group anagram experiments, agent-based modeling, and hypothesis generation and testing. *The proposed abductive process can be considered as a general methodology for other social science researches. Although the methodology is general, we provide considerable detail in both the experiments and modeling, and their interactions (e.g., how they complement each other), that we believe will be helpful for other social science studies.* For example, our method of model construction from data (see Contribution 2 above) can be used to capture other temporal human action sets among interacting agents.

4. Statistical analysis of numbers of samples required for modeling (Sect. 6) We evaluate the quality of our state transition matrices of our ABMs using a root of mean squared errors (RMSE) approach. Specifically, we are interested in how many test samples are required to achieve a specified small error in predicted transition probabilities, which are an integral part of our ABMs, as compared to measured transition probabilities. We use our feature vector from the ABM and break each element down into bins, and add to it the dimension of number of neighbors that a player has in a game. By evaluating all of the successive state transition pairs among the actions, i.e., action $a(t)$ at time t and the next action $a(t+1)$ at time $(t+1)$, within each of the resulting 324 distinct bins of data, we find that the minimum number of observations (samples) for each state transition clearly demarcates small from large RMSE. The data show that small RMSE values result when a state transition has at least 100 observations.

5. New experimental understanding of the formation of collective identity (CI), (Sects. 4 and 7) First, we demonstrate that our group anagram game can prime game players to form CI (Sect. 7.3). Second, we discover three novel insights on the formation of CI by coupling the team anagram game and DIFI score. (a) Players' DIFI scores increase with increasing numbers of neighbors in the anagram game. Since DIFI score is our proxy for CI, this implies that a player's sense of CI increases with increasing numbers of neighbors (Sect. 7.4). (b) The number of interactions increases as number of neighbors of a player increases from two to four. However, the numbers of interactions, relatively speaking, saturates with further increases in degree, up to a degree of eight (Sect. 7.4). (c) Despite this saturation in numbers of interactions, the DIFI score continues to increase with degree, suggesting complicated interactions among game parameters (Sect. 7.4).

Third, what we did not find is interesting (Sect. 7.3). Specifically, the number of words formed does not correlate with players' feelings of CI (as measured by the DIFI score). We conjectured that since players' earnings are directly tied to the number of words that they generate, players would deem this outcome important. That is, the more words formed by the group, the greater their earnings and hence the greater the success achieved by the group. Success breeds cohesion. We therefore hypothesized that the greater the number of words formed, the greater the CI among team members. We found that this is not the case: number of words formed is not significantly correlated with DIFI score. Rather, numbers of interactions correlated more strongly with DIFI score (Sect. 7.3).

Fourth, these experimental observations, and hypotheses that go along with them, are made in the context of the abductive loop (see Sect. 7). To the best of our knowledge, these are the first experimental results of this kind. In this process of conducting two iterations, we demonstrate: constructing hypotheses, testing theories, falsifying hypotheses, finding support for hypotheses in full or partially, and finding support for multiple hypothesis from the same observations that require disambiguation in more abductive iterations. We cast our hypotheses in general terms, using only degrees of players in networks, numbers of interactions in games, and game rewards, i.e., abstracting away our particular game conditions, thus enabling testing of our findings by other researchers, potentially using different games.

1.7 Extensions from the conference paper

This paper was originally published as Ren et al. (2018). Extensions of that work, presented herein, are summarized as follows. (1) Introduction has been expanded to give fuller treatment of background, motivation, and problem context. (2) Related work is expanded with more detail and new topics. (3) Game description has more detail. (4) Experimental data from the game are given with new insights on player behavior. (5) Fuller treatment of the development of each of the three ABMs (M0, M1, and M2). (6) Fuller treatment of comparisons of model predictions with experimental data. (7) Additional model evaluation and data, comparing model predictions to experimental results across games. (8) Enhanced description and results in error analysis, comparing experiments and models.

1.8 Paper organization

Figure 3 shows the technical sections of this paper, and their relationships. An overview of the abductive loop process is presented in Sect. 2, providing a framework for the rest of the paper. Related work is in Sect. 3. The group anagram experiments and results are described in Sect. 4. Models of the experiments are developed in Sect. 5. Model predictions are compared to experimental data. Section 6 contains error analyses of the models. Sections 4 through 6 contain the major technical components of the abductive loop that is overviewed in Sect. 2. These analyses and results enable a more streamlined description of the particular abductive loops executed in this work in studying CI in Sect. 7, so that the abductive process is clear. In Sect. 7, the relevant sections of the experiments and modeling are referenced. This also makes the experiments and modeling more clear. Furthermore, we discuss the generalizable knowledge gained from the CI study, how CI is formed and not formed, and because we state our hypotheses in terms of interactions, how other experiments can be undertaken to confirm or

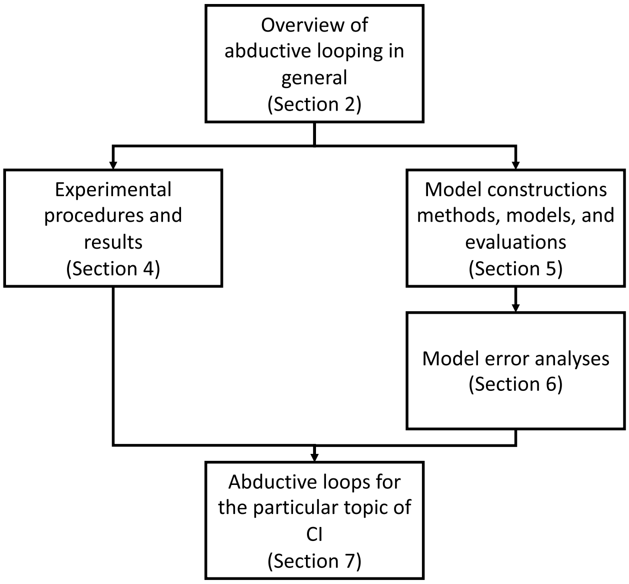

Fig. 3 Relationships among the sections of the paper that describe the technical work. Section 4 presents both the CI formation game (group anagram game) and the CI measurement task (DIFI). Section 5 describes the CI formation models. The model for predicting CI measurement (DIFI score) is given within the abductive iterations in Sect. 7. Section 6 analyzes errors in the anagram model predictions, compared to experimental data. Section 7 takes key points from the experiments and modeling and presents them within the framework of abductive looping. This section also contains additional analyses of experimental data, hypotheses about CI and their evaluations, and selected modeling results. Section 7 is a culmination of all the work in the preceding sections. Sections not appearing in the figure are Related Work (Sect. 3), Limitations of the work (Sect. 8), and Summary (Sect. 9)

contradict our results. Limitations of this work are presented in Sect. 8. Section 9 summarizes. Sections 4 and 5 are substantial in size. Consequently, we provide tables within these sections to organize the work and guide the reader, and we present many of the results in the Ph.D. dissertation of Cedeno (2019).

2 Overview of Abductive Loop

Figure 4 illustrates our iterative abductive process, which includes inductive and deductive steps and hypothesis testing. All work in this paper takes place within this framework. This structure follows that of Haig (2005), Timmermans and Tavory (2012), which are based on Piercian abduction (Pierce 1931), but augments it in key areas. Note that in contrast to confirmatory (deductive) analyses, where theories, hypotheses, and models are developed *first*, and used to predict results of future candidate experiments, one-step abduction first generates data through experiments or observations. (Abduction uses data to drive the scientific

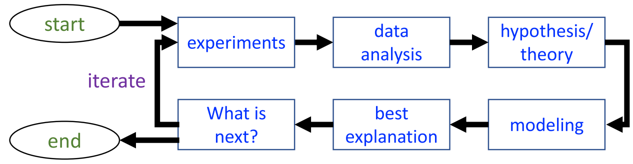

Fig. 4 Steps in our iterative abductive analysis/loop

discovery process.) Then, data analysis consists of searching for *patterns* and generalizing these into *phenomena*, which is an inductive step. These results are used to formulate hypotheses based on theories whose purpose is to explain the data. Hypotheses may exist (e.g., from a previous loop) or may be proposed in this step, and can be removed (e.g., via falsification). Multiple candidate theories may be posed for a given phenomena. Models are developed from the data, with the objective of generating outputs that help evaluate hypotheses and theories, and/or help guide experiments for the next loop. The best explanation, or hypothesis/theory appraisal, is the process of identifying the best explanation for the phenomena (Thagard 1989); this includes hypothesis falsification. Finally, the last step in an iteration is to determine what to do next, in terms of designing new experiments. The iterative process may terminate for any number of reasons; e.g., a best explanation has been found.

This description provides the structure for the rest of the paper. The experimental work of Fig. 4 is described in Sect. 4, after related work. The modeling work in Fig. 4 is presented in Sects. 5 and 6. We provide the experimental and modeling methodologies, data and results in these sections because they are too large to fit within a discussion of results from the abductive iterations. Following these sections, we return to the abductive loop and reference experimental and modeling results as appropriate, to make the looping process and results more streamlined and cogent (and provide additional results).

3 Related work

Related work topics are provided in Table 2, along with each topic's relevance to our work. Each subsection below provides research for one row in the table.

3.1 Overviews of CI

Overviews of CI are provided in Tajfel (1974), Abrams and Hogg (1990), Owens (2006), Vryan et al. (2003), Fiske et al. (2010), Hogg and Abrams (2007), Snow (2001). Peek (2005) provides an interesting view of CI as a combination of social structure (through roles) and processes (via perceptions and interactions) (Melucci 1995).

Table 2 Research literature topics addressed in Sect. 3, related work. Selected sections are presented here, while others appear in "Appendix A"

Section of related work	Name	Relevance
3.1	Overviews of CI	CI is a broad topic. These are surveys of CI for the interested reader
3.2	Individual Anagram Games: Experiments	Individual anagram games are precursors to group anagram games and have been extensively studied for more than 60 years to analyze the effects of goal setting, compensation types, internal-external attributions, and test anxiety. It includes a broad range of disciplines like sociology, economics, management science, and (social) psychology. For our work, anagram games are priming activities
A.1	Individual Anagram Games: Modeling	With all of the experimental work on anagram games, it is surprising that very little work has been done in modeling and simulating these games
A.2	Individual Anagram Games: Experiments and Modeling	Few works combining experiments and modeling of individual anagram games exist (Feather 1969; Feather and Simon 1971a, b)
3.3	Collective Identity-Based Experiments: Formation of CI	Our work is motivated by CI, and in particular the CI formation process. These works study different methods from ours in generating CI
3.4	Collective Identity-Based Experiments: Implications of CI	Along with the Introduction, this section provides works that demonstrate the implications of CI, thus motivating why we study it
3.5	Measurement of CI	Methods used in research to measure (quantify) CI are important.
3.6	Combined Group Anagram and CI Experiments	This section emphasizes that there is only one work on group anagram game. That work motivated our work. However, there are differences between that work and ours
A.3	Modeling of CI	Demonstrates that there are few modeling studies of CI, and no works like ours
A.4	Agent-Based Models of Anagram Games and Formation of CI	This puts our preliminary results into context. The first and only work, to our knowledge, in modeling human group anagram games is our work Ren et al. (2018)
A.5	Studies of Phenomena Related to CI	As described in the Introduction, CI is relevant for and closely related to, many other phenomena like cooperation and collective action. These works provide some background on these works
A.6	Data-Driven: Combining Experiments and Data-Driven Modeling	Demonstrates that combined experimental and modeling studies, as we do here, are used for other phenomena besides CI
3.7	Modeling of Time Sequences of Actions	These are studies that investigate time series models. Our modeling and ABMs are essentially time series models
3.8	Evaluation of Model Predictions	Methods for comparing experimental and model prediction distributions, as we do here, are presented
3.9	Abduction and Abductive Loop	We use abductive iterations as a framework for our experimental and modeling work. We survey other abductive works

3.2 Individual anagram games: experiments

Over 20 experimental works use anagram games—with *individual* players (e.g., Mayzner and Tresselt 1958; Russell and Sarason 1965; Tresselt 1968; Warren and Thomson 1969; Dominowski 1969; Feather 1969; Feather and Simon 1971a, b; Davis and Davis 1972; Sarason 1973b, a; Miller and Ross 1975; Goldman et al. 1977; Gilhooly and Johnson 1978; Stones 1983; Locke and Latham 1990; Latham and Locke 1991; Vance and Colella 1990; Schweitzer et al. 2004; Cadsby et al. 2007, 2010)). An individual game means no interactions (e.g., sharing letters) between subjects playing a game at the same time.

We review anagram game studies that are purely experimental. In Stones (1983), experiments of anagram games are used to test player's specification of causality for their performance (e.g., did a player attribute good performance to skill or luck?). It was found that people more likely to be responsible for their own actions attributed success or failure to their own behavior, versus assigning outcomes to chance. Miller and Ross (1975) analyzed how individuals engage in attributions of causality. Situational factors were studied through anagram games in Davis and Davis (1972).

Effects of goal setting are analyzed with anagram tasks in Locke and Latham (1990), Vance and Colella (1990), Latham and Locke (1991), Schweitzer et al. (2004). In Vance and Colella (1990) players played the anagram game and their assigned goals became increasingly difficult. For example, for each goal trial, subjects were assigned a goal for the number of words they have to form. After each goal trial, subjects recorded their performance (i.e., the number of words formed) as well as their assigned goal for the next trial. Difficulty of assigned goal was increased by two words per trial. Before beginning the next trial, subjects completed a form on which they calculated their GDF (goal discrepancy feedback: performance minus assigned goal) and PDF (performance discrepancy feedback: performance this trial minus performance last trial). Assigned goals were rejected when GDF became sufficiently negative. GDF and PDF differed both in sign and magnitude of effects on acceptance and personal goals, indicating that subjects used these feedback discrepancies differently in the goal evaluation process. Unusually, personal goals and performance remained high even after assigned goals were rejected. In Locke and Latham (1990), Latham and Locke (1991), theories of goal settings are developed. In Schweitzer et al. (2004), it was found that people with unmet goals were more likely to engage in unethical behavior than people attempting to do their best.

Goldman et al. (1977) use the anagram task to examine three factors and their effects on group performance: intergroup competition or cooperation, intragroup competition or cooperation, and task means interdependence. In Russell and Sarason (1965), Sarason (1973b), Sarason (1973a), studies look at anxiety generated from performing a task, where the task is the anagram game. In Cadsby et al. (2007), pay-for-performance and fixed-salary compensation were compared using an anagram task. In Cadsby et al. (2010), an anagram game was employed as the experimental task to evaluate a target-based compensation system, a linear piece-rate system and a tournament-based bonus system. Larger amounts of cheating occurred under target-based compensation. In Mayzner and Tresselt (1958), Dominowski (1969), Warren and Thomson (1969), Gilhooly and Johnson (1978), the effects of letter order and word frequency on anagram game performance are analyzed.

3.3 Collective identity-based experiments: formation of CI

The following references study or theorize on the CI formation process. That is, they study processes by which a group of individuals that does not possess CI can form CI by, for example, interacting or undergoing a priming task.

In Brewer and Gardner (1996), laboratory experiments of CI with no interactions between subjects are performed using priming. They argue that the personal, relational, and collective levels of self-definition (shift from personal to collective) represent distinct forms of self-representation with different origins, sources of self-worth, and social motivations. They suggest the concept “we” primes social representations of the self that are more inclusive than that of the personal self-concept. In a preliminary investigation of the implications of different levels of the social self-concept, a set of three experiments were conducted to explore the effects of priming various “we” schemas on individual judgments and self-descriptions. In the priming task, participants read a descriptive paragraph with instructions to circle all the pronouns that appeared in the text, as part of a proofreading and word search task. After completing this word search task, participants were escorted to another room and asked to judge, as quickly as possible, whether the statements were similar or dissimilar to their own views by pressing a number key on the keyboard, ranging from 1 (very dissimilar) to 4 (very similar). They found that individuals primed with “we” would entail an expanded sense of self that would lower thresholds for agreement and assimilation.

In Chen and Li (2009), laboratory experiments with no interactions between subjects measure the effects of induced group identity on participant social preferences. They show that participants are more altruistic toward an in-group match. They evaluate different ways of creating group identity in the laboratory, to explore the formation of groups and to investigate the foundation of what group identity is. When participants are matched with an in-group member (as opposed to an out-group member) they show a 47% increase in charity concerns when they have a higher payoff and a 93% decrease in envy when they have a lower payoff. Also, participants are 19% more likely to reward an in-group match for good behavior, but 13% less likely to punish an in-group match for misbehavior. Participants are significantly more likely to choose social-welfare-maximizing actions when matched with an in-group member.

In Pilny et al. (2017), online experiments with no interactions between subjects are performed. To expand upon perspectives on the commons dilemma (e.g., do I contribute to the common resource or do I free ride), Pilny et al. (2017) developed an online experiment grounded on group decision making. They create manipulations based on three modalities of structure: dense versus sparse networks (domination), collective versus individual identity (signification), and social sanction versus non-social sanction (legitimation). The online experiments reveal that modalities of signification positively influence contribution rates on the commons dilemma, when participants were provided information meant to stimulate a CI. This is analogous to the findings of Charness et al. (2014); see Sect. 3.6. They mention how challenging it is for an online experiment to create CI, because the individual is sitting alone playing the game on a computer. In their experiments they try to stimulate CI by communicating three additional pieces of information regarding collective outcomes: (1) total collective score, rather than just an individual collective score, (2) collective rank compared to previous sessions, and (3) the score of the highest collective score from previous sessions. By including more collective, rather than individual, information, the user may come to behave more in a collective fashion and contribute to the public good.

In Ackland and O'Neil (2011), an online experiment using data collected from the websites of over 160 environmental activist organizations is developed. A model is presented where social movement actors exchange practical and symbolic resources in the guise of website text content and hyperlinks, as part of a process of online CI formation. The hyperlink and online frame networks are compared on three measures of centralization: degree, betweenness and closeness.

Wendt (1994) argues that international cooperation among independent states can be fostered through CI. He describes different mechanisms that may lead to CI, and takes examples from past events or general ideas. For example, he states that trade relations among states can foster CI through the emergence of the feeling of a common fate, but there are no experiments nor historical observations. It focuses more directly on identities and interests as the dependent variable and investigates whether, how, and why identities change.

Greenhill (2008) evaluates self and recognition theories—recognition theory states that an individual or group places recognition of itself by others as a very high-priority goal—to determine whether these two ideas can combine to produce CI. The reasoning is that as the self acknowledges others, and this process is replicated by all participants, a collective identity is formed. However, social identity theory-based experiments do not support this line of reasoning. This work is more akin to a meta-study, summarizing existing results.

Peek (2005) studies CI generation among Muslims in the USA. It is an empirical study of the formation of religious CI among 127 subjects, using focus groups, individual interviews, and participant observations. She presents three consecutive steps to form CI: religion as an ascribed identity; religion as chosen identity; and finally religion as declared identity.

Choup (2008) studies the relationships among specific (poor) constituent groups and governments, and how these groups use their shared (collective) identity to position themselves. She also uses observations (of group meetings) and interviews of group leaders to produce a model of CI formation and its effect on collective action.

Swanson (2015) uses small groups of music students (sizes of 2–5 students) to study the formation of CI. Again, as with several previous works, surveys, interviews, and observational studies are used to document the CI formation processes as students work together.

Brunsdon (2017) examines South Africa and the fracturing of the nation among different societal groups. Factors contributing to the lack of a national CI (e.g., a lack of trust among sub-groups and misunderstandings) are also discussed. Finally, the article posits that one way to heal these divisions and form a national CI is through religious understanding.

A final work in CI formation is experiments with interactions among subjects performed in Charness et al. (2014). This work, in a general way, motivated our anagram game experiments (although there are many differences between our work and that in Charness et al. (2014)). Consequently, we address this work separately below.

Dismissing for the moment this last reference, it is clear that none of the above works on CI formation are like ours. In contrast, our work uses controlled online laboratory experiments to produce CI through priming groups of subjects using a cooperative anagram game.

3.4 Collective identity-based experiments: implications of CI

The effects of religious (group) identity on individual behavior is studied in Benjamin et al. (2016). Subjects (self-identified as Protestant, Jew, Catholic, or agnostic/atheist) were primed or not primed with respect to religion. Priming consisted of having players unscramble a set of words that form a sentence, and that sentence has religious content. The unprimed subjects unscrambled words to form a sentence with no religious content. The purpose of priming is to make salient the religious identities of players, if they exist. Subjects then played a number of games, including public goods games, risk aversion games, discount rate elicitation games (i.e., delayed gratification games), among others. In a public goods game, players are given some amount of money. They have the option of contributing a portion of their money to the group. The pooled money that is contributed to the group by all members is then typically multiplied by some factor and redistributed to the players. Hence, there may be some incentive to contribute to the group. There are several interesting results. Among them is that religious identity salience (i.e., priming) produced an increase in Protestant subjects' contributions to Public Goods Games (PGG), while it generated a decrease in Catholic subjects' contributions.

In related experimental economics work using Indian caste and other nonreligious identities, Eckel and

Grossman (2005), Hoff and Pandey (2006, 2014), Charness et al. (2007), Chen and Chen (2011), Cohn et al. (2014), Chen et al. (2014), Cohn et al. (2015) find that group identity effects on behavior strengthen with the salience of group membership. Chen and Yeh (2014) manipulate the norms (expressed by legal rulings) that subjects are exposed to and study how these norms affect their self-identification.

The following works study the implications of *identity fusion*, where individuals may feel fused with (i.e., strongly connected to) a group (Swann et al. 2009, 2010a, b; Gomez et al. 2011a; Swann et al. 2014; Gomez et al. 2011b). We interpret identity fusion to be synonymous with, or very similar to, CI.

In Swann et al. (2009), the authors use online experiments to test the notion that fusion represents a distinctive form of allegiance to groups. They propose that when people become fused with a group, their personal and social identities become functionally equivalent. To measure identity fusion they used a modified version of a fusion scale developed by Schubert and Otten (2002). They prove that activating either personal or social identities of people who were fused with their group increased the extent to which they were willing to fight or even die for the group. Thus, even when people become deeply aligned with a group, their personal identities remain potent.

In Swann et al. (2010a), using an intergroup version of the trolley problem, the authors explored participants' willingness to sacrifice their lives for their group. Studies showed that nonfused participants expressed reluctance to sacrifice themselves, and identification with the group predicted nothing. To measure identity fusion they used the same scale as in Swann et al. (2009).

In Swann et al. (2010b), they assume that autonomic arousal will increase agency (i.e., the capacity to initiate and control intentional behavior) for fused and nonfused persons. In four experiments, increasing autonomic arousal through physical exercise elevated heart rates among all participants. Fused participants, however, uniquely responded to arousal by translating elevated agency into endorsement of pro-group activity. To measure identity fusion they used the same scale as in Swann et al. (2009).

In Gomez et al. (2011b), online experiments showed that when people are ostracized (i.e., rejected and excluded) by either an out-group or an in-group, they may either withdraw or engage in compensatory activities designed to reaffirm their social identity as a group member. The authors proposed that individual differences in identity fusion (an index of familial orientation toward the group) would moderate the tendency for people to display such compensatory activity. Four experiments showed that irrevocable ostracism increased endorsement of extreme, pro-group actions (fighting and dying for the in-group) among fused persons but not among nonfused persons. To measure identity fusion they used the same scale as in Swann et al. (2009).

In Gomez et al. (2011a), the authors determine what fusion is and the mediating mechanisms that lead fused individuals to make extraordinary sacrifices for their group. For measure of group identification, they proposed a seven-item verbal scale with greater fidelity than the earlier pictorial measure of identity fusion from Swann et al. (2009).

In Swann et al. (2014), online experiments explored the cognitive and emotional mechanisms that underlie the endorsement of self-sacrifice. Using participants responses to moral dilemmas, they found that only those who were strongly fused with the group preferentially endorsed self-sacrifice. Identity fusion was measured using the seven-item verbal fusion scale from Gomez et al. (2011a).

3.5 Measurement of CI

Researchers measure or declare the existence of CI in different ways. This is in part because there are many definitions for, and types of, CI (see Sect. 1.1).

Many references on CI formation (Wendt 1994; Peek 2005; Choup 2008; Greenhill 2008; Swanson 2015), presented in Sect. 3.3, pronounce that CI has been formed based on expert evaluation of textual comments of participants, survey responses, and interviews. These are subjective approaches for determining the existence of CI. They require an expert to interpret the data, and multiple experts may arrive at different conclusions.

In PGGs (Ledyard 1994), players are given some amount of money. They have the option of contributing a portion of their money to the group. The pooled money that is contributed to the group by all members is then typically multiplied by some factor and redistributed to the players. Hence, there may be some incentive to contribute to the group. Charness et al. (2014), Chen and Li (2009) use PGG contributions as a proxy for CI. In Charness et al. (2014), the percentage of a persons money that they contribute to the group is taken as their identification with the group: those with greater group identity contribute more of their money to the team.

In Swann et al. (2009) a modified version of a fusion scale developed by Schubert and Otten (2002) is proposed. To capture fusion in a manner that emphasized perceived overlap and nothing else, participants choose from five pictures which best represented the way they perceived their relationship with the group. Each figure in the scale shows two circles of different sizes. The small circle represents "the self", the big circle represents "the group". When participants need to choose from the scale, five figures with symmetrical degrees of overlap (0%, 25%, 50%, 75%, and 100%) are presented. For example, the first figure shows the two circles not intercepting, the second figure show a 25% interception and the fifth figure show a 100% interception with the small circle. To measure identity fusion, the following works use this scale (Swann et al. 2009, 2010a, b; Gomez et al. 2011a).

In Gomez et al. (2011a), a seven-item verbal scale is proposed to obtain greater fidelity in the measurement of identity fusion, compared to the pictorial measure from Swann et al. (2009). The levels in the verbal scale are represented with the following sentences (1) "I am one with my group", (2) "I feel immersed in my group", (3) "I have a deep emotional bond with my group", (4) "My group is me", (5) "I'll do for my group more than any of the other group members would do", (6) "I am strong because of my group", (7) "I make my group strong". Swann et al. (2014) use this scale to measure identity fusion.

In Jiménez et al. (2016) the DIFI is introduced to combine the simplicity of the single pictorial item (Swann et al. 2009) with the higher predictive fidelity of the verbal scale (Gomez et al. 2011a). The scales presented in Swann et al. (2009), Gomez et al. (2011a) are not dynamic. In Jiménez et al. (2016) the DIFI is defined as a continuous measure of identity fusion, introducing a dynamic behavior for web-based questionnaires. The DIFI shows a figure formed by two circles of different sizes in the screen of the computer. The small circle represents "the self", and the big circle represents "the team". The player can move the small circle by clicking and dragging with the mouse to measure the degree to which the player feels part of the team.

3.6 Combined group anagram and CI experiments

A group anagram game entails cooperation in requesting and receiving letters, with the goal of forming more words with additional letters received from teammates. The only *face-to-face* cooperative team-play of an anagram game is reported in Charness et al. (2014). Their goal, like ours, is to foster CI among teammates. While this motivated our experiment, there are several differences in procedures and context. Major differences include (1) the game setup: we used larger fixed team compositions, while in Charness et al. (2014), the four-person team composition varied in time (by people voting themselves and others onto and off of teams); (2) in Charness et al. (2014), games were played face-to-face among participants in the same room cooperatively manipulating Scrabble-like tiles on a table, while we used remote players interacting in a game through a web application; and (3) in Charness et al. (2014), they measure CI with the proxy of PGG contributions, while we use DIFI score.

3.7 Modeling of time sequences of actions

We review modeling of time sequences because our ABMs are essentially in this class of models.

Many complex action sequences from human behavior are being collected from different environments, like sensors (Guralnik and Haigh 2002; Aipperspach et al. 2006; Bergmann et al. 2014; Tanaka et al. 2018) or computer-based applications (Kinnebrew et al. 2013; Chierichetti et al. 2014; Kurashima et al. 2018). Sequence mining techniques to model and predict human behavior in the real world can be used in different types of applications to improve a person's life (e.g., mobile health (Kurashima et al. 2018), education patterns (Kinnebrew et al. 2013), smart-home optimization (Guralnik and Haigh 2002; Aipperspach et al. 2006)).

Sequence analysis is an important task to understand human behavior (Abbott 1995). The sequential pattern mining problem was first introduced by Agrawal and Srikant (1995), where the main focus is on the patterns present in the sequential order of different transactions. But the complexity of human behavior with time-varying, interdependent and periodic action sequences (Kurashima et al. 2018) makes accurate analysis and predictions a challenging task.

Kurashima et al. (2018) use activity data from logging applications to model the task of predicting future user actions and their timing through a mixture of Gaussian intensities. The model captures short-term and long-term periodic interdependencies between actions through Hawkes process-based self-excitations (Hawkes 1971). Accurate recommendations could improve a person's health through the personalization of these applications. In Kinnebrew et al. (2013), a combination of sequence mining techniques uses data from computer-based learning environments to model students learning behavior patterns. Guralnik and Haigh (2002) use sequential pattern learning to model an agent-based system to aid elderly people in living longer in their homes. Aipperspach et al. (2006) use pervasive home sensors, like motion sensors, door close sensors, and floor pressure pads, to model and predict discrete human actions with smoothed *n*-grams.

We are not modeling specific actions in our work. Rather, we are modeling the sequencing of actions during an anagram game. The above works use primarily data from in situ environments, while our data come from human subjects experiments.

3.8 Evaluation of model predictions

Predictive models can have many forms. For example, simple classifier algorithms try to predict discrete class labels. Another technique used in predictive modeling is regression analysis, which tries to predict the mean value of a quantitative response variable. Also, the factor analysis approach, tries to predict the distribution of a set of correlated quantitative variables (i.e., predicts the values of some variables from knowing the values of others). The evaluation of prediction models can be developed using a variety of different methods and metrics. For classification, the usual measure of error is the fraction of cases mis-classified, called the mis-classification rate or the error rate. For linear regression, the measure of accuracy is R^2 and the measure of error is the sum of squared errors or $1 - R^2$. For the method of factor analysis, when a model predicts a whole distribution, the negative log-likelihood is the usual measure of error, but sometimes a direct measure of the distance between the predicted and the observed distribution is used (Hand et al. 2001).

In this work, we are primarily concerned with using well-known measures to characterize the difference between two statistical distributions. In our work, one distribution is generated from experimental data, and one distribution is generated from predictions of models from Sect. 5. Gibbs and Su (2002) list ten metrics on probability measures: (1) Discrepancy, (2) Hellinger distance, (3) Relative entropy (or Kullback–Leibler divergence), (4) Kolmogorov (or Uniform) metric, (5) Lévy metric, (6) Prokhorov metric, (7) Separation distance, (8) Total variation distance, (9) Wasserstein (or Kantorovich) metric, and (10) χ^2 distance.

It is clear that there are many measures for comparing two probability distributions, and different ones are used in different settings. For our needs, we have chosen to use KL divergence (also called relative entropy). The KL divergence was introduced by Solomon Kullback and Richard Leibler in 1951 as the directed divergence between two distributions (Kullback and Leibler 1951).

The most important measure in information theory is called entropy and measures the uncertainty associated with a random variable. The entropy of a random variable X denoted $H(X)$ is a lower bound on the average length of the shortest description of the random variable (Cover and Thomas 1991). The concept of information entropy was introduced by Shannon (1948). The Shannon entropy, defined in Shannon (1948), measures how close a random variable is to being uniformly distributed. Shannon entropy estimates the average minimum number of bits needed to encode a string of symbols based on an alphabet size, and the frequency of the symbols is calculated using the following formula $H(X) = -\sum_{x \in X} P(x) \log P(x)$. The KL divergence measures the discrepancy between two probability distributions, and from which Shannon entropy can be constructed. For discrete probability distributions P and Q defined on the same probability space, the KL divergence between P and Q is defined to be

$$D_{\mathrm{KL}}(P||Q) = -\sum_{x \in X} P(x) \log\left(\frac{Q(x)}{P(x)}\right).$$

In the simple case, a KL divergence of 0 indicates that the two distributions in question are identical. The KL divergence is not symmetric. The evaluations of our models are described in Sect. 6.

Fig. 5 Steps for the overall online game include: recruitment of players from Amazon Mechanical Turk (AMT), directions for the use of the platform, DIFI1 score procedure, anagram game, and DIFI2 score procedure. This figure directly maps onto the experimental components in Fig. 1

3.9 Abduction and abductive loop

Works on constructive procedures for implementing abductive analyses include Haig (2005), Timmermans and Tavory (2012). We extend those works for abductive looping by making modeling a first-class process, and by adding the task of determining what to do in the next iteration. In addition to the applications cited in the Introduction, abduction was used to understand emergency room personnel's efforts to save injured people in terms of "social viability" (Timmermans 1999). Perhaps the work closest to ours is Singla and Mooney (2011) in that they develop models and make predictions based on data. However, their data are either artificially generated or address isolated individuals, and they use abduction rather than abductive iterations. Several additional works are provided in Sect. 1.3.

4 Experiments

In Sect. 4.1, we provide a description of the experiment and overview the web application (web app) software system for running games. An experiment consists of an anagram game and two executions of the dynamic identity fusion index (DIFI) procedure. We present analyses of the experimental data that illustrate how players interact in the anagram games in Sect. B.

4.1 Experiment description

The elements of an experiment, as specified in Fig. 5, are:

1. Players are recruited from Amazon Mechanical Turk (AMT), to play our anagram game.
2. Players receive directions on how to use the platform, including a description of the game and how to play it, and information about remuneration at the end of the game.
3. Players play the Dynamic Identity Fusion Index (DIFI) 1, DIFI1, procedure individually.

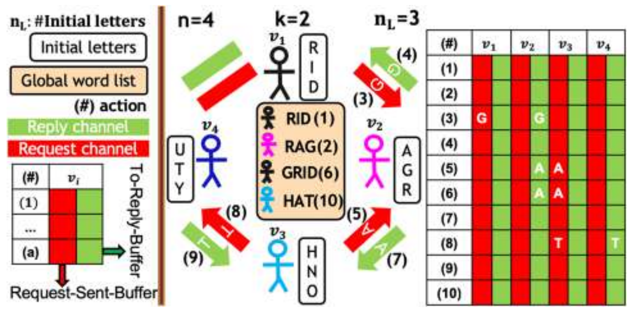

Fig. 6 Group anagram game configuration with a $k = 2$ regular graph on $n = 4$ players (v_1, v_2, v_3, v_4) with number of initial letters $n_L = 3$ assigned to each player, as shown in the boxes next to the players. Requests for letters and replies are sent across the channel links (red to request letters, green to reply with letter). Request-Sent-Buffer keeps track of player v_i's letter requests. To-Reply-Buffer contains letter requests from other players to v_i. Numbers (#) denote actions by players during a game, and a table illustrating the time sequencing of the actions appears at right. Table 3 shows a detailed description of these actions (color figure online)

Table 3 Action table detailing the sequences of actions by all players during the group anagram game example from Fig. 6. The first column defines the number of the sequence of actions during the game, appearing in Fig. 6. For this example, the duration of the game is 10 actions. The second column shows the player initiating the action. The third column shows the name of the action. The fourth column provides a description of the action

(#)	Player	Action	Description
(1)	v_1	Form word	v_1 forms word "RID"
(2)	v_2	Form word	v_2 forms word "RAG"
(3)	v_1	Request letter	v_1 requests v_2 for letter "G"
(4)	v_2	Reply letter	v_2 replies v_1 with letter "G"
(5)	v_3	Request letter	v_3 requests v_2 for letter "A"
(6)	v_1	Form word	v_1 forms word "GRID"
(7)	v_2	Reply letter	v_2 replies v_3 with letter "A"
(8)	v_3	Request letter	v_3 requests v_4 for letter "T"
(9)	v_4	Reply letter	v_4 replies v_3 with letter "T"
(10)	v_3	Form word	v_3 forms word "HAT"

4. Players play the anagram game in a cooperative group setting.
5. Players play the DIFI2 procedure individually.

The terms DIFI1 and DIFI2 are used to indicate the first and second uses of the DIFI procedure (Fig. 5). The two DIFI procedures are the same.

4.1.1 Group anagram game description

The group anagram game is a word construction game, where n players cooperate in sharing letters to form and submit words of length ≥ 3 letters. Communication channels between pairs of agents mean that they can request and share letters with each other. An edge between nodes (players) v_i and v_j means that v_i and v_j can share letters with each other; v_i and v_j are neighbors. We use random regular graphs of degree k on the n players so that everyone has the same number of neighbors. A random k-regular graph on n players is an undirected graph such that each player has k neighbors assigned uniformly at random. Over all abductive loops, experiments are run in groups with nominal values of $10 \le n \le 20$ and with regular degrees $2 \le k \le 8$. We describe our motivation for using random regular graphs in Sect. 4.1.2.

An example game configuration and system states are provided in Fig. 6. The game configuration can be represented as a graph $G(V, E)$ where V is the set of nodes that represent players and E is the set of edges that are communication channels between pairs of nodes. Red channels are for letter request and green channels are for letter replies. The number of players is $n = 4$ with players v_1, v_2, v_3, and v_4, the degree of each player is $k = 2$, and the number of initial letters per player is $n_L = 3$. The players have the following initial letters: $L^{\text{init}}_{v_1} = \{RID\}$, $L^{\text{init}}_{v_2} = \{AGR\}$, $L^{\text{init}}_{v_3} = \{HNO\}$, and $L^{\text{init}}_{v_1} = \{UTY\}$. Key (#) shows the sequence of actions by all the players during a game. In Fig. 6, the sequence of actions is detailed in Table 3, which narrates the actions. The To-Reply-Buffer and the Request-Sent-Buffer of Fig. 6 are buffers, per player, that contain outstanding requests-to-be-fulfilled and requests of letters, respectively. For example, in step (5) of Table 3, v_2 has a request from v_3 for the letter A. Therefore, v_3 has an entry A in its Request-Sent-Buffer and v_2 has an entry A in its To-Reply-Buffer. If/when v_2 fulfills that request [in the example this happens in step (7)], v_3's "received letters" will contain an A, A will be removed from v_3's Request-Sent-Buffer, and v_2's To-Reply-Buffer will become empty.

Team members earn money by forming as many words as possible. Players are told that the total team earnings e_t are split evenly; each player receives e_t/n, so that it is in their interests to assist their neighbors. Players must form words with at least three letters. A single letter can be used any number of times in a word, e.g., a player can form the word *TOT* if she has a T and an O among her current letters (own letters and those received from neighbors) because the T can be used twice. Moreover, players do not lose letters that they use. Hence, a player has infinite multiplicity of each letter they possess so that letters can be reused any number of times. This means that a player only has to request a letter (and receive it) one time. Therefore if a player forms *TOT*, she still possesses T and O with which to form more words. A player can only share their initial letters with her neighbors; letters received from neighbors cannot be shared with others. These rules were designed to foster word

Table 4 Description of anagram game configurations played with players recruited from Amazon Mechanical Turk (AMT). There were 47 games with 289 players, of which 34 games and 224 players were used in analysis and modeling. Others were scoping experiments

Degree, k	No. Players, n	No. Games
2	10	18
2	20	10
3	15	1
4	15	9
5	15	2
6	15	3
8	15	4

construction, to increase earnings potential, and to foster team cohesion.

A total of 105 players participated in 47 games. The anagram game is played for 5 min; Table 4 shows all the game configurations played.

We provide an overview of the web application (app) game platform that we built. The web app software platform consists of the oTree infrastructure (Chen et al. 2016) for recruiting players from Amazon Mechanical Turk (AMT) and interactions during the game; Django Channels for player interactivity; and JavaScript and HTML for generating the screens for a consent form, instructions, information, a survey, and game interactions. Experiments and data analyses *are part of* the abductive loop of Sects. 2 and 7 and Fig. 4. This game platform was constructed as part of our work.

A screen shot of one player's screen at one point in time is shown in Fig. 7. Each player is given $n_L = 3$ letters that she can use to form words and that she can share with others. She has an infinite supply of letters so that sharing letters does not inhibit her own use of letters. A player can also request letters from her neighbors and if the neighbors provide those letters, then she can use those letters in words, but she cannot pass on the received letters.

Initially, a player sees her n_L own letters and those of all of her neighbors, but has access only to her own letters. Over the 5-min anagram game duration, players can form words, request letters from their neighbors and reply to requests.

4.1.2 Choice of random regular networks for experiments

There may be many possible network configurations to explore in some games. These include Erdos–Renyi random graphs, small-world networks, and scale-free networks. We select random regular networks for the following reasons.

First, since all players in a game have the same degree, all players have the same number of neighbors and the same

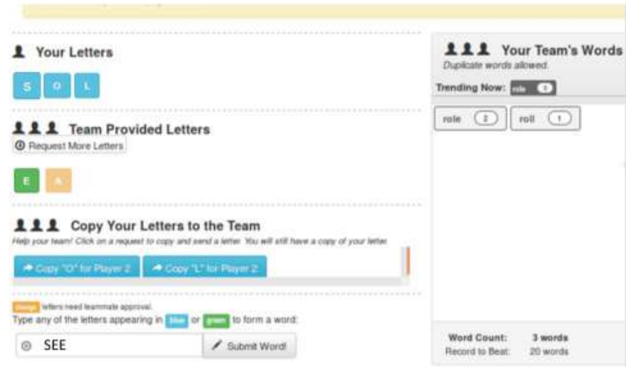

Fig. 7 The anagram game screen of the web app for one player. This player has own letters "S," "O," and "L" and has requested an "E" and "A" from neighbors. The "E" is green, so this player's request has been fulfilled and so "E" can be used any number of times in forming words. But the request for "A" is still outstanding so cannot be used in words. Below these letters, it shows that player 2 has requested "O" and "L" from this player; this player has to reply to these requests, if this player so chooses. Below that is a box where the player types and submits new words, like "SEE" (color figure online)

maximum number of interactions. No player can be viewed as special.

Second, our experiments use between five and twenty players per game, consistent with other studies of this kind. With these numbers of players, one cannot construct Erdos–Renyi, scale-free, and exponential decay graphs with the requisite degree distributions because there is an insufficient number of nodes. For example, to generate a scale-free or power law degree distribution, one needs about 5000 to 10,000 nodes in a graph. (One can demonstrate this by generating a scale-free network on 1000 nodes in NetworkX, and you will see that the degree distribution is quite "choppy"). Hence, instances of these latter classes of graphs cannot be generated with roughly five to twenty nodes. Critically, though, we can *model* these classes of network. That is, we can specify, say, a scale-free network on $n = 5000$ players and model each player based on their degree in the graph, as described below.

Third, our analyses and modeling in Sect. 5 are based on representations of individual agents. We account for the local structure of a node (agent) in a game by considering its closed neighborhood (i.e., its degree, plus one), which consists of itself and its immediate neighbors. Consequently, our models account for each player, and each player's closed neighborhood. This is, in particular, the Model M2 of Sect. 5.6. In this way, we can "connect" agents and their neighborhoods to form larger networks for modeling than the sizes of graphs we were able to test.

Fourth, by controlling degree in each experiment, we can conduct experiments at $k = 2$, 4, 6, and 8, and build models that interpolate for intermediate values of k (see Sect. 5). This enables a systematic approach to model building for individual agent behavior. Note that this systematic approach includes the ability to model systems where numbers of neighbors is *heterogeneous* across agents, as in Erdos–Renyi, scale-free, and other network classes, precisely because we generate individual agent models. For example, with our models, we can simulate a single game where players have arbitrarily assigned node degrees, and the model assigned to each player is based on that node's degree.

Fifth, eight neighbors is used as the upper end of the degree range. With each player sharing three letters, a player with degree eight has access to 27 letters (including their own three letters), and hence a high probability of having access to almost all letters of the alphabet. In our models, as a first approximation, if a player has degree greater than eight, we use the $k = 8$ model for that agent.

Sixth, using a uniform degree for all players in a game enables us to generate multiple (replicate) sets of data across multiple individuals within one game. The importance of this replication in degree values is demonstrated in Sect. 6. In that section, we conduct rigorous analyses to identify the number of experimental observations required to drive down errors between model predictions and experimental data. An important result is that had we experimented with different network structures and hence a greater number of values for k (node degree), then the errors in our models would have been greater because of the "dispersion" (i.e., wider range of degree k) in experimental conditions that would have resulted. For example, if we had created one network with node degrees of 2, 3, 5, and 7 (among others), then we would have generated data over more k values, which would have dispersed the data over more k, and led to fewer data points at each k. This would result in larger errors in model predictions.

Seventh and finally, there is a real-time, pragmatic issue. We recruit game players through AMT. But not all players who promise to show up for the game do so at the appointed time and date. By using even regular degree values, we can produce valid random regular graphs as long as $n > k$. With graph structures such as scale-free (were it even possible), if players do not show up, one is faced with the ambiguous decision about which nodes in the communication graph to delete.

For all of these reasons, we chose to use random regular graphs in our study: it enables us to study behavior based on player degree and to construct and evaluate models. Additional graph structures (i.e., heterogeneous degree graphs) may be studied as part of future work.

4.1.3 DIFI description

The DIFI procedure precedes and follows the anagram game. Each player executes individually the DIFI procedure (Swann et al. 2010b), to measure the degree to which a player feels part of a team (i.e., associates their identity with that of a team). Each player does this individually by moving a circle in a browser, relative to a fixed team circle. The DIFI score is in the range [− 100, 125], with a score < 0 representing no overlap of circles, and therefore indicating no CI; = 0 representing the circles just touching; and > 0 indicating overlap of the two circles and hence formation of some level of CI. See Fig. 8. There are screens in the web app that also step each player through the steps in the DIFI game/procedure.

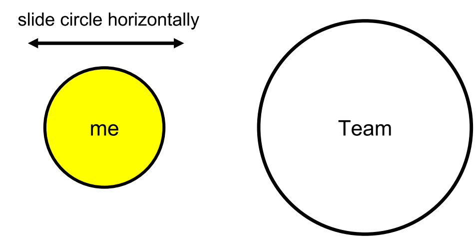

Fig. 8 DIFI game where player v_i moves the smaller circle, representing the player, either over (partially), or away from, the bigger circle that represents the team. The team circle is stationary. The distance δ between centroids of circles is measured. The distance is such that $\delta = 0$ corresponds to the small and large circles just touching; $\delta < 0$ means that the two circles are disjoint; and $\delta > 0$ means the two circles overlap. The distance δ is transformed into a DIFI value. The range in DIFI value is: $-100 \le \delta \le 125$. The DIFI score is a proxy for CI. This is an individual player game, played in isolation; results are never shared among players, so there is no concern over reprisals by other players, to foster honest actions

4.2 Experimental data

Experimental data are provided in "Appendix B".

5 Agent-based models (ABMs) of the group anagram game and modeling results

We present three progressively more sophisticated ABMs of the anagram game that are used in the abductive loop analyses to follow in Sect. 7. All models were developed *as part of* the abductive loop process, but are presented here to emphasize their construction and evaluation, and to obviate the need for a large digression for the models in the description of the AL process in Sect. 7. Each model represents the behavior of one player or agent. The models are data-driven, and hence *inductive inference* is used with data in three ways: to inform model structure, model parameters, and to compute parameter values. Some of the figures appear in "Appendix C".

In all models, we represent the set V of players and the set E of their communication channels (edges) as an undirected graph $G(V, E)$. The game is modeled as a discrete-time

Table 5 Actions of players in the model. The set A of actions is $A = \{a_1, a_2, a_3, a_4\}$

Item	Variable	Name	Description
1	a_1	Idling	Thinking
2	a_2	Reply	Replying to a neighbor with a requested letter
3	a_3	Request	Requesting a letter from a neighbor
4	a_4	Words	Forming and submitting a word

stochastic process, where at each time step, a player performs one of the actions from the action set A, consisting of: (1) a_1: idling (i.e., thinking); (2) a_2: replying to a neighbor with a requested letter, (3) a_3: requesting a letter from a neighbor, and (4) a_4: forming and submitting a word. Table 5 shows the actions.

5.1 Discrete-time stochastic process

In all ABMs, actions are taken at integer numbers of seconds; that is, simulations of interacting agents take place as time advances in discrete 1-second increments from 0 to 300. This time increment is based on the experimental data where no player takes two or more actions in one second.

We chose ABMs for their generative properties, fine granularity, and ability to model temporal effects. These enable us to more readily quantify "what if" scenarios (counter factuals) as part of parametric studies and sensitivity analyses. Also, ABM maps well onto the actual experiments: players have connections in a network arrangement and they interact through their edges, taking actions at discrete times as in Fig. 6.

The choice of discrete time or discrete event simulation arises. If we selected discrete event simulations, then we would also have to predict the time at which the next action for a player takes place (at some Δt into the future). However, with discrete time, we know we are always predicting for the next time unit (here, one second). We also used a multinomial logistic regression model; other approaches could have been employed.

5.2 KL-divergence

To measure the performance of our models, we use *Kullback–Leibler divergence* between our model prediction on x and the experimental observation of x, *throughout this manuscript*. That is, we are comparing distributions of data: distributions of experimental data against distributions of model predictions. Most relevant for our work is Boltzmann's (Bach 1990) concept of generalized entropy, where the entropy of a physical system is a measure of disorder related to it. Kullback and Leibler (1951) derived an information measure, now referred to as the KL divergence, the negative of Boltzmann's entropy. The motivation for Kullback and Leibler's work was to provide a rigorous definition of information. The Kullback–Leibler distance can be conceptualized as a directed distance between two models, say a and b (Kullback 1959). This is a measure of discrepancy. It is not a simple distance because the measure from a to b is not the same as the measure from b to a. It is a directed, or oriented, distance. The KL divergence $D_{\mathrm{KL}}(a, b)$ is always positive, except when the two distributions a and b are identical (i.e., $D_{\mathrm{KL}}(a, b) = 0$ if and only if $a(x) = b(x)$ everywhere). Entropy is zero if there is unit probability at a single point. If the distribution is widely dispersed over a large number of individually small probabilities, then the entropy is high (e.g., $D_{\mathrm{KL}} > 1$).

5.3 Overview of the three agent-based models

ABM M0 is a baseline model, where each player makes a probabilistic transition from action $a_i \in A$ to action $a_j \in A$. The transition matrix is time invariant and is the same for all players. Data from the experiments is used to infer the model parameters using a ring topology (degree of each node is 2) of player connectivity within an anagram game. Model M1 is similar to M0 but with the crucial difference that the transition matrix is time variant. Model M2 is similar to M1 but now instead of a ring topology, we used other topologies and infer model parameters (degree from 2 to 8). Models M0, M1 and M2 predict the actions of A for a player but are generic in that letter request a_3, letter reply a_2, and submit word a_4 are not associated with particular letters. For example, if the player action is a_4, then the model assumes that the player can form a word. Table 6 shows a description of the three progressively sophisticated models.

Models M0, M1, and M2 are presented in Sects. 5.4, 5.5, and 5.6, respectively. In each of these subsections, model development and results are provided.

Throughout, we use k to denote the number of neighbors (degree) of an agent $v \in V$. Also, we evaluate five variables and their distributions, across all players in a set of games, in comparing models and experiments: $x = (x_1, x_2, x_3, x_4, x_5)$, where x_1 is the number of letter replies received (*RplR*); x_2 is the number of replies sent (*RplS*); x_3 is the number of letter requests received (*RqsR*); x_4 is the number of requests sent (*RqsS*); and x_5 is the number of words formed (*Wrds*). Table 7 summarizes these variables.

In the results sections for each model, simulations are performed using ABMs that implement each of the described models. These simulations produce, for each player, time histories of the actions in Table 7. One hundred simulations are run and results are averaged across these simulations, i.e., are averaged across all players in each simulation. These data are post-processed to generate distributions of

Table 6 Progressively sophisticated models of the group anagram game are developed in this work. Models were constructed in order M0, M1, and M2. The incremental improvements in models are given in columns two and three, in terms of transition probabilities and degrees of players in the games

Model	Transition probabilities	Degree k
M0	Fixed	2
M1	Temporal	2
M2	Temporal	2, 4, 6, 8

Table 7 Variables that are measured in experiments for each player, and predicted with models for each agent, where vector $x = (x_1, x_2, x_3, x_4, x_5)$. All x_i, $1 \leq i \leq 5$, are time dependent

Item	Variable	Name	Description
1	x_1	RplR	Number of replies received
2	x_2	RplS	Number of replies sent
3	x_3	RqsR	Number of requests received
4	x_4	RqsS	Number of requests sent
5	x_5	Wrds	Number of words formed

the variables in Table 7. These distributions from ABM predictions are compared against corresponding distributions generated from experiments.

Note that fixing $n = 10$ in all simulations does not introduce errors because the distributions that we use are density distributions, not counts. Thus, the number of players is normalized out of all comparisons of distributions of experimental data and model predictions.

Table 8 shows the structure of comparisons of results for each of the models M0, M1, and M2. First, comparisons are made between distributions of experimental results and model predictions, for each x_i of Table 7, at the end of a game (i.e., over all 5 min of an anagram game). Then, these data are broken down into 1-min intervals to assess temporally the distributions of data and predictions. Next, we compute KL-divergence values that provide a scalar representing how well the model predictions of the distributions of x_i compare with those of the experimental data. From Sect. 5.2, $D_{KL} = 0$ means the model distribution agrees very well with the corresponding experimental distribution. As D_{KL} increases from zero, model predictions worsen. Table 8 denotes that these comparisons are performed over all 5 min of the anagram game (number 3), corresponding to the end of the group anagram game, and for each 1-min interval over the game (number 4) of Table 8. Finally, we compare these sets of computed D_{KL} across all x_i of Table 7. The reason for the temporal breakdown is to

Table 8 Summary of the model comparison plots for each of the models M0, M1, and M2. For each model, we collect the data into the five groups shown. See the text for details and justification. The fifth column indicates the time period, in minutes, over which experimental data are compared to model predictions. These plots facilitate comparisons of final outcomes and temporal performance of models

No.	Method	Plot	Variables, player actions	Time
1	Comparisons of distributions at end of game	(a)	x_1	0–5
		(b)	x_2	0–5
		(c)	x_3	0–5
		(d)	x_4	0–5
		(e)	x_5	0–5
2	Temporal comparisons of distributions	(a)	x_1, x_2, x_3, x_4, x_5	0–1
		(b)	x_1, x_2, x_3, x_4, x_5	1–2
		(c)	x_1, x_2, x_3, x_4, x_5	2–3
		(d)	x_1, x_2, x_3, x_4, x_5	3–4
		(e)	x_1, x_2, x_3, x_4, x_5	4–5
3	Comparisons of KL-divergence distributions at end of game		x_1, x_2, x_3, x_4, x_5	0–5
4	Temporal comparisons of KL-divergence distributions	(a)	x_1, x_2, x_3, x_4, x_5	0-1
		(b)	x_1, x_2, x_3, x_4, x_5	1–2
		(c)	x_1, x_2, x_3, x_4, x_5	2–3
		(d)	x_1, x_2, x_3, x_4, x_5	3–4
		(e)	x_1, x_2, x_3, x_4, x_5	4–5
5	Comparisons of KL-divergence distributions combining all variables		x_1, x_2, x_3, x_4, x_5	0–5

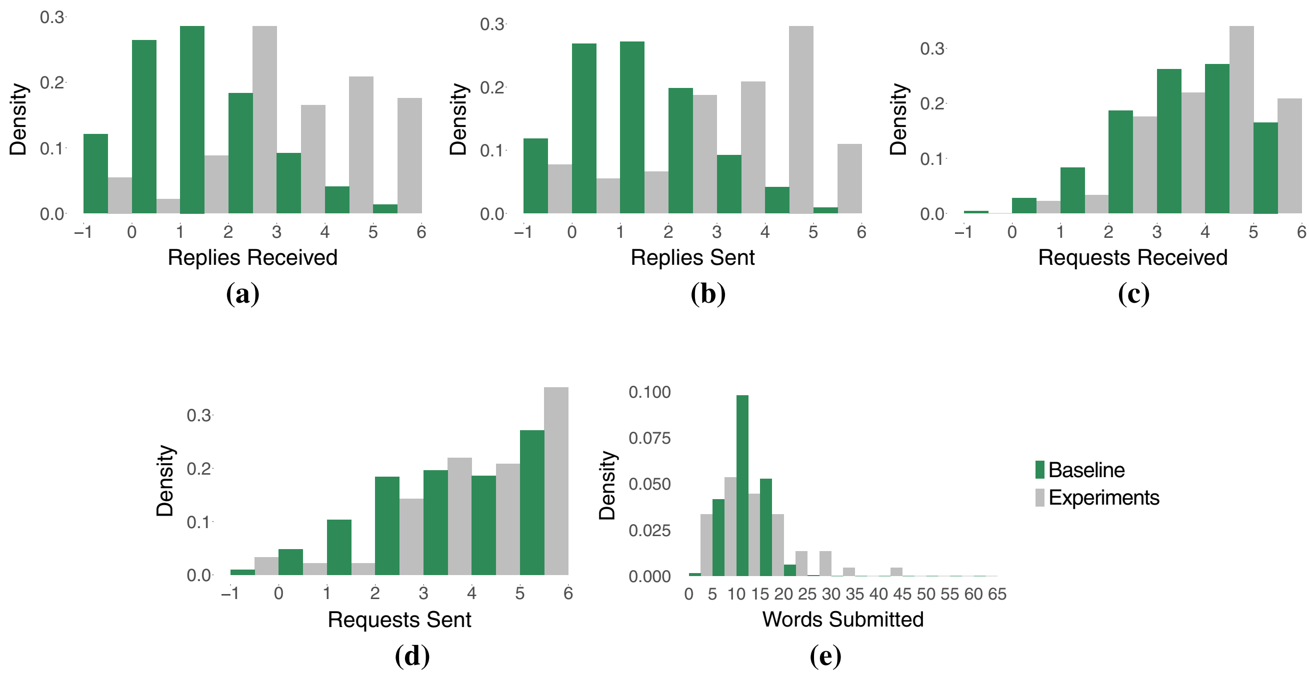

Fig. 9 ABM baseline model M0 predictions of the $k = 2$ experiments (in green) and experimental data (in gray), over the entire 5-min group anagram game. The probability density function is show for **a** distribution of replies received, **b** distribution of replies sent, **c** distribution of requests received, **d** distribution of requests sent, and **e** distribution of words formed, each at the end of the 5-min anagram game (gray bars are experimental data) for all $k = 2$ experiments. The Baseline model M0 predictions are from 100 simulations of a $n = 10$ player game. It is clear from visual inspection that model M0 predictions are in better agreement with the experiment data for the requests received and requests sent variables. We make this comparison more precise using KL divergence below in Fig. 10 (color figure online)

examine model predictions over time. Temporal comparisons are hidden in numbers 1, 3 and 5 of Table 8, which examine aggregated data.

5.4 Baseline agent-based model M0

5.4.1 ABM M0 development

The goal is to accurately quantify the transition probability from one action $a(t) = a_i$ at time t to the next action $a(t+1) = a_j$ for each agent $v \in V$, $i, j \in [1..4]$ and $a(t) \in A$. For clarity, we use i and j to represent the actions a_i and a_j. Agent v executes a stochastic process driven by transition probability matrix $\Pi = (\pi_{ij})_{m \times m}$, where $m = |A|$ (here, $m = 4$) and

$$\pi_{ij} = Pr(a(t+1) = j | a(t) = i) \quad \text{with} \quad \sum_{j=1}^{m} \pi_{ij} = 1. \tag{1}$$

The transition matrix Π is formed from the data by using successive pairs of actions of players in experiments so that the 16 values of π_{ij} in Eq. (1) are *constant*, i.e., time invariant. The matrix in Eq. (2) shows the transition probabilities for Model M0 (the baseline model) generated from experiment data with $n = 10, k = 2$. For example, given that the action of a player v_i at time t is a_2 (replying to a letter request), the probability that v_i's next action, at time $(t+1)$, is a_1 (thinking) is 0.84.

$$\Pi = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{array} \begin{array}{c} \begin{array}{cccc} a_1 & a_2 & a_3 & a_4 \end{array} \\ \left\Vert \begin{array}{cccc} 0.93 & 0.01 & 0.02 & 0.04 \\ 0.84 & 0.16 & 0 & 0 \\ 0.98 & 0.01 & 0.01 & 0 \\ 0.93 & 0.01 & 0 & 0.06 \end{array} \right\Vert \end{array} \tag{2}$$

5.4.2 ABM M0 (baseline) results

We address all of the results in Table 8 for model M0.

Comparisons of distributions between model and experiments for individual variables at the end of the anagram game Figure 9 shows the ABM M0 predictions of the $k = 2$ experiments. Figure 9a shows the distribution of replies received, Fig. 9b shows the distribution of replies sent, Fig. 9c shows the distribution of requests received, Fig. 9d shows the distribution of requests sent, and Fig. 9e shows the distribution of words formed, each at the end of the 5-min anagram game (gray bars) for all $k = 2$ experiments, compared to Baseline M0 predictions (green) for 100 simulations of an $n = 10$ player game. It is clear from visual inspection that model M0 predictions

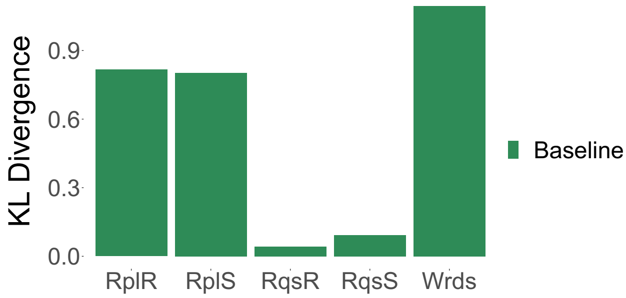

Fig. 10 KL-divergence values for the Baseline Model M0 across the five parameters of x: lower values are better. M0 does a better job predicting the number of Requests Received and Requests Sent. Analyses are based on the data of Fig. 9, over 5 min, at the end of a game

are in better agreement with the experimental data for the requests received and requests sent variables. We make this comparison more precise using KL divergence in Fig. 10.

Temporal comparisons of distributions between model and experiments for individual player actions Appendix B.1.5 in Cedeno (2019) shows the figures resulting from the temporal analysis by minute of distributions between Model M0 and experiments for $k = 2$. Each plot contains data over a time window for each variable of x from Table 7. Often, but not always, the largest discrepancies between the model predictions and experiments occur in the first minute of the game.

Comparisons of KL-divergence values between model and experiments for individual variables at the end of the anagram game. Figure 10 shows the KL-divergence values for the baseline M0 across the five parameters of x: lower values are better. M0 does a better job predicting the number of requests received and requests sent at the end of a game. These data span the entire 5-min game. That is, the request-related operations are better predicted than reply operations.

Temporal comparisons of KL-divergence values between model and experiments for individual player actions. Figure 26 shows the temporal KL-divergence values for the baseline M0 across the five parameters of x, at 1-min intervals: lower values are better. Each figure contains data over a time window: Fig. 26a shows the 0–1 min, Fig. 26b shows the 1–2 min, Fig. 26c shows the 2–3 min, Fig. 26d shows the 3–4 min, and Fig. 26e shows the 4–5 min results of the 5-min anagram game. These plots show that request-related predictions are better than reply-related predictions for the first 3 min, but are worse for the last 2 min, based on KL-divergence. Reply-related predictions are better in the second half of the 5-min anagram games, but Fig. 23 shows that in experiments, there are fewer replies in the second half of the games.

Comparisons of KL-divergence values between model and experiments for combining all variables. Figure 3.15 in Cedeno (2019) shows the distribution of KL-divergence values for comparing distributions of model output with corresponding distributions of experimental data for the anagram game. The model is the $(n = 10, k = 2)$ baseline. The data sets used in the comparison are $(n = 10, k = 2)$. There are 30 values in the distribution, with five values for variables x_i over the 5-min game, at the end of the game; and 25 values for the five variables of x over five intervals of 1 min duration. It shows that for model M0, some KL-divergence values are high (e.g., > 0.5), indicating poor agreement between model predictions and the experiment data. As we see in Figs. 10 and 26, M0 does not do a good job predicting the number of replies received, replies sent, and words formed.

5.5 Agent-based model M1

Model M1 is similar to M0 but with the important enhancement that the transition matrix Π is time variant.

5.5.1 ABM M1 development

To make Π [and its components π_{ij} in Eq. (1)] dynamic in time and account for history effects, four variables are introduced in Eq. (3): number $z_L(t)$ of letters that v has available to use (i.e., in hand) at t; number $z_W(t)$ of valid words that v has formed; size $z_B(t)$ of the buffer of letter requests that v has yet to reply to; and number $z_C(t)$ of consecutive time increments that v has taken the same action. See Table 9. Thus, letting $z = (1, z_L, z_W, z_B, z_C)_{g\times 1}$, we can model π_{ij} as a function of these covariates, among other variables.

We use a multinomial logistic regression to model π_{ij}—the probability of a player taking action a_j at time $t+1$, given that the player took action a_i at time t—as

$$\pi_{ij} = \frac{\exp(z' \boldsymbol{\beta}_j^{(i)})}{1 + \sum_{l \neq i} \exp(z' \boldsymbol{\beta}_l^{(i)})}, \quad (3)$$

where $\boldsymbol{\beta}_j^{(i)} = (\beta_{j1}^{(i)}, \dots, \beta_{j,g}^{(i)})'$, for $j \neq i$, and $\boldsymbol{\beta}_i^{(i)} = \mathbf{0}$, prime indicates vector transpose, and $\beta_{j,h}^{(i)}$ are the elements of $\beta_j^{(i)}$, with $1 \le h \le g$ being the index of the element of the z vector. For a given i, the parameter set can be expressed as

$$\boldsymbol{B}^{(i)} = \begin{pmatrix} \beta_{11}^{(i)} & \beta_{12}^{(i)} & \dots & \beta_{1,g}^{(i)} \\ \beta_{21}^{(i)} & \beta_{22}^{(i)} & \dots & \beta_{2,g}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{41}^{(i)} & \beta_{42}^{(i)} & \dots & \beta_{4,g}^{(i)} \end{pmatrix}, \quad (4)$$

with matrix entries $\beta_{j,h}^{(i)}$. Parameters in Eq. (4) are inferred from the $k = 2$ experimental data using the framework of maximum likelihood estimation for the multinomial distribution.

Table 9 The feature vector $z = (z_L(t), z_W(t), z_B(t), z_C(t))$ used in the models M1 and M2. These capture history effects in determining the next action of a player

Variable	Name	Description
z_B	Size of reply buffer	Number of current letter requests to which this player may reply. Captures the notion that the more letter requests that have not been replied to, the more likely v is to reply
z_L	Number of letter in hand	Number of unique letters in hand to form words. Captures the idea that the more letters v has in hand, the more likely the agent is to form words
z_W	Number of words formed	Number of words formed. Captures the notion that the more words that have been formed, the larger the vocabulary of the player
z_C	Number of consecutive actions	Number of consecutive time steps at which player takes the same action. Captures the notion that the more time v is idle (thinking), the more likely v will take some other action at the next timestep

5.5.2 Inductive inference

We address the three dimensions of inference stated above: (1) model structure; (2) model parameters; and (3) parameter values. First, the model structure is informed by the $k = 2$ data, by design, as described above. Second, the parameters identified in the feature vector z are described and justified in Table 9. In fact, we claim that identifying this feature vector has elements of art. Third, parameters in Eq. (4) are inferred from the $k = 2$ experimental data using the framework of maximum likelihood estimation for the multinomial distribution.

The reason to emphasize inductive inference is because this is an integral part of the abductive looping process, and of abduction itself: ***the data drive the model and theory development and hypothesis identification****, and not the other way around.*

5.5.3 ABM M1 results

Results for Model M1 are provided according to Table 8 as was done for Model M0. In many cases, we compare KL-divergence values for M0 and M1 to show improvements in performance. These results, like those for model M0, are compared against the $k = 2$ data in Table 4.

Comparisons of distributions between models and experiments for individual variables at the end of the anagram game Figure 11 shows M0 and M1 model predictions and experimental data distributions for all variables in Table 7. These data are over all 5 min of the anagram game for all $k = 2$ experiments. Model predictions are averages over 100 simulations with $n = 10$ players. Figure 11a shows the distributions of replies received, Fig. 11b shows the distributions of replies sent, Fig. 11c shows the distributions of requests received, Fig. 11d shows the distributions of requests sent, and Fig. 11e shows the distributions of words formed. It is clear from visual inspection that model M1 predictions are in better agreement with the experiment data than are M0 predictions. We make this comparison more precise using KL divergence in Fig. 12.

Temporal comparisons of distributions between models and experiments for individual variables Appendix B.1.6 in Cedeno (2019) shows the figures resulting from the temporal analysis by minute of distributions between Models M0, M1 and Experiments for $k = 2$. Each plot contains data over a 1-min time window for each variable of x from Table 7. It is clear from visual inspection that model M1 predictions are in better agreement with the experiment data than are M0 predictions.

Comparisons of KL-divergence values between models for individual variables at the end of the anagram game Figure 12 shows KL-divergence values for comparing distributions of model outputs with corresponding distributions of experimental data for the anagram game. The models are (baseline) M0 and M1 for the ($n = 10$, $k = 2$) experiments. The comparisons are at the end of the game, i.e., at $t = 5$ min, over the entire game. For each experiment/model combination, the variables (and hence distributions) compared are: number of replies received, number of replies sent, number of requests received, number of requests sent, and number of words formed. Lower values are better. This figure shows that M1 generates predictions much closer to the experimental data than does M0. For example, M1 significantly reduces the reply-related and words formed KL-divergence values (weaknesses of model M0 as shown in Fig. 10).

Temporal comparisons of KL-divergence values between models for individual player actions Figure 27 shows the temporal KL-divergence values for the baseline M0 and M1 across the five parameters of x: lower values are better. Each plot contains data over a time window: Fig. 27a for 0–1 min, Fig. 27b for 1–2 min, Fig. 27c for 2–3 min, Fig. 27d for 3–4 min, and Fig. 27e for 4–5 min time intervals of the 5-min anagram game.

The plots demonstrate that KL-divergence values for the model M1 predictions are closer to the experimentally-determined data distributions than are those from model M0. While Model M0 has good predictions for the minute 3 and minute 5 (with the exception of the words formed), Model M1 has better predictions for the minute 3 and minute 5 for

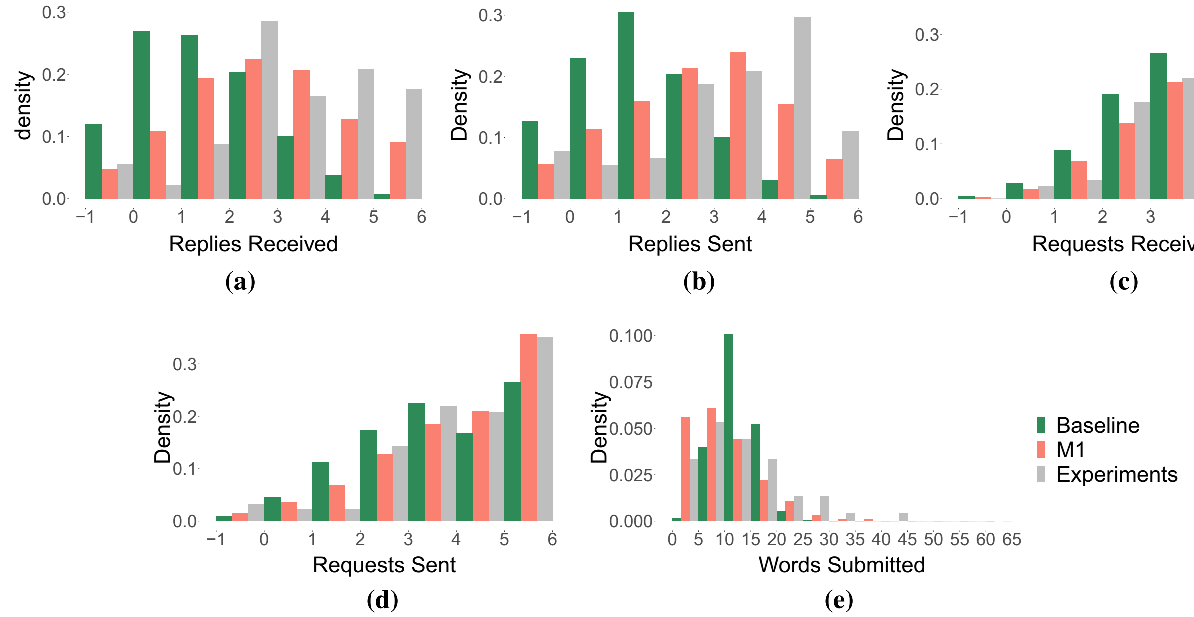

Fig. 11 Baseline Model M0 and Model M1 predictions of the $k = 2$ experiments, along with the experimental data. The probability density function is show for **a** distribution of replies received, **b** distribution of replies sent, **c** distribution of requests received, **d** distribution of requests sent, and **e** distribution of words formed, each at the end of the 5-min anagram game (gray bars) for all $k = 2$ experiments, compared to M1 predictions (red) for 100 simulations of an $n = 10$ player game. M1 predictions (red) for 100 simulations of an $n = 10$ player game. The Baseline Model M0 is shown in green for comparison. It is clear from visual inspection that model M1 predictions are in better agreement with the experiment data than are M0 predictions. We make this comparison more precise using KL divergence in Fig. 12 (color figure online)

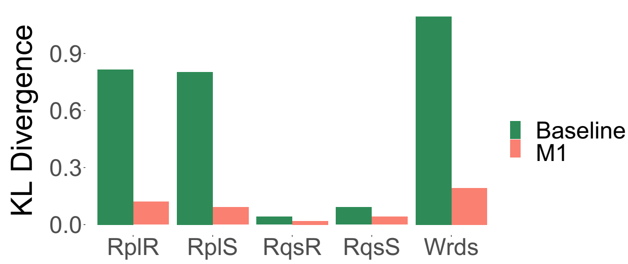

Fig. 12 KL-divergence values for the Baseline Model M0 and Model M1 across the five parameters of x: lower values are better. The modeling conditions are those of the experiments with $k = 2$. This figure shows that M1 greatly improves a weakness of model M0 in poorly representing RplR (number of replies received), RplS (number of replies sent), and Wrds (number of words formed)

all five x variables of Table 7. These data are significant because they evaluate the quality of the models to predict behavior temporally. That is, just because a model can produce predictions at the end of some scenario, this does not mean that it can capture the trajectory (or time evolution) of phenomena. With these types of plots, we demonstrate that our models do capture temporal behavior.

Comparisons of KL-divergence distributions between models and experiments for combining all variables Figure 3.19 in Cedeno (2019) shows the distribution of KL divergence for comparing distributions of model output with corresponding distributions of experimental data for the anagram game. The models are $(n = 10, k = 2)$ M0 and M1. The data sets used in comparison are experiments: $(n = 10, k = 2)$. There are 30 values in the distribution, with five values for each variable x at the end of the game, and 25 values for the five variables x over five intervals of 1-min increment. It shows that for Model M1, the great majority of KL-divergence values are less than 0.2, while they can be much greater for Model M0.

Summary of M0 and M1 model comparisons. Clearly, ABM M1 is in better agreement with the experimental data compared to the baseline model. From KL-divergence values in Fig. 12, it is clear that the predictions of M1 represent the experimental data better than those of the baseline model.

In addition, we use M1 to make predictions for anagram games with $k > 2$, resulting in more interactions. Counterintuitively, as shown in Fig. 13, the number of replies does not change as k increases. These results call for more experiments at larger k. Note that we exercise M1 learned from experiments with $k = 2$. The results in Fig. 13 indicate that M1 predicts no changes in the number of letter replies received as k increases, which seems counter intuitive. One would expect more letter requests and replies with increasing numbers of neighbors. These types of data lead us to construct model M2.

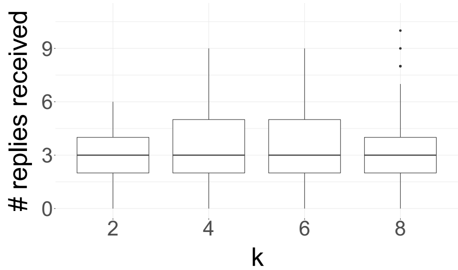

Fig. 13 M1 model distributions predicted for the number of replies received at the end of game ($n = 10$, 100 simulations), for different regular degrees k of the game network G. This partially motivated our development of ABM M2, since model M1 predictions do not vary significantly with the number of a player's neighbors

We remark that we also fitted M1 using experimental data with $k = 4$ (call this Model M1b), and consequently made predictions for the case of $k = 2$. We compared the distributions of x between prediction and experimental results using KL-divergence, and determined values in the range 0.11 to 0.46, indicating good predictions. Note that Model M1b is interpolating when it predicts $k = 2$ experimental data, while Model M1 is extrapolating to predict $k = 4$ experimental data.

5.6 Agent-based model M2

5.6.1 ABM M2 development

Model M1 was developed with data where all game players have the same degree $k = 2$. To generalize M1 to incorporate various k, we conducted additional experiments with $2 < k \leq 8$ as a part of the second AL (Sect. 7.4). We resume from the description of Model M1 in Sect. 5.5 and Eqs. (3) and (4).

Recall that for Model M1, $\beta_{jh}^{(i)}$ in $\boldsymbol{B}^{(i)}$ denotes parameters that are used to compute the transition probability π_{ij} based on a player taking action a_i at time t and action a_j at $(t + 1)$. Since there are $m = 4$ possible player actions in action set A, $1 \leq i, j \leq m$. As in Sect. 5.5, h is the index of an element of the z vector with $1 \leq h \leq g$. The $\beta_{jh}^{(i)}$ are determined from analyses of the transitions in the experimental data.

Here, we build a hierarchical model to incorporate the effect of agent degree k. For different values of k, the parameter coefficients in $\boldsymbol{B}^{(i)}$ of Eq. (4), used in Eq. (3), are now a function of k, denoted as $\boldsymbol{B}^{(i)}(k)$. We use an orthogonal polynomial basis to construct a continuous and smoothing function for $\beta_{jh}^{(i)}(k)$ for any given i, j, h, as

$$\beta_{jh}^{(i)}(k) = \alpha_{j,h}^{(i,0)} + \alpha_{j,h}^{(i,1)} \xi_l(k) + \alpha_{j,h}^{(i,2)} \xi_q(k), \tag{5}$$

where ξ_l and ξ_q are the linear and quadratic functions of the orthogonal basis in terms of k. We have

$$\boldsymbol{B}^{(i)}(k) = \boldsymbol{C}_0^{(i)} + \boldsymbol{C}_1^{(i)} \xi_l(k) + \boldsymbol{C}_2^{(i)} \xi_q(k), \tag{6}$$

where

$$\boldsymbol{C}_r^{(i)} = \begin{pmatrix} \alpha_{11}^{(i,r)} & \alpha_{12}^{(i,r)} & \dots & \alpha_{1,g}^{(i,r)} \\ \alpha_{21}^{(i,r)} & \alpha_{22}^{(i,r)} & \dots & \alpha_{2,g}^{(i,r)} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{41}^{(i,r)} & \alpha_{42}^{(i,r)} & \dots & \alpha_{4,g}^{(i,r)} \end{pmatrix}, \quad r = 0, 1, 2, \tag{7}$$

correspond to the constant, linear, and quadratic coefficient matrices in Eq. (6), with $\alpha_{ih}^{(i,r)} = 0$ for any r and h. Here, $\alpha_{jh}^{(i,r)}$ in Eqs. (5) and (7) are the elements of $\boldsymbol{C}_r^{(i)}$ determined from the successive pairs of actions a_i at time t and a_j at time $t + 1$, with current values of z.

Equation (4) is a special case of Eq. (6) when $\boldsymbol{C}_1^{(i)} = \boldsymbol{0}$ and $\boldsymbol{C}_2^{(i)} = \boldsymbol{0}$, i.e., when the coefficient matrix $\boldsymbol{B}^{(i)}$ is not a function of k.

5.6.2 Inductive inference

We address the three dimensions of inference, as for M1: (1) model structure; (2) model parameters; and (3) parameter values. In this case, the model structure we employ to capture the effect of k was identified a priori. However, if the model structure was found lacking, we would have tried another approach. The model parameters given in Eqs. (6) and (7) were also anticipated owing to the development of ABM M1. Hence, these first two steps were not solely driven by the data. To estimate the parameters sets $\boldsymbol{C}_0^{(i)}, \boldsymbol{C}_1^{(i)}, \boldsymbol{C}_2^{(i)}$, we use maximum likelihood estimation across the experimental observations for $k = 2$, 4, 6, and 8. For a given i and k, denote the corresponding observational data as $\mathcal{D}_k^{(i)}$. Then we conduct parameter estimation by

$$\hat{\boldsymbol{C}}_0^{(i)}, \hat{\boldsymbol{C}}_1^{(i)}, \hat{\boldsymbol{C}}_2^{(i)} = \arg\max \sum_{k=d_{min}}^{d_{max}} \log L(\boldsymbol{C}_0^{(i)}, \boldsymbol{C}_1^{(i)}, \boldsymbol{C}_2^{(i)} | \mathcal{D}_k^{(i)}),$$

where $L(\boldsymbol{C}_0^{(i)}, \boldsymbol{C}_1^{(i)}, \boldsymbol{C}_2^{(i)} | \mathcal{D}_k^{(i)})$ is the likelihood function with respect to the data $\mathcal{D}_k^{(i)}$ collected under the setting of k neighbors in the experiments of Sect. 4.

5.6.3 ABM model M2 results

Results for model M2 are provided according to Table 8. Results are often compared to those for model M1.

Comparisons of distributions between models and experiments for individual variables at the end of the anagram game Figure 14 shows data distributions at the end of the 5-min anagram game (gray bars) for all $k = 2$ experiments, compared to M2 predictions of distributions (blue) for 100 simulations of an $n = 10$ player game. These results are over all 5 min of the group anagram game. Figure 3.22 in Cedeno

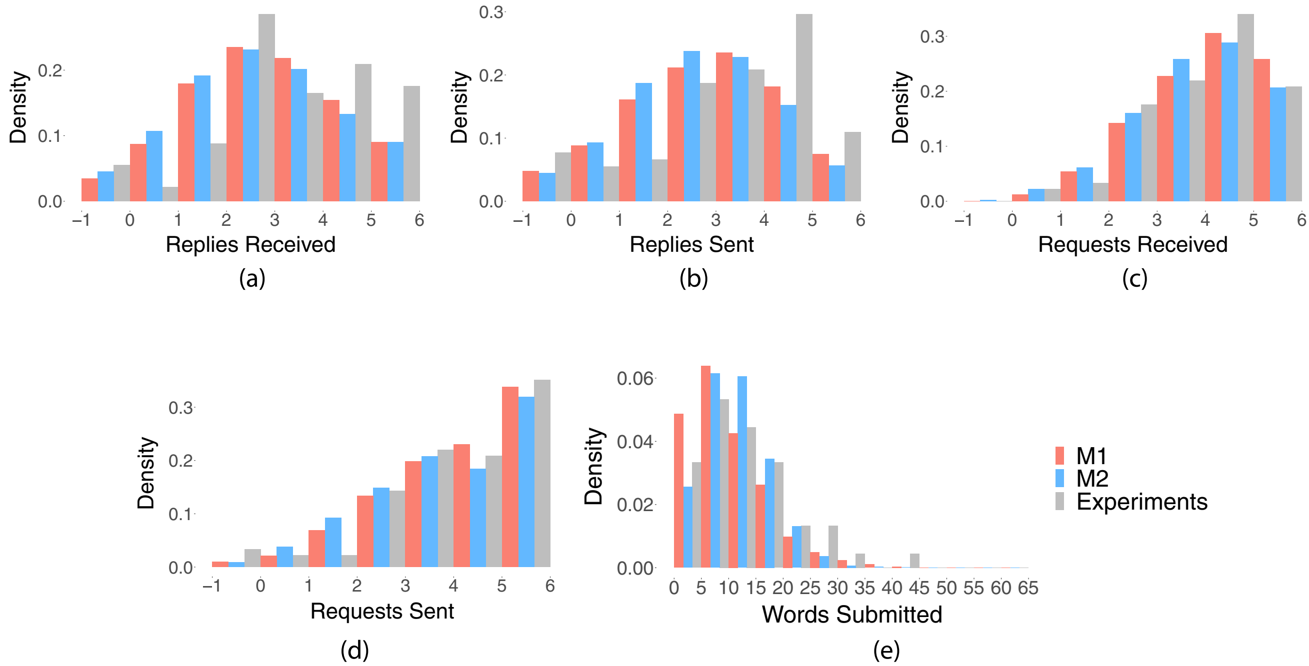

Fig. 14 Model M1 and Model M2 predictions of the $k = 2$ experiments, and experimental data, over all 5 min of the group anagram games. The probability density distributions are shown for **a** distribution of replies received, **b** distribution of replies sent, **c** distribution of requests received, **d** distribution of requests sent, and **e** distribution of words formed, each at the end of the 5-min anagram game (gray bars are experimental data) for all $k = 2$ experiments, compared to M2 predictions (blue) for 100 simulations of an $n = 10$ player game. The Model M1 predictions are shown in red for comparison. It is clear from visual inspection that models M1 and M2 generate similar predictions, in agreement with the experiment data, as M1 is learned solely from $k = 2$ experimental data. We make this comparison more precise using KL divergence in Fig. 15 (color figure online)

(2019) shows data distributions at the end of the 5-min anagram game (gray bars) for all $k = 4$. In Appendix B.1.7, in Cedeno (2019), Figure B.15 shows data distributions at the end of the 5-min anagram game (gray bars) for all $k = 6$. Figure B.16 Model M1 is shown in red for comparison. In all of these figures, Figure (a) shows the distributions of replies received, Figure (b) shows the distributions of replies sent, Figure (c) shows the distributions of requests received, Figure (d) shows the distributions of requests sent, and Figure (e) shows the distributions of Words Formed. M2 gives much better performance, as expected, as it explicitly accounts for agent degree. As expected, M1 and M2 perform equally well for $k = 2$ as M1 is learned from $k = 2$ experimental data.

Temporal comparisons of distributions between models and experiments for individual variables Appendix B.1.8 in Cedeno (2019) shows the temporal analysis by minute of distributions for models M1 and M2 and experiments for $k = 2$. Each plot contains data over a time window of 1 min. For $k = 2$ experiments, Figure B.17 shows temporal analysis for the number of Replies Received at the end of each minute. Figure B.18 shows temporal analysis for the number of Replies Sent at the end of each minute. Figure B.19 shows temporal analysis for the number of Requests Received at the end of each minute. Figure B.20 shows temporal analysis for the number of Requests Sent at the end of each minute. Figure B.21 shows temporal analysis for the number of Words Formed at the end of each minute.

Collections of plots for each of $k = 4$, 6, and 8 are analogously provided in Appendix B.1.8 in Cedeno (2019) . As expected, M1 and M2 perform equally well for $k = 2$, as M1 is learned from $k = 2$ experimental data. For $k > 2$, M2 performs better. We make this comparison more precise using KL divergence below.

Comparisons of KL-divergence distributions between models and experiments for individual variables at the end of the anagram game Figures 15 and 16 in this section, and Figures B.37 and B.38 in Appendix B.1.9 in Cedeno (2019) show KL-divergence values for comparing distributions of model outputs with corresponding distributions of experimental data, for the group anagram game. The figures are for, respectively, $k = 2$, $k = 4$, $k = 6$, and $k = 8$ experiments. The models are M1 (red) and M2 (blue). These four figures show clear and interesting behavior. Model M1 agrees better with experiments than does Model M2 for $k = 2$, since Model M1 was specifically developed with $k = 2$ data. However, for larger k ($4 \le k \le 8$), Model M2 does better than M1. This is because Model M2 was developed using data

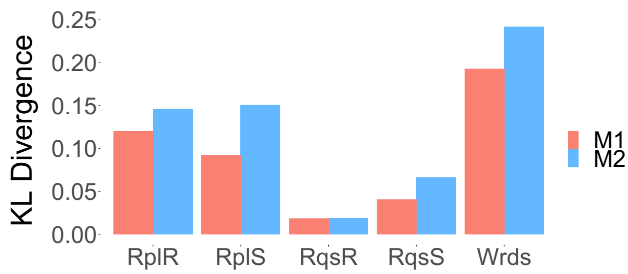

Fig. 15 The plot shows on the x-axis KL-divergence values for the M1 and M2 model predictions at the end of the 5 min anagram game. Here we compare $k = 2$ data for M1 and M2 model predictions with the experiments across the five parameters of x: lower values are better. This figure shows that model M1 and M2 generate similar predictions to the experimental data, with Model M1 slightly better (color figure online)

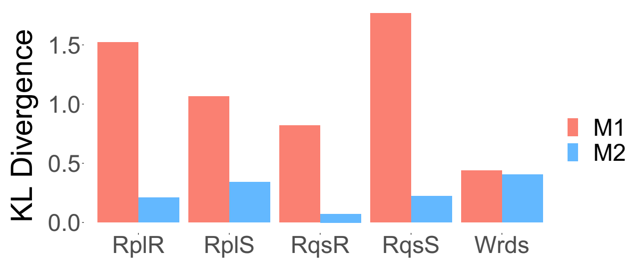

Fig. 16 The plot shows on the x-axis KL-divergence values for the M1 and M2 model predictions at the end of the 5 min anagram game. Here we compare M1 and M2 model predictions of the $k = 4$ data with the experiments across the five parameters of x: lower values are better. This figure shows that M2 gives much better performance than M1 in predicting the time to generate an action for an agent. M2 gives better performance, as expected, as it explicitly accounts for agent degree (color figure online)

across all of these k values. Hence, to obtain a wider range in input space for simulations, our Model M2 does slightly worse for a particular k ($k = 2$).

Temporal comparisons of KL-divergence distributions between models and experiments for individual player actions This section shows the temporal KL-divergence values for the model M1 and M2 predictions across the five parameters of x. Lower values are better. Figure 3.25 in Cedeno (2019) shows $k = 2$ experiments, and Figure 3.26 in Cedeno (2019) shows $k = 4$ experiments. In Appendix B.1.10 in Cedeno (2019) Figure B.39 shows $k = 6$ experiments, and Figure B.40 shows $k = 8$ experiments. Each plot contains data over a 1-min time window, as in previous analyses. It is clear from visual inspection that model M2 predictions are in better agreement with the experiment data than are M1 predictions for $k > 2$. As noted above, however, Model M1 does slightly better for $k = 2$. That is, the comparisons between models M1 and M2, for temporal variations in 1-min time intervals over the 5-min group anagram

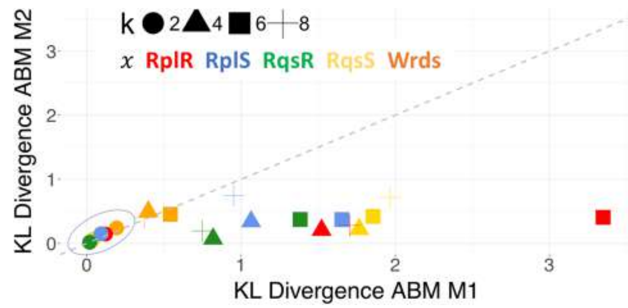

Fig. 17 A scatter plot of KL divergence for M1 (x-axis) and M2 (y-axis) for four k values and five x variables. For $k > 2$, M2 performs better than M1, as M2 incorporates experimental data with $2 \le k \le 8$. Interestingly, M1 and M2 perform equally well (highlighted) for $k = 2$ as M1 is learned from $k = 2$ experimental data (M1 is slightly better). These are data over the total 300 s anagram game (color figure online)

game, are similar to those comparisons when combining all data into one analysis over the entire 5-min game.

Summary of M1 and M2 model comparisons In addition to Figure 3.27 in Cedeno (2019) discussed immediately above, Fig. 17 compares Model M1 and Model M2 for each of the five actions in x, accumulated through the 5-min group anagram game. Model M2 does not perform quite as well as Model M1 for the $k = 2$ data, but does better than M1 for $k = 4$, 6, and 8. Thus we sacrifice some quality for $k = 2$ and get in return capabilities over a range of k. Hence, Model M2 is of greater value, since it covers a broader range of inputs for simulations.

In Appendix B.1.11 in Cedeno (2019), Figure B.41 shows the boxplots grouped by type of $k = 2, 4, 6, 8$, where each box contains five values of KLD corresponding to the five x variables at the end of each minute. The plot show that our models show highest median values on the first 2 min of the game.

6 Model evaluation

This section contains evaluations of Model M2 from Sect. 5. Our goal is to understand the conditions for which our estimated model transition probabilities π_{ij} are sufficiently accurate.

To evaluate the goodness of fitting for the proposed hierarchical model, we compare the estimated (model) transition probability matrix $\hat{\Pi} = (\hat{\pi}_{ij})$ with the empirical (data) transition probability matrix $\tilde{\Pi} = (\tilde{\pi}_{ij})$ under different settings of covariates (the z vector of Table 9). Here, the empirical transition probability matrix $\tilde{\Pi}$ is obtained under the settings by grouping the value of each covariate into three levels, as described in Table 10, to obtain comparable numbers of samples across bins.

Table 10 Three bins and ranges of values for the z variables from Sect. 5.5. These bins are created for each of the four values of $k \in \{2, 4, 6, 8\}$

Level	Buffer (z_B)	Letter (z_L)	Word (z_W)	Consec. (z_C)
1	0	0–3	0–1	0–3
2	1	4–6	2–8	4–11
3	≥ 2	≥ 7	≥ 9	≥ 12

For each setting, there is a level combination of the four covariates. We compute a counting matrix $\mathcal{N} = (n_{ij})$, where n_{ij} is the number of data instances observed for the transition from action i to next action j across all players in group anagram games. (Here, actions i and j refer to actions a_i, $a_j \in A$ in Table 5.) We then calculate the empirical probability $\hat{\pi}_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$. There are 324 settings in total from the grouping of variables in Table 10 (three settings of each of four variables, and four k values), and 279 of them have valid empirical transition probability matrices. For the estimated transition probability matrix $\hat{\Pi} = (\hat{\pi}_{ij})$, the value of $\hat{\pi}_{ij}$ is estimated by the proposed model under each setting of covariates, where the averaged value at each level of the covariate is used in the estimated model.

The squared Root of Mean Squared Errors (RMSE) is used to quantify the difference between $\hat{\Pi} = (\hat{\pi}_{ij})$ and $\tilde{\Pi} = (\tilde{\pi}_{ij})$. RMSE is calculate as follows:

$$\text{RMSE} = \sqrt{\frac{1}{4|\mathcal{I}|} \sum_{i \in \mathcal{I}} \sum_{j=1}^{4} (\hat{\pi}_{ij} - \tilde{\pi}_{ij})^2} \tag{8}$$

where $\mathcal{I} = \{i : \min_j n_{ij} > 0\}$ is the index set of the rows where the empirical probability can be obtained.

Figure 18 shows the scatter plot between the RMSE and $n_{\min}$ for the 279 settings for which there are sufficient data, where the plot is in log10-log10 scale. From the figure, the proposed method generally provides an accurate estimation of probability transition matrix in most of settings. Clearly, the value of RMSE decreases as the Min.Count $n_{\min}$ increases. When $n_{\min} \geq 100$, the value of RMSE is smaller than 0.069, showing a very good model fitting. When $n_{\min}$ is small, the RMSE is relatively high. One explanation is that the empirical probabilities cannot be calculated accurately when $n_{\min}$ is small.

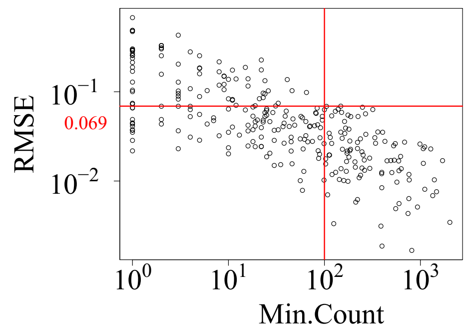

Fig. 18 Scatter plot of RMSE against Min.Count in different settings of covariates in Table 10. See Eq. (8) for RMSE and text for Min. Count. One hundred observations in a category drives RMSE down to 0.069

7 Abductive loop analyses and results

7.1 Overview

In this section, we present the results of iterative abductive analyses, described in Fig. 4. First, we "unroll" the abductive loop to illustrate several iterations of abduction and different analysis paths that can be taken, depending on results generated up to that point; see Sect. 7.2. Then, we present two ALs in the next two subsections. At the end of loop-2, we describe how ABM M2 can be used in further loops. We then summarize the findings of the abductive iterations, and discuss their generality, and candidate research questions for future work. We note that the experiments (Sect. 4) and modeling (Sect. 5) are major components of the abductive looping process, and were separated out to make this section more streamlined.

7.2 Abductive iterations with hypotheses

Figure 19 provides a tree structure representation of several *candidate* abductive loops. Specifically, an iteration or loop of the abductive looping process, shown in Fig. 4, is represented by a node in the graph of Fig. 19. Each loop (Fig. 4) specifies and evaluates at least one hypothesis. Figure 19 emphasizes a hypothesis H_{ij} at each node. These hypotheses are stated in Table 11. The last step in a loop is to determine what is to be done next, and the options are represented by the edges out of a node, extending toward its child nodes in Fig. 19 (a node may have any number of children). Hence, the tree graph in Fig. 19 is motivated by the usefulness of representing several abductive loops and their dependencies in a compact fashion, as opposed to writing down Fig. 4 for each loop.

The hypotheses in Fig. 19 and Table 11 are candidates because they depend on the data generated as successive abductive iterations are completed (Haig 2005; Timmermans and Tavory 2012). Hence, it may be modified with iterations. The tree here is not unique. Different analysts may compose different hypotheses and different trees, and a tree will generally need to be modified as analyses unfold. Nonetheless, it is a useful exercise to construct such trees as part of reasoning about a problem.

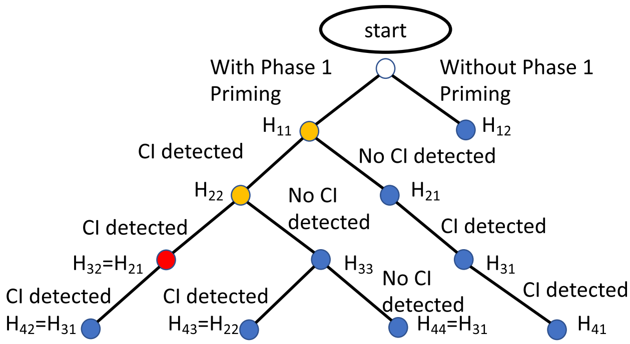

Fig. 19 An abductive tree representing candidate abductive loops with dependencies. Each node in the figure is one loop (see Fig. 4). Since each loop has at least one hypothesis, we label the nodes here with hypotheses that are provided in Table 11. Edges to child nodes are labeled with outcomes from hypothesis evaluations within a loop, and indicate, based on this evaluation, which abductive loop to perform next. A node can have any number of outgoing edges to child nodes. The orange colored nodes correspond to abductive iterations presented herein. The red node is a candidate next loop. This tree is not unique; different analysts can devise different trees, and they can be modified as analyses proceed. The purpose of this construction is to provide a succinct representation of multiple candidate abductive loops, and their dependencies (color figure online)

We now overview the two abductive iterations detailed in subsequent sections. The root node of the tree in Fig. 19 is the starting point. We conduct experiments with the Phase 1 (group anagram priming game), so we take the left branch (edge) from the root node, labeled "With Phase 1 Priming". We perform an abductive loop, where we form and evaluate hypothesis H_{11}. This loop is detailed in Sect. 7.3. Since it is clear that CI was formed, we follow the "CI detected" path out of H_{11} to arrive at the node hypothesis H_{22}. This AL-2 is described in Sect. 7.4. Since we do obtain a CI signal from these experiments, we follow edge "CI detected" to hypothesis H_{32}. Details are provided below, and we note that modeling results guide decisions about what experiments to perform in the next abductive iteration, illustrating the value of modeling. This is one reason why our abductive loops promote modeling to a central role. We also note that a hypothesis can appear at multiple nodes within the abductive tree, e.g., the group anagram game conditions may be altered after particular loops. Finally, a node need not have two children; e.g., fewer or more children are possible.

7.3 Abductive Loop 1 (AL-1)

We describe the steps of the abductive loop in Fig. 4, in turn.

Experiments A set of 18 experiments with a total of 87 players was completed where $k = 2$. See Sect. 4.

Data Analysis For this loop, data analysis and modeling are intertwined and so both are described under the *models* step below. It is critical to note that the data analyses below came *before* the specification of hypotheses, because a critical element of abduction is that patterns in the data drive the hypotheses—not the other way around.

Hypothesis/Theory Hypothesis H1 (H_{11}): *In the team-based anagram game, the CI formed is driven more by the number of words a player forms than the number of interactions of a player (requests and replies).* Social Exchange Theory (Homans 1961) focuses on the individual and suggests that the number of words resonates more in forming CI because they are directly related to reward in the game. Theory of Social Interactions (Becker 1974) indicates that interactions are important for forming an interdependent

Table 11 Candidate hypotheses to evaluate in abductive iterations. Not all of the hypotheses are evaluated herein. The goal of these hypotheses, coupled with Fig. 19, is to illustrate that there are many possible hypotheses that can be formulated, and it is up to analysts to decide which ones to pursue. An analyst will be guided by the results of completed iterative abductive analyses

Hypothesis Number	Description
H_{11}	In the team-based anagram game, the collective identity formed is driven more by the number of words a player forms than the number interactions of a player (requests and replies)
H_{12}	Playing the group anagram game will produce greater individual DIFI scores than not playing the group anagram game
$H_{21} = H_{32}$	As the number and quality of letters assigned to a person decreases (i.e., as the letters assigned to a player occur less frequently in common words), collective identity of the player will increase
$H_{22} = H_{43}$	(a) As the number of neighbors of a player increases in the anagram game, the number of interactions of a player increases, but only up through degree $k = 4$. For further increases in k, there will be no increase in numbers of interactions (b) The trend in numbers of interactions with k in hypothesis (a) will be reflected in the CI produced in group anagram games: CI will increase for $2 \le k \le 4$, but will saturate thereafter because the numbers of interactions saturate
$H_{31} = H_{42} = H_{44}$	Playing the game with players face to face will produce greater individual DIFI scores (by enabling the players to communicate and pick up on visual and verbal cues)
H_{33}	Lesser payouts in the group anagram game means that players do not have enough incentive to engage their neighbors
H_{41}	Having the group anagram game score of another team displayed during the game will increase CI because it will create a stronger in-group/out-group paradigm

organization. Reciprocity Theory suggests that v_i will respond to v_j's requests because v_i wants v_j to respond to hers, so that interactions are important.

Models There are two types of models constructed. One type is the models of the group anagram game: Baseline Model M0 and M1 from Sect. 5. The other is a regression model to predict DIFI2 score as a function of outputs from the group anagram games (e.g., number of requests sent n_{RqsS}, number of replies received n_{RplR}).

The group anagram game models M0 and M1 of Sects. 5.4 and 5.5 were constructed from the time histories of actions of players for experiments with $k = 2$. The results relevant to this abductive iteration are provided in Figs. 11 and 12. Model M1 is much better at capturing the dynamics in the experiments than is Baseline Model M0.

From data of the actions $a_i \in A$ from the anagram games, and the measured DIFI2 scores after the group anagram games, a linear regression was performed to correlate DIFI2 score with the numbers of actions of each kind for game players. The DIFI2 score is given as

$$\widehat{\text{DIFI2}} = c_1 + c_{RplR}\, n_{RplR} + c_{RplS}\, n_{RplS} + c_{RqsR}\, n_{RqsR} + c_{RqsS}\, n_{RqsS} + c_{Wrds}\, n_{Wrds} \tag{9}$$

where Table 12 provides the equation coefficients and the definitions of variables.

Table 13 provides the regression results that identify the player actions that correlate with DIFI score. At the 0.05 level, replies received, replies sent, and requests sent are all significant.

Best Explanation Results of a linear regression in Table 13 indicate that hypothesis H1 is falsified because the number Wrds, i.e., the number of words formed, is not significant, while numbers of RplR, RplS and RqsS (i.e., interactions) are significant. Thus, Social Exchange Theory can be eliminated as a theory of CI formation in this experiment. It is somewhat surprising that Wrds is not significant because it is the variable that is most closely associated with the reward (earnings). A conjecture was made that the greater the monetary reward given to players, the greater their affinity would be for the team; these data do not support this conjecture. In the social sciences, and in many domains, eliminating candidate theories is a valuable result (that is, an analysis does not always have to identify the best theory). Thus, at this point, the best explanation is Reciprocity Theory and Theory of Social Interactions because the analyses results in Table 13 show that interactions correlate most strongly with DIFI score.

A key result indicated by this first iteration is that the group anagram game can produce CI (as measured by the proxy DIFI score).

Table 12 Constants in the regression of Eq. (9) to predict DIFI2 score of a player from the player actions in the team anagram game

Coefficient	Value
Intercept c_1	102.7
c_{RplR} on number of replies received n_{RplR}	14.95
c_{RplS} on number of replies sent n_{RplS}	− 12.99
c_{RqsR} on number of requests received n_{RqsR}	6.406
c_{RqsS} on number of requests sent n_{RqsS}	− 16.43
c_{Wrds} on number of words formed n_{Wrds}	− 0.2134

What is Next? Figure 13 indicates that Model M1 predicts player behavior that is invariant with respect to the degrees k of players [and hence the number of letters that neighbors possess] (plots of other variables of x are similar). We want to determine whether there is an effect of k, and hence the next experiments are specified to study increasing k (i.e., $k > 2$). Thus, the ABM M1 (driven by the data) is guiding what to do next. While Social Exchange Theory was eliminated in this loop, Reciprocity Theory and Theory of Social Interactions are carried forward into the next loop(s), where they may be supported or refuted.

7.4 Abductive Loop 2 (AL-2)

We execute the steps of the abductive loop in Fig. 4, as described next.

Experiments A set of 16 experiments with a total of 137 players was completed where $k = 4$, 6, and 8 in turn. See Sect. 4.

Data Analysis We continued the same types of analyses described in AL-1, but with the added dimension of k. Figure 20 shows the frequency distributions for replies received, for the four values of k. Note the large change in distributions in going from $k = 2$ to $k = 4$, but relatively minor changes for further increases in k. This indicates two regimes of behavior: (1) $2 \leq k \leq 4$, and (2) $k > 4$. In the first regime, numbers of interactions increases with k, and in the second regime, the numbers of interactions does not appreciably increase with further increases in k. Additional data analyses are presented in step *best explanation* below.

Hypothesis/Theory Two hypotheses are formed based from the preceding data. Hypothesis H2 (= H_{22}): *(a) As the number of neighbors of a player increases in the anagram game, the number of interactions of a player increases, but only up through $k = 4$. For further increases in k, there will be no increase in numbers of interactions. (b) The trend in numbers of interactions with k in hypothesis (a) will be reflected in the CI produced in group anagram games: CI will increase for $2 \leq k \leq 4$, but will saturate thereafter because the numbers of interactions saturate.*

Table 13 Results of linear regression of variables in x (see Table 7) against dependent variable DIFI2 score, indicating that interactions are more significant than number of words formed in producing CI

Var.	Interc.	RplR	RplS	RqsR	RqsS	Wrds
est.	103.	15.0	− 13.0	6.41	− 16.4	− 0.213
p val.	0.001	**0.019**	**0.011**	0.332	**0.011**	0.735

The values in bold indicate that the number of replies received (RplR), number of replies sent (RplS) and number of requests sent (RqsS) are significant

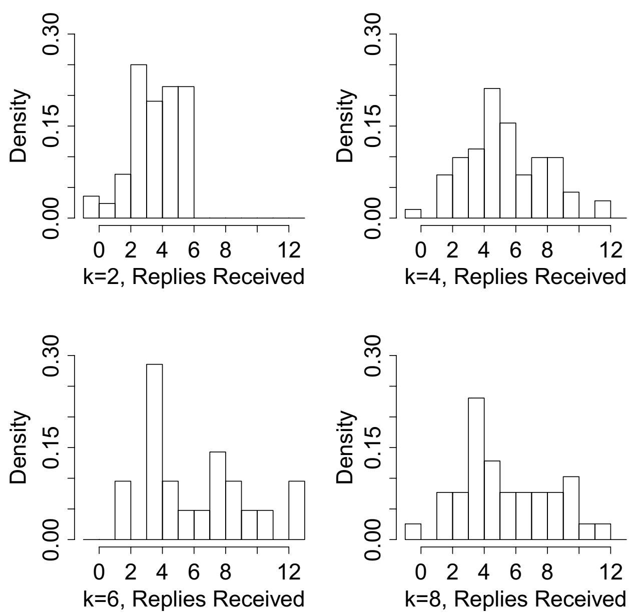

Fig. 20 Statistical analysis correlation results of the anagram game parameters. The probability density of replies received changes markedly from $k = 2$ to $k = 4$, but relatively little for further increasing k up to eight

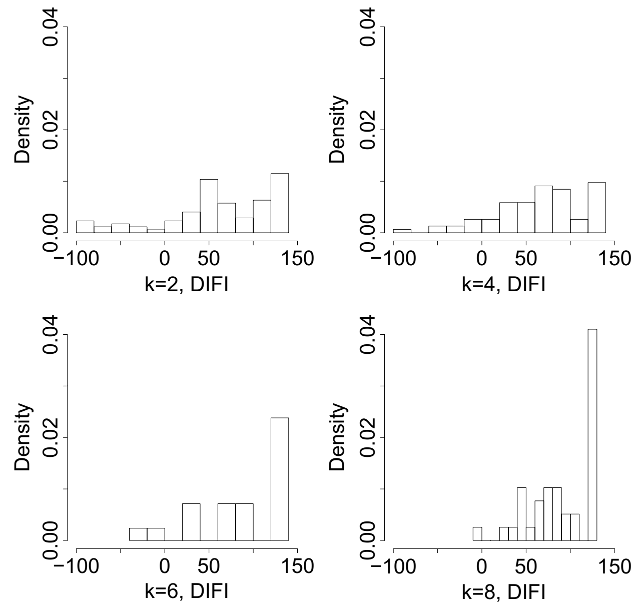

Fig. 21 Statistical analysis correlation results of the anagram game parameters and DIFI2 score. The probability density of DIFI2 score moves to larger DIFI2 score with increasing k, between two and eight

Model Model M2 of Sect. 5.6 was constructed from the time sequences of actions of players, from the combined data from *both* iterations. Model results relevant to this iteration are provided in Figs. 14 through 17. ABM M2 captures trends in degree k more effectively than ABM M1, for all parameters of x.

Best Explanation Figures 20 and 21 provide results that address the hypotheses.

For $H_{22}(a)$, we return to Fig. 20 and the observations under *data analysis*. We note a saturation in the distributions of replies received (distributions for other actions are similar). As the number of neighbors of a player increases from two to four, the numbers of interactions increases; but for $k > 4$, the number of interactions does not change appreciably with further increases in k. This can be explained by Utility Theory and by Cognitive Load Theory. Utility theory argues that 15 letters is enough for a player to form words (three letters from each of four neighbors and three own letters of a player), so a player derives no marginal utility from more neighbors and more letters. Cognitive load theory states that a player cannot reason about forming words with more than 15 letters, so no attempt is made to acquire more letters. Hence, hypothesis H_{22}(a) is not falsified. It is possible in future work to conduct more experiments, possibly by modifying the priming procedures, to disambiguate these two explanations.

We note the agreement between our hypothesis and the results of the experiments. This is a direct result of the abductive approach: the data guide the hypotheses. We stated in H_{22}(a) that the transition between the two regimes occurred at $k = 4$ because the data in Fig. 20 indicated this. It is the goal of hypothesis and theories to *explain* the behavior.

For hypothesis H_{22}(b), we use Fig. 21, showing how the probability density of DIFI score moves to higher scores as k increases from two through eight. Thus, Figs. 20 and 21 together support H_{22}(b), but only for $k \leq 4$. The explanation in the hypothesis—that greater DIFI scores is caused by greater numbers of interactions—is not supported by the combined view of Figs. 20 and 21 for $k > 4$. The latter figure shows an ever increasing probability density of greater DIFI score with increasing k, but the former figure shows an essential saturation of numbers of actions with increasing k beyond $k = 4$. Hence, the numbers of interactions may be contributing to increasing DIFI scores for $2 \leq k \leq 4$, but appears not to be not the reason for increasing DIFI scores for $k > 4$. Consequently, the Theory of Social Interactions can explain the results for $2 \leq k \leq 4$, but this theory is falsified for $k > 4$.

What is Next? At this point we halt the iterative abduction process for this paper. In a next iteration, we could try to isolate the effects of number of interactions versus the number of neighbors in different experiments on DIFI score, to more fully explain the results in AL-2. Several other directions are possible, guided by the hypotheses in Table 11: evaluating the quality of letter assignments, varying the number of letters per neighbor, playing face-to-face games, adding competition for the team. We could also perform a deductive (confirmatory) analysis by making specific quantitative predictions for experiments using ABM M2 as part of AL-2, and running corresponding experiments in AL-3.

7.5 Summary of experimental contributions to the understanding of CI and possible extensions

Abductive loops one and two are presented in Sects. 7.3 and 7.4. Here, we summarize our findings and identify how our findings might be tested or extended by presenting a series of research questions for others to consider. First, we consider the generality of our specified hypotheses.

Hypotheses H_{11} and H_{22} are stated in Table 11. Since the number of words that players form is directly proportional to the monetary reward of players in the game, H_{11} may be restated in more general terms as *Hypothesis* H^{*}_{11}*: The collective identity formed is driven more by player reward in a game than the number interactions of a player with its neighbors*. Note that we have not removed the notion of interactions—which we effect by inducing a network on the players, but many games use networks to control interactions, so this is not a problem in our view. The hypotheses of H_{22} can be restated in more general terms as *Hypothesis* H^{*}_{22} *: (a) As the number k of neighbors of a player increases, the number of interactions of a player increases, but only up to a (saturation) point denoted by a critical value of degree* k^{*}*. For further increases in* $k > k^{*}$*, there will be no increase in numbers of interactions. (b) The trend in numbers of interactions with k in hypothesis (a) will be reflected in the CI produced: CI will increase up to the saturation point, but will remain relatively constant thereafter because the numbers of interactions saturate.*

There is no mention of the group anagram game or our DIFI score task in these restated hypotheses. Specifically, our hypotheses and findings need not be specified in terms of the particulars of our game, such as particular letters requested of neighbors or specific words formed. Rather, our hypotheses and findings may be stated in terms of game rewards (i.e., earnings), numbers of connections (i.e., network degree) of players, and numbers of player interactions. These are basic features. One reason to state the hypotheses in terms of the games, as we did in Sects. 7.3 and 7.4, is because it makes clearer how the analyses tie in with the hypotheses. We return to the implications of these restatements after summarizing our findings.

In Sect. 7.3, we find that CI is *not* produced owing to the parameter most closely aligned with the payout (i.e., earnings) to the group: the number of words formed. Since a player's monetary benefit is most closely and directly tied to the number of words formed, one might conjecture that this variable would most closely correlate with the DIFI score (i.e., our measure of CI). That is, we speculated that success in the group anagram game, in terms of increases in monetary reward, would result in increases in players' affinities for the team. Our analyses indicate that this is not the case. Hence, financial rewards did not translate into greater CI. We found instead that numbers of interactions are more significantly correlated with greater DIFI scores, our proxy for CI, for the $k = 2$ data.

In Sect. 7.4, where we study the effects of $2 \leq k \leq 8$, we find that the numbers of interactions increase most rapidly as number of neighbors increases from two to four. As player degree increases from four to six, and then to eight, the numbers of interactions remains essentially constant. However, the DIFI score increases as player degree increases from two to eight. Thus, DIFI score increases with increasing numbers of interactions, as described in the abductive loops above, but increases further as numbers of interactions remain relatively constant, but numbers of neighbors increases.

As stated earlier in this subsection, our hypotheses and their evaluations can be stated in general terms. Consequently, it should be possible to test our findings using other group interaction approaches. These could be online or offline settings, computer-based or not computer-based. Player (participant) interactions can be of various forms: verbal communication, written communication, transactions or partial transactions in the form of actions, or selection (i.e., choosing a subset of neighbors from a player's complete set of neighbors). We use a network setting; this is not required, although it does provide the ability to control who might interact with whom. Also, we use DIFI score as a proxy for CI, based on the literature, but other means of inferring the existence of CI are also viable. These decisions are of course the prerogative of the researcher.

Nonetheless, we provide some specific questions that others may consider in their work. This list is not exhaustive, but illustrates how our findings may be tested and/or extended by others:

1. Do other graph structural properties, such as clustering coefficient, influence CI formation?
2. Does the proxy or measurement of CI increase with the number of neighbors (interactors) that a player has?
3. Does the proxy or measurement of CI increase with the number of interactions that a player has?
4. Is a one-shot, binary decision game as effective as a multi-action (or multi-decision) game where actions may be repeated over time, in producing CI?
5. Can the CI priming process (for us, the group anagram game) be more effective by introducing competition with another team? (There are many ways to do this, including using a fictional team, using data from a previous experiments where other teams did well, or poorly.)

As we stated in Sect. 1.4 when describing our technical challenges, we believe that our work is not the final word on CI, but rather is closer to a starting point, providing detailed methods and results for others to scrutinize and extend.

7.6 Abductive loops: role of analyst and bigger picture

Two ALs have been demonstrated. Many additional loops are possible, as illustrated in Fig. 19, which depicts several hypotheses, including the two addressed above (in orange). These additional loops would require more experiments. Figure 19 and Table 11 make clear the important role of an analyst in this process, as she guides the direction of the looping. So, while a plan such as that in Fig. 19 may be useful, the actual tree structure will evolve with analyst decisions as the looping progresses and as data are generated, because hypotheses are based on newly-generated data in abduction.

8 Limitations and additional work

More experiments, particularly at greater k would be useful. Also, we would like to alter the number of letters and to control the "quality" of letters that are assigned to players (e.g., e is a more desirable letter than q) in additional experiments, so that we can run experiments that will more stringently test the models. We would like to study more network structures (i.e., connectivity among players in a game), such as a clique structure. Beyond collecting more data, we could test more conditions, and also specify and evaluate more hypotheses about CI. We attempted to correlate player behavior with survey information in the online experimental platform. For example, we tried to correlate DIFI score with player age, gender, nationality, ethnic group, and education level. We did not get a strong signal in any of these correlation studies. This would be a huge step forward if such correlations exist and can be found because it would relate macro-player features with player behavior. With respect to modeling, we can improve the models for the player actions (e.g., the process of forming words); this work is in progress. We can improve the modeling in translating results in the group anagram game to the DIFI scores, to better understand the connection between priming and CI formation.

9 Summary

We formalize an abductive loop, implement it computationally, and exercise it in an experimental setting (the group anagram game) designed to induce CI, as operationalized by Swann's DIFI score. However, our abductive looping process is not tied to CI. As part of the abductive iterations, we provide novel experimental insights into CI and build and evaluate three ABMs. This work establishes the potential of iterative abductive looping for the (computational) social sciences.

Acknowledgements We thank the anonymous reviewers for their helpful feedback. We thank our colleagues at NSSAC and computer system administrators: Dominik Borkowski, Jason Decker, Miles Gentry, Jeremy Johnson, William Marmagas, Douglas McMaster, and Kevin Shinpaugh. This work has been partially supported by NSF CRISP 2.0 (Grant No. 1832587), DARPA Cooperative Agreement D17AC00003 (NGS2), DTRA CNIMS (Contract HDTRA1-11-D-0016-0001), DTRA Comprehensive National Incident Management System Contract HDTRA1-17-D-0023, NIH 1R01GM112938-01. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

A Supplemental related work

Related work topics that augment those in Sect. 3 are provided here. See Table 2 for a listing of all related work topics.

A.1 Individual anagram games: modeling

In Tresselt (1968), problem solving and verbal cues are analyzed with an anagram game. Tresselt (1968) modifies the H. Kendler and S. Kendler (1962) mediational model of problem-solving behavior (introducing word length and letter position), to understand anagram problem solving. This is a theoretical model of individual anagram games.

A.2 Individual anagram games: experiments and modeling

These works combine experiments and modeling. In Feather (1969), it was found that subjects who were initially confident of passing an anagram game test tended to attribute success to ability and failure to bad luck. However, subjects who were initially not confident tended to attribute success to good luck and failure to lack of ability. Results are discussed in terms of Heiderian theory and a valence-difficulty model. In Feather and Simon (1971a, b), two individuals played anagram games simultaneously but independently to test whether a person attributed her success (if she performed better) to skill versus good fortune, and failure to inferior skill or bad luck. Attributions were found to be dependent on expectations of players. Results are discussed in terms of models involving Heider's principle of balance and his analysis of the causes of action, in terms of positivity biases in social perception, and as indicating effects of the social context of performance upon attribution and valence.

A.3 Modeling of CI

Lustick (2000), Rousseau and van der Veen (2005) use ABMs to study identity diffusion. An agent adopts (changes) her type of identity to that of a neighbor with a stronger (higher valued) type of identity. Hence, these are contagion processes and are implemented much like voter models (de Oliveira 1992; Pereira and Moreira 2005). Other works modeling collective identity (van Zomeren et al. 2008; Chen and Li 2009; Benjamin et al. 2016; Ackland and O'Neil 2011) are presented in Sect. 3.3.

A.4 Agent-based models of anagram games and formation of CI

The Charness et al. (2014) work in Sect. 3.6 has no modeling for the group anagram game. This motivated the *online* experiments and ABMs in Ren et al. (2018). This article is an expansion of Ren et al. (2018). In this work, we model the priming process of producing CI, which is the group anagram game. There are no ABMs (or models of any kind) of group anagram games, to our knowledge, other than ours.

A.5 Studies of phenomena related to CI

Many phenomena, such as in-group and out-group effects are related to CI. In Brewer and Silver (1978), Perdue et al. (1990), laboratory experiments with no interactions between subjects are performed. In Brewer and Silver (1978), it was found that bias in favor of the in-group on a reward allocation task was unaffected by the arbitrariness of classification into groups. An effort was made to assure that subjects in the arbitrary condition would not perceive the out-group as dissimilar. They found that similarity-dissimilarity of the out-group did not affect allocation bias as long as the in-group was perceived as similar to the subject. Subjects were divided clearly into groups labeled "dark" and "light". Subjects then were asked to indicate their ratings first of "the other members of my group" and then of "the members of the other group" on a series of six-point bipolar scales (friendly-unfriendly; trustworthy-untrustworthy; cooperative-competitive; intelligent-stupid; weak-strong; generous-stingy; likeable-unlikeable). In Perdue et al. (1990), classical conditioning in-group and out-group descriptors (e.g., "us" and "them") are used to establish evaluative responses to novel, unfamiliar targets. Nonsense syllables unobtrusively paired with in-group designating pronouns (e.g., "we") were rated as more pleasant than syllables paired with out-group designators (e.g., "they").

Paris et al. (1972) study how the anticipated interaction between groups determines the representations that groups have of each other. When students are categorized into groups, discrimination occurs such that the in-group is more favorably represented than the out-group before interaction takes place and also when no interaction is anticipated. Such discrimination is stronger when competitive interaction is anticipated in an important situation. In this condition, intergroup differences are also more easily projected on physical traits. Categorization is shown to be not only an independent variable but also a dependent variable in intergroup relations.

In Kahn and Ryen (1972), Own Group Bias (OGB) was measured by differences in pre and postgame scores on the evaluative scales of the Semantic Differential (SD).

In Shank et al. (2015), an experiment on Amazon Mechanical Turk was used to develop an agent-based simulation to understand how people's motivations and behaviors within public goods dilemmas interact with the properties of the dilemma to lead to collective outcomes. They predict how the public good's benefit and size, combined with controlling individual versus group properties, produce different levels of cooperation in public goods dilemmas.

In Sethi and Somanathan (2006), a simple model of collective action is presented as a framework for empirical research into the issue of when collective action in the commons will be successful.

In van Zomeren et al. (2008), an integrative social identity model of collective action (SIMCA) is developed that incorporates three socio-psychological perspectives on collective action. Instructions for coders were to answer different questions like "Does the measure of identification (used in this study) refer to a disadvantaged group or a social movement?", "Is this group incidentally disadvantaged or structurally disadvantaged?". Coders also rated the extent to which collective disadvantage was structural on a 5-point Likert-type scale ranging from 1 (not at all) to 5 (very much).

In Salganik and Watts (2009), new insights into the role of individual behavior on collective outcomes are obtained using a multiple-worlds experimental design in a web-based experiment in which 2930 participants listened to, rated, and download 48 songs by up-and-coming bands.

In Suri and Watts (2011), laboratory experiments with interactions between subjects are performed. Web-based experiments are conducted where 24 individuals played a local public goods game arranged on one of five network topologies that varied between disconnected cliques and a random regular graphs. It was found that although players did generally behave like conditional cooperators, they were as likely to decrease their contributions in response to low contributing neighbors as they were to increase their contributions in response to high contributing neighbors. They also found that positive effects of cooperation were contagious only to direct neighbors in the network.

In Capraro (2013), online experiments using Amazon Mechanical Turk were used to develop a predictive model of human cooperation able to organize a number of different experimental findings that are not explained by the standard model.

In Rousseau and van der Veen (2005), an agent-based computer simulation of identity change explores how changes in the attributes of the individual and/or elements of the environment influence the dependent variable: the degree of shared identity in a population.

There is a host of other studies that investigate phenomena such as cooperation and a person's affinity for a group that are closely related to CI. In Worchel et al. (1977), Charness et al. (2007) laboratory experiments with interactions between subjects are performed. They study concepts such as group attraction and salience, respectively, which are related to CI. In Worchel et al. (1977), study groups worked cooperatively on two tasks and results were interpreted as showing that both previous interaction and success of combined effort are important variables in determining when intergroup cooperation will increase intergroup attraction. In Charness et al. (2007), groups perform two stage games as priming tasks, the Battle of the Sexes and Prisoner's Dilemma. Results show that the salience of the group affects behavior of members, as well as the behavior of people in another group, and that participants anticipate these effects.

A.6 Data-driven: combining experiments and data-driven modeling

This section reports on works that combine experiments with data-driven modeling. These works cover explore-exploit networked experiments with limited modeling (Mason and Watts 2012); individual models of single-choice (i.e., one-shot) evacuation decisions (Nguyen et al. 2017); ABM of emotion and information contagions spreading on a network and comparisons with a single event (Li et al. 2014); and ABM of solar panel adoption and comparisons with data in San Diego county (Zhang et al. 2016). See Zhang and Vorobeychik (2019) for a review of innovation diffusion models. None of these works use ABMs to model networked experiments where individuals take a series of actions (that may be repeated) over time, to study CI, as we do.

In Luhmann and Rajaram (2015), small-scale laboratory experiments and an ABM were used to analyze the dynamics of collaborative inhibition. In Gates et al. (2017), the model in Luhmann and Rajaram (2015) was tested against human data collected in a large-scale experiment to find that participants demonstrate non-monotonicities not evident in the predictions. These unexpected results motivate more recent work in elucidating the algorithms underlying collaborative memory. In Paxton et al. (2018), using real-time online social experiments data, a statistical model is used to study interpersonal coordination in a "minimally interactive context" to explore how people become coupled in their perceptual and memory systems while performing a task together.

In contrast to the above works, where *controlled experiments* are used to produce data that are then used for modeling, there are many models based on *observational* data. We survey some of these works here.

In Korolov et al. (2016), the possibility of predicting a social protest (planned, or unplanned) based on social media messaging is studied. In Nguyen et al. (2016), to help increase the performance of retweet prediction, a flexible model under the framework of Random Forest classifier captures a number of behavior signals affecting user's retweet decision. In Hu et al. (2014), a semantic model that can naturally represent various academic social networks, especially various complex semantic relationships among social actors, is presented. In Qin et al. (2017), the proposed method integrates topology and content of networks, and introduces a novel adaptive parameter for controlling the contribution of content with respect to the identified mismatch degree between the topological and content information. In Attema et al. (2015), data-driven multi-agent models predict Twitter trends. In van Maanen and van der Vecht (2013), a method that implements, validates, and improves an individual behavior model is proposed. The multi-agent model contains the social network structure, individual behavior parameters, and the scenario that are obtained from empirical data. In Lee and Oh (2013), emergence and propagation of reputations in social networks is modeled with a distributed algorithm. In Chierichetti et al. (2014), using several Twitter data sets, focusing in particular on the tweets sent during the soccer World Cup of 2010, a model of how users switch between producing information or sentiments

Table 14 Summary of the analyses in the Experimental Data Sect. B, and the questions we answer. Section B.1 presents histograms for the timestamps for letter requests. Section B.2 presents histograms for the timestamps for letter replies. Section B.3 presents histograms for the timestamps of the time duration between replies received and requests sent. Section B.4 presents histograms for the timestamps for words formed

Section	Histograms	Questions for analysis
B.1	Timestamps for letter request	When do players request letters during the game?
		How does the number of neighbors affect the behavior of a player to request a letter in the game?
B.2	Timestamps for letter reply	When do players reply to letter requests during the game?
		How does the number of neighbors affect the behavior of a player to reply a letter in the game?
B.3	Timestamps for time duration(reply received–request sent)	How long does it take players to reply to a letter request?
		How does the number of neighbors affects the time duration between the timestamps of the letter reply action and the letter request action?
B.4	Timestamps for word formed	When do players submit words during the game?
		How does the number of neighbors and the number of available letters affects the number of words formed by a player?

and sharing others news or sentiments is developed. In Korolov et al. (2015), a theoretical analysis is developed for how social-chatter quantitatively relates to action via a superlinear scaling law.

Other works include using data from geotagged social media messages and data from mobile health applications (Tran and Lee 2016; Kurashima et al. 2018) In Tran and Lee (2016), to understand citizen reactions regarding Ebola, a large-scale data-driven analysis of geotagged social media messages is performed. In Kurashima et al. (2018), data from mobile health applications is used to develop a statistical model, called TIPAS (Time-varying, Interdependent, and Periodic Action Sequences). This approach is based on personalized, multivariate temporal point processes that model time-varying action propensities through a mixture of Gaussian intensities. Their model captures short-term and long-term periodic interdependencies between actions through Hawkes process-based self-excitations.

Clearly, much of the modeling of observational data is motivated by social media.

B Experimental data

This Appendix describes data from the game experiments of Sect. 4. In this section we present an analysis of the experimental data that illustrates how players interact in the anagram games. We focus on experimental data that are useful in modeling. We identify four main actions $a_i \in A, 1 \leq i \leq 4$, in the set A of actions for a player during the game: (1) request letter from neighbor, (2) reply with letter to a request from a neighbor, (3) form and submit valid word, and (4) think (i.e., a no-op).

We define the following variables for the actions in the game:

- When v_i sends a requests for a *letter* to v_j, a **request sent** occurs.
- When v_j receives the *letter* request from v_i, a **request received** occurs.
- When v_j replies with the *letter* requested from v_i, a **reply sent** occurs.
- When v_i receives the *letter* reply from v_j, a **reply received** occurs.
- When v_i uses its own letters to form a word, a **word formed** occurs.

Table 14 shows a summary of the section plots and the questions we answer with the analyses.

B.1 Timestamp for letter request

The number of letters a player can request through a game depends on the number of its neighbors. Each neighbor can share up to three letters (the initial three letters), so if a player has $k = 2$ neighbors, then six letters can be requested throughout the game. If a player has $k = 8$ neighbors, then 24 letters can be requested. We want to analyze the behavior of players with reference to the letter request action and answer the following questions. When do players request letters during the game? How does the number of neighbors affects the behavior of a player to request a letter in the game?

Figure 22 shows a histogram with 10 bins of 30-s each of timestamps for **request sent**, for 47 experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used to estimate the probability density function. It indicates that more letters are being requested during the first half of the 300-second anagram game. To analyze whether the number of neighbors affects the letter request, Figure B.1 in Appendix B.1.1 from Cedeno (2019) shows

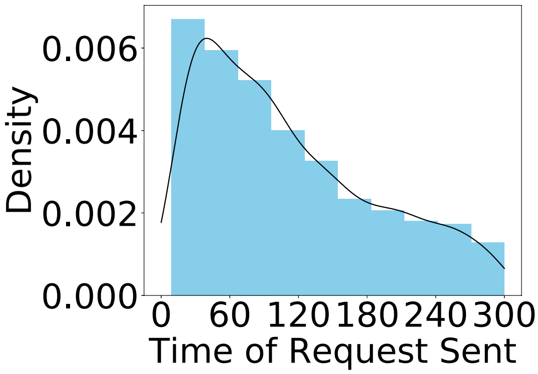

Fig. 22 Probability density distribution for time of request sent over the 300-s anagram game. Each of the bins on the *x*-axis corresponds to 30-s intervals. It shows experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used to estimate the probability density function. Letter requests are made throughout the game, rather than solely at the outset

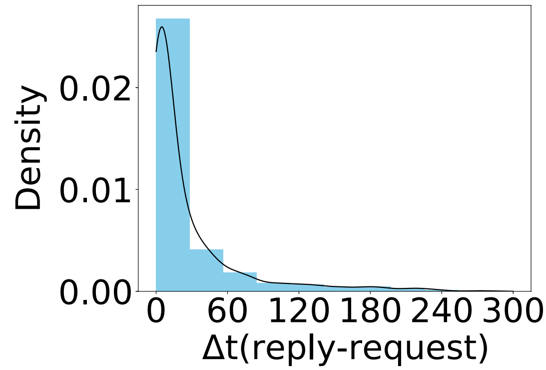

Fig. 24 Probability density distribution for time duration between requesting a letter and replying to the request, over the 300-s anagram game. Each of the bins on the x-axis corresponds to 30-s intervals. It shows experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used to estimate the probability density function. Players generally respond relatively quickly to their neighbors letter requests, with replies typically made within 30 s of the request

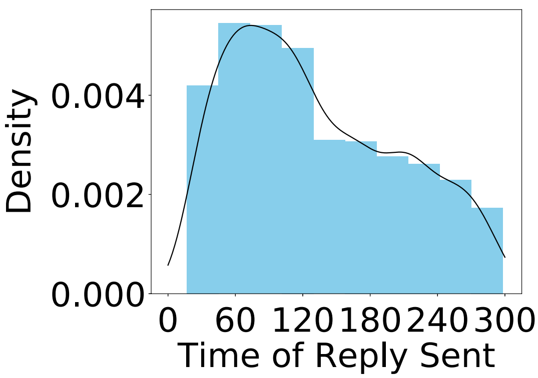

Fig. 23 Probability density distribution for time of reply sent over the 300-s anagram game. Each of the bins on the *x*-axis corresponds to 30-s intervals. It shows experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used to estimate the probability density function. Letter replies are made throughout the game, rather than solely at the outset

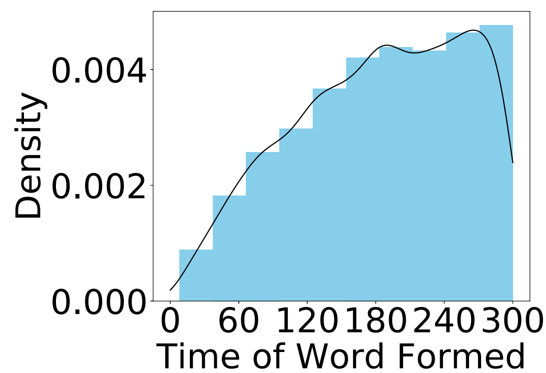

Fig. 25 Probability density distribution for time of forming words over the 300-s anagram game. Each of the bins on the *x*-axis corresponds to 30-s intervals. It shows experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used to estimate the probability density function. Word submissions are made throughout the game, and the numbers of neighbors and available letters do not affect this type of action

histograms with 10 bins of 30-s each for **request sent** for experiments with k = 2, 3, 4, 5, 6, 8. The same trends exist for each value of *k*. However, if there are few neighbors (k = 2) and consequently fewer available letters (3 letters per neighbor), there are fewer letter requests and letter replies near the end of the game.

B.2 Timestamp for letter reply sent

The number of letters a player can reply with, in response to letter requests, through a game depends on the number of its neighbors. Each neighbor can share up to 3 letters, so if a player has $k = 2$ neighbors, then 6 letters can be replied (when requested) at any time through the game, since the number of letters assigned initially is three. We want to analyze the behavior of players with reference to the letter reply action and answer the following questions. When do players reply letters during the game? How do the number of neighbors affects the behavior of a player to reply a letter in the game?

Figure 23 shows a histogram with 10 bins of 30 s each, for **reply sent**, for 47 experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used to estimate the probability density function. It indicates that letter requests are being replied to throughout the game, but more so at the earlier stages of the game. To analyze whether the number of neighbors affects the letter request, Figure

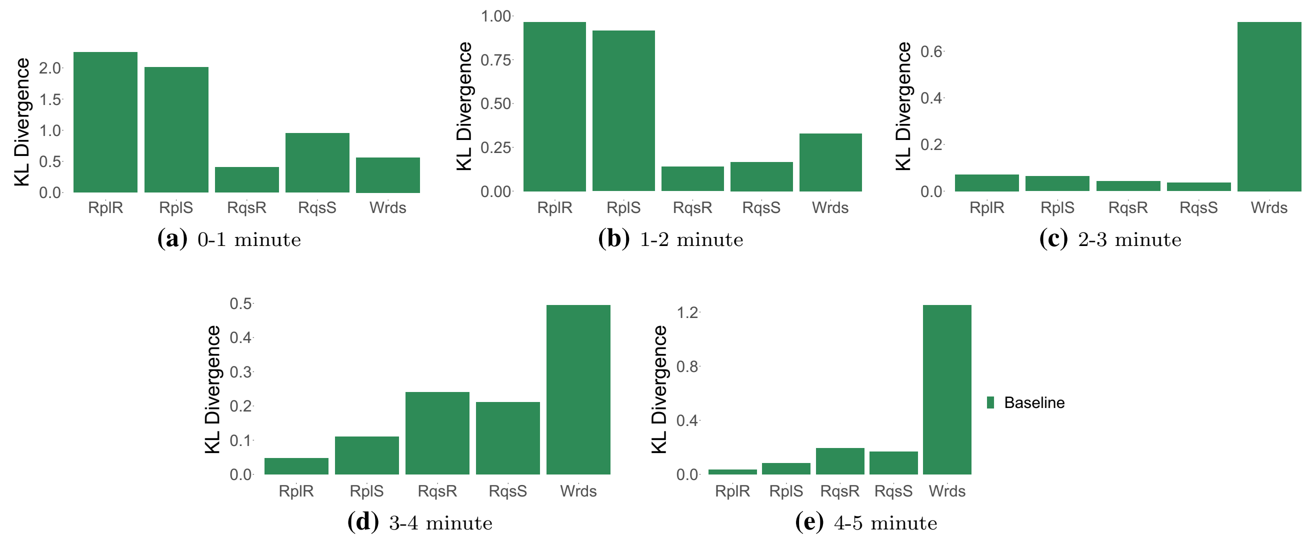

Fig. 26 Within each subfigure we show KL-divergence values for Baseline Model M0 across the five parameters of x at 1-min intervals: lower values are better. Each plot contains data over a time window: **a** 0–1 min, **b** 1–2 min, **c** 2–3 min, **d** 3–4 min and **e** 4–5 min, of the 5-min anagram game. The data are for conditions $(n = 10, k = 2)$. These plots show that request-related predictions are better than reply-related predictions over all time intervals. The reply-related predictions are better in the second half of the 5-min anagram games, but Fig. 23 shows that in experiments, there are fewer replies in the second half of the games

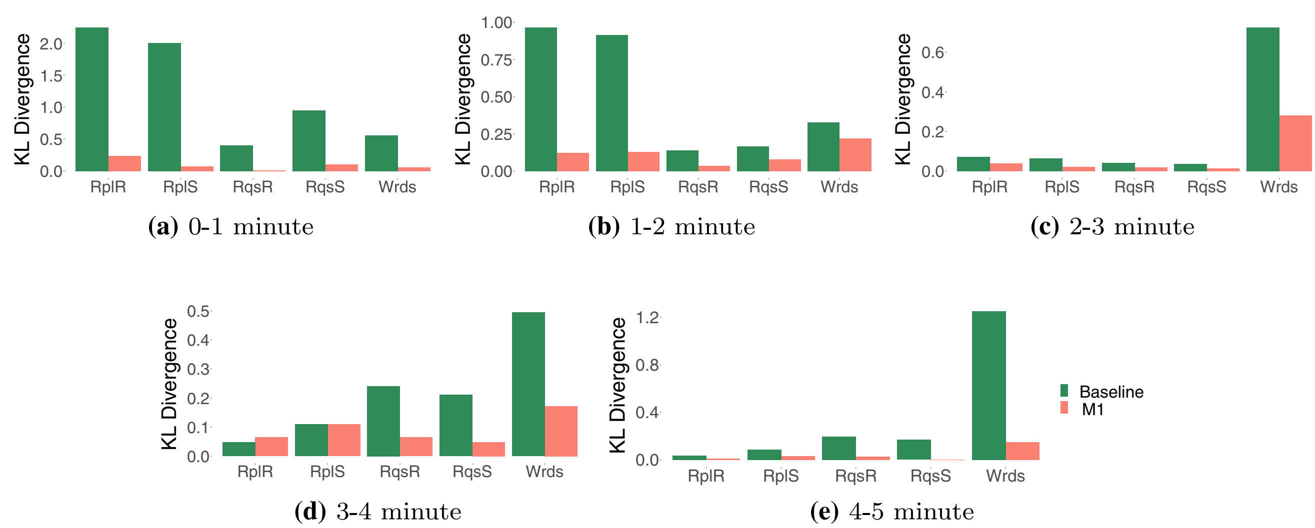

Fig. 27 Within each subfigure we show KL-divergence values for the Baseline Model M0 (in green) and Model M1 (in red) across the five parameters of x: lower values are better. The modeling conditions are experiment with $k = 2$. Each plot contains data over a time window: **a** 0–1 min, **b** 1–2 min, **c** 2–3 min, **d** 3–4 min and **e** 4–5 min, of the 5-min anagram game. While Model M0 has good predictions for minute 3 and minute 5 (with the exception of the words formed), Model M1 has better predictions for minute 3 and minute 5 for all five x variables (color figure online)

B.2 in Appendix B.1.2 in Cedeno (2019) shows histograms with 10 bins of timestamp for **reply sent** for experiments with $k = 2, 3, 4, 5, 6, 8$. Similar trends are obtained when data are broken down by k. We find that letter reply are made throughout the game, rather than solely at the outset.

B.3 Time duration from sending a letter request to receiving the requested letter

When v_i requests a letter of v_j, it has to wait for v_j to respond. Once v_j replies with the letter, then v_i is allowed to use the received letter and form words to contribute to the team. This time duration between request sent and reply received reveals how long players take to reply to their neighbors' requests. A player only has to request a letter (and receive it) on one occasion to use it as any number of times in forming words. Remember that these rules were designed to foster word construction, to increase earnings potential, and to foster team cohesion. We want to analyze the behavior of players with reference to the time duration between the timestamps of the letter reply action and the letter request action, to answer the following questions. How long does it take for players to reply to a letter request? How does the number of neighbors affect the difference between the timestamps of the letter reply action and the letter request action?

Figure 24 shows a histogram with 10 bins of 30-s each, for the time difference between **reply received** and **request sent**, for 47 experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used estimate the probability density function. Players generally respond relatively quickly to their neighbors letter requests with replies typically made within 30 s of the request.

To analyze whether this behavior is common while increasing the number of k neighbors in a game, Figure B.3 in Appendix B.1.3 in Cedeno (2019) shows histograms with 10 bins of 30-s each of timestamp change between reply received and request sent for experiments with $k = 2, 3, 4, 5, 6, 8$. The number of neighbors doesn't affect this type of action, players generally respond relatively quickly to their neighbors letter requests with replies typically made within 30 s of the request.

B.4 Timestamp for word formed

At any time during a game, a player can form a word and submit it for validation to our web application. If a player possesses letters to form a valid word, then she forms and submits a word, the application validates it, and the word is added to the game screen. We want to analyze the behavior of players with reference to the action of word formed and answer the following questions. When do players submit words during the game? How does the number of neighbors and the number of available letters affects the number of words formed by a player?

Figure 25 shows a histogram with 10 bins of 30-s each for **word formed**, for 47 experiments with $k = 2, 3, 4, 5, 6, 8$. A kernel-density estimation with Gaussian kernels is used estimate the probability density function. It suggests that words are being formed throughout the game, and even up through the end of the game. This justifies a 5-min anagram game duration. To analyze whether the number of neighbors affects the word formation, Figure B.4 in Appendix B.1.4 in Cedeno (2019) shows histograms with 10 bins of 30-s each for timestamp of **word formed** for experiments with $k = 2, 3, 4, 5, 6, 8$. Word submissions are made throughout the game, and the number of neighbors and available letters does not affect this behavior.

C Modeling supplement

This Appendix contains several figures that support the modeling of Sect. 5. See Figs. 26 and 27.

References

Abbott A (1995) Sequence analysis: new methods for old ideas. Ann Rev Sociol 21:93–113

Abrams DE, Hogg MA (1990) Social identity theory: constructive and critical advances. Springer, New York

Ackland R, O'Neil M (2011) Online collective identity: the case of the environmental movement. Soc Netw 33:177–190

Agrawal R, Srikant R (1995) Mining sequential patterns. In: Proceedings of the eleventh international conference on data engineering, IEEE Computer Society, ICDE'95, Washington, pp 3–14

Aipperspach R, Cohen E, Canny J (2006) Modeling human behavior from simple sensors in the home. In: Proceedings of the IEEE conference on pervasive computing

Alexander JC, Eyerman R, Giesen B, Smelser NJ, Sztompka P (2004) Cultural trauma and collective identity, 1st edn. University of California Press, Berkeley

Andrews EAM, Bonner AJ (2011) Explaining genetic knock-out effects using cost-based abduction. In: Proceedings of the 22nd international joint conference on artificial intelligence (IJCAI), pp 1635–1640

Atran S, Axelrod R, Davis R (2007) Sacred barriers to conflict resolution. Science 317:1039–1040

Atran S, Sheikh H, Gomez A (2014a) Devoted actors sacrifice for close comrades and sacred cause. Proc Natl Acad Sci 111(50):17702–17703

Atran S, Sheikh H, Gomez A (2014b) For cause and comrade: devoted actors and willingness to fight. Cyclodynamics 5:41–57

Attema T, van Maanen PP, Meeuwissen E (2015) Development and evaluation of multi-agent models predicting twitter trends in multiple domains. In: 2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp 1133–1140

Bach A (1990) Boltzmann's probability distribution of 1877. Arch Hist Exact Sci 41:1–40

Becker GS (1974) A theory of social interaction. J Political Econ 82:1063–1093

Benjamin DJ, Choi JJ, Fisher G (2016) Religious identity and economic behavior. Rev Econ Stat 98(4):617–637

Bergmann JHM, Langdon PM, Mayagoitia RE, Howard N (2014) Exploring the use of sensors to measure behavioral interactions: An experimental evaluation of using hand trajectories. PLoS One 9(2):1–10

Bornstein G, Yaniv I (1998) Individual and group behavior in the ultimatum game: Are groups more rational players? Exp Econ 1:101–108

Brewer MB (1991) The social self: on being the same and different at the same time. Personal Soc Psychol Bull 17:475–482

Brewer MB, Gardner W (1996) Who is this "we"? Levels of collective identity and self representations. J Personal Soc Psychol 71:83–93

Brewer MB, Silver M (1978) Ingroup bias as a function of task characteristics. Eur J Soc Psychol 8(3):393–400

Brody H (2000) The other side of Eden: hunters, farmers, and the shaping of the world. Farrar, Straus and Giroux, New York

Brunsdon AR (2017) #MisconstruedidentitiesMustFall collective: identity formation in the current South African context: a practical theological perspective. HTS Theol Stud 73:1–7
Butler J (1988) Performative acts and gender constitution: an essay in phenomenology and feminist theory. Theatre J 40(4):519–531
Cadsby CB, Song F, Tapon F (2007) Sorting and incentive effects of pay for performance: an experimental investigation. Acad Manag J 50:387–405
Cadsby CB, Song F, Tapon F (2010) Are you paying your employees to cheat? An experimental investigation. BE J Econ Anal Pol 10:1–30
Cameron D (1997) Performing gender identity: Young men's talk and the construction of heterosexual masculinity. In: Johnson S, Hanna U (eds) Language and masculinity. Blackwell, Oxford, pp 47–64
Capraro V (2013) A model of human cooperation in social dilemmas. PLoS One 8:e72427–1–e72427–6
Cedeno VI (2019) Pipelines for computational social science experiments and model building. Ph.D. dissertation, Virginia Polytechnic Institute and State University, https://vtechworks.lib.vt.edu/bitstream/handle/10919/91445/Cedeno_VI_D_2019.pdf. Accessed 8 Jan 2019
Chamiak E, Santos E (1992) Dynamic map calculations for abduction. In: Proceedings AAAI conference on artificial intelligence, pp 552–557
Charness G, Rigotti L, Rustichini A (2007) Individual behavior and group membership. Am Econ Rev 97:1340–1352
Charness G, Cobo-Reyes R, Jimenez N (2014) Identities, selection, and contributions in a public-goods game. Games Econ Behav 87:322–338
Chen DL, Yeh S (2014) The construction of morals. J Econ Behav Organ 104:84–105
Chen R, Chen Y (2011) The potential of social identity for equilibrium selection. Am Econ Rev 101(6):2562–89
Chen Y, Li S (2009) Group identity and social preferences. Am Econ Rev 99:431–457
Chen Y, Li SX, Liu TX, Shih M (2014) Which hat to wear? Impact of natural identities on coordination and cooperation. Games Econ Behav 84:58–86
Chen DL, Schonger M, Wickens C (2016) otree-an open-source platform for laboratory, online and field experiments. J Behav Exp Financ 9:88–97
Chierichetti F, Kleinberg J, Kumar R, Mahdian M, Pandey S (2014) Event detection via communication pattern analysis. In: Proceedings of the 8th international conference on weblogs and social media, ICWSM, pp 51–60
Choup AM (2008) The formation and manipulation of collective identity: a framework for analysis. Soc Mov Stud 7(2):191–207
Cohn A, Fehr E, Maréchal MA (2014) Business culture and dishonesty in the banking industry. Nature 516:86–89
Cohn A, Marechal MA, Noll T (2015) Bad boys: how criminal identity salience affects rule violation. Rev Econ Stud 82(4):1289–1308
Cover TM, Thomas J (1991) Elements of information theory. Wiley, Hoboken
Contractor N (2019) How can computational social science motivate the development of theories, data, and methods to advance our understanding of communication and organizational dynamics? In: Foucault Welles B, González-Bailón S (eds) The Oxford Handbook of Networked Communication. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190460518.013.7
Currarini S, Jackson M, Pin P (2009) An economic model of friendship: homophily, minorities and segregation. Econometrica 77:1003–1045
Davis WL, Davis DE (1972) Internal-external control and attribution of responsibility for success and failure. J Personal 40:123–136
de Oliveira M (1992) Isotropic majority-vote model on a square lattice. J Stat Phys 66:273–281
DeChurch LA, Mesmer-Magnus JR (2010) The cognitive underpinnings of effective teamwork: a meta-analysis. J Appl Psychol 95(1):32–53
Dominowski RL (1969) The effect of pronunciation practice on anagram difficulty. Psychon Sci 16(2):99–100
Drouvelis M, Metcalfe R, Powdthavee N (2010) Priming cooperation in social dilemma games. In: IZA Discussion Papers 4963, Institute for the Study of Labor (IZA)
Durkheim E (1951) Suicide. Free Press, Mumbai
Echenim M, Peltier N, Tourret S (2013) An approach to abductive reasoning in equational logic. In: Proceedings of the 23rd international joint conference on artificial intelligence (IJCAI), pp 531–537
Eckel C, Grossman P (2005) Managing diversity by creating team identity. J Econ Behav Organ 58:371–392
Epstein JM (2007) Generative social science: studies in agent-based computational modeling. Princeton University Press, Princeton
Eriksen TH (2010) Ethnicity and nationalism: anthropological perspectives, 3rd edn. Pluto Press, London
Erikson EH (1980) Identity and the life cycle. W. W. Norton & Company, New York
Feather NT (1969) Attribution of responsibility and valence of success and failure in relation to initial confidence and task performance. J Personal Soc Psychol 13:129–144
Feather NT, Simon JG (1971a) Attribution of responsibility and valence of outcome in relation to initial confidence and success and failure of self and other. J Personal Soc Psychol 18:173–188
Feather NT, Simon JG (1971b) Causal attributions for success and failure in relation to expectation of success based upon selective or manipulative control. J Personal Soc Psychol 39:527–541
Feher O, Wonnacott E, Smith K (2016) Structural priming in artificial languages and the regularisation of unpredictable variation. J Mem Lang 91:158–180
Fiske ST, Gilbert DT, Lindzey G (2010) Handbook of social psychology, 5th edn. Wiley, Hoboken
Flach PA, Kakas AC (2010) Abduction and induction: essays on their relation and integration. Springer, New York
Fominaya CF (2010) Collective identity in social movements: central concepts and debates. Sociol Compass 4:393–404
Gates MA, Suchow JW, Griffiths TL (2017) Empirical tests of large-scale collaborative recall. In: CogSci
Gibbs AL, Su FE (2002) On choosing and bounding probability metrics. Int Stat Rev 70:419–435
Gilhooly KJ, Johnson CE (1978) Effects of solution word attributes on anagram difficulty: a regression analysis. Q J Exp Psychol 30(1):57–70
Gilwhite FJ (2001) Are ethnic groups biological species to the human brain? Essentialism in our cognition of some social categories. Curr Anthropol 42(4):515–553
Ginges J, Atran S (2013) Sacred values and cultural conflict. In: Advances in culture and psychology, pp 273–305
Goldberg CA (2003) Haunted by the specter of communism: collective identity and resource mobilization in the demise of the Workers Alliance of America. Theory Soc 32(5/6):725–773
Goldman M, Stockbauer JW, McAuliffe TG (1977) Intergroup and intragroup competition and cooperation. J Exp Soc Psychol 13:81–88
Gomez A, Brooks ML, Buhrmester MD, Vazquez A, Jetten J, Swann WB (2011a) On the nature of identity fusion: Insights into the construct and a new measure. J Personal Soc Psychol 100:918–933
Gomez A, Morales J, Hart S, Vazquez A, Jetten J, Swann WB (2011b) Rejected and excluded forevermore, but even more devoted:

irrevocable ostracism intensifies loyalty to the group among identity-fused persons. Personal Soc Psychol Bull 37:1574–1586
Granovetter M (1978) Threshold models of collective behavior. Am J Sociol 83(6):1420–1443
Greene B (2000) Letters to a young poet. New World Library, Novato
Greenhill B (2008) Recognition and collective identity formation in international politics. Eur J Int Relat 14(2):343–368
Guralnik V, Haigh KZ (2002) Learning models of human behaviour with sequential patterns. In: Proceedings of the AAAI-02 workshop “Automation as Caregiver”, aAAI Technical Report WS-02-02, pp 24–30
Haig BD (2005) An abductive theory of scientific method. Psychol Methods 10:371–388
Hand DJ, Smyth P, Mannila H (2001) Principles of data mining. MIT Press, Cambridge
Hawkes AG (1971) Spectra of some self-exciting and mutually exciting point processes. Biometrika 58(1):83–90
Heller N (2019) The philosopher redefining equality: Elizabeth Anderson thinks we’ve misunderstood the basis of a free and fair society. The New Yorker 07 January. https://www.newyorker.com/magazine/2019/01/07/the-philosopher-redefining-equality. Accessed 2 Feb 2019
Hoff K, Pandey P (2006) Discrimination, social identity, and durable inequalities. Am Econ Rev 96(2):206–211
Hoff K, Pandey P (2014) Making up people-the effect of identity on performance in a modernizing society. J Dev Econ 106(C):118–131
Hogg MA, Abrams D (2007) Intergroup behavior and social identity. In: Hogg MA, Cooper J (eds) The sage handbook of social psychology. Sage, New York, pp 335–360
Homans G (1961) Social behavior: its elementary forms. Harcourt Brace, San Diego
Hu J, Liu M, Zhang J (2014) A semantic model for academic social network analysis. In: 2014 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM 2014), pp 310–313
Jiménez J, Gomez A, Buhrmester MD et al (2016) The dynamic identity fusion index: a new continuous measure of identity fusion for web-based questionnaires. Soc Sci Comput Rev 34:215–228
Juba B (2016) Learning abductive reasoning using random examples. In: Proceedings of the thirtieth AAAI conference on artificial intelligence
Judd S, Kearns M, Vorobeychik Y (2010) Behavioral dynamics and influence in networked coloring and consensus. PNAS 107:14978–14982
Juergensmeyer M (2003) Terror in the mind of god: the global rise of religious violence, 3rd edn. University of California Press, Berkeley
Kahn A, Ryen AH (1972) Factors influencing the bias towards one’s own group. Int J Group Tens 2:33–50
Kearns M, Judd S, Tan J, Wortman J (2009) Behavioral experiments on biased voting in networks. PNAS 106:1347–1352
Kearns M, Judd S, Vorobeychik Y (2012) Behavioral experiments on a network formation game. In: Economics and computation (EC), pp 690–704
Kendler H, Kendler ST (1962) Vertical and horizontal processes in problem solving. Psychol Rev 69:1–16
Kinnebrew JS, Loretz KM, Biswas G (2013) A contextualized, differential sequence mining method to derive students’ learning behavior patterns. J Educ Data Min 5(1):190–219
Korolov R, Lu D, Wang J, Zhou G, Bonial C, Voss C, Kaplan L, Wallace W, Han J, Ji H (2016) On predicting social unrest using social media. In: Proceedings of the 2016 IEEE/ACM international conference on advances in social networks analysis and mining, ASONAM 2016, Institute of Electrical and Electronics Engineers Inc., pp 89–95
Korolov R, Peabody J, Lavoie A, Das S, Magdon-Ismail M, Wallace W (2015) Actions are louder than words in social media. In: 2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp 292–297
Kozegi B (2006) Ego utility, overconfidence, and task choice. J Eur Econ Assoc 4:673–707
Kozlowski SW, Ilgen DR (2006) Enhancing the effectiveness of work groups and teams. Psychol Sci Publ Interest 7(3):77–124
Kullback S (1959) Information theory and statistics. Wiley, New York
Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22(1):79–86
Kurashima T, Althoff T, Leskovec J (2018) Modeling interdependent and periodic real-world action sequences. In: World Wide Web Conference WWW, ACM, pp 803–812
Latham GP, Locke EA (1991) Self-regulation through goal setting. Organ Behav Hum Decisi Process 50:212–247
Ledyard JO (1994) Public goods: a survey of experimental research. Public economics, University Library of Munich, Germany
Lee JY, Oh JC (2013) A model for recursive propagations of reputations in social networks. In: Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining, ASONAM ’13, ACM, New York, pp 666–670
Li B, Sun D, Lin Z, Ou C (2014) Agent-based simulation research on group emotion evolution of public emergency. In: ASONAM
Locke EA, Latham GP (1990) A theory of goal setting and task performance. Prentice-Hall, Englewood Cliffs
Luhmann CC, Rajaram S (2015) Memory transmission in small groups and large networks: an agent-based model. Psychol Sci 26(12):1909–17
Lustick I (2000) Agent-based modelling of collective identity: testing constructivist theory. Journal of Artificial Societies and Social Simulation, vol 3. no. 1. http://jasss.soc.surrey.ac.uk/JASSS.html
MacGregor N (2018) Living with the gods: on beliefs and peoples, 1st edn. Knopf, New York
Manchester W (1993) A world lit only by fire: the medieval mind and the renaissance: portrait of an age. Little, Brown and Company, Boston
Mason W, Watts DJ (2012) Collaborative learning in networks. Proc Natl Acad Sci 109(3):764–769
Mayzner MS, Tresselt ME (1958) Anagram solution times: a function of letter order and word frequency. J Exp Psychol 56(4):376
McAuliffe K, Dunham Y (2015) Group bias in cooperative norm enforcement. Philos Trans R Soc B 371:20150073-1–20150073-9
McFarland DA, Moody J, Diehl D, Smith JA, Thomas RJ (2014) Network ecology and adolescent social structure. Am Sociol Rev 79:1–34
Melucci A (1989) Nomads of the present: social movement and identity needs in contemporary society. Temple University Press, Philadelphia
Melucci A (1995) The process of collective identity. In: Klandermans B, Johnston H (eds) Social movements and culture. University of Minnesota Press, Minneapolis, pp 104–130
Miller DT, Ross M (1975) Self-serving biases in the attribution of causality: fact or fiction? Psychol Bull 82:213–225
Muller W (1996) How then, shall we live? Four simple questions that reveal the beauty and meaning of our lives. Bantam Books, New York
Nagel J (1996) American Indian ethnic renewal: red power and the resurgence of identity and culture. Oxford University Press, New York
Nguyen DA, Tan S, Ramanathan R, Yan X (2016) Analyzing information sharing strategies of users in online social networks. In: 2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp 247–254

Nguyen C, Schlesinger KJ, Carlson JM (2017) Data-driven models for individual and group decision making. In: ASONAM, pp 852–859
Oldenquist A (1982) Loyalties. J Philos 79:173–193
Olson M (1965) The logic of collective action: public goods and the theory of groups. Harvard Univ. Press, Cambridge
Owens TJ (2006) Self and identity. In: Delamater J (ed) Handbook of social psychology. Springer, New York, pp 205–232
Paris WD, Budapest GC, Konstanz HDD, Bristol CG, Oregon KL, Stirling AO (1972) An experimental investigation into the formation of intergroup representations. Eur J Soc Psychol 2(2):202–204
Paxton A, Morgan TJH, Suchow JW, Griffiths T (2018) Interpersonal coordination of perception and memoryin real-time online social experiments. In: Proceedings of the 40th annual meeting of the cognitive science society, Austin
Peek L (2005) Becoming Muslim: the development of a religious identity. Sociol Relig 66(3):215–242
Perdue C, Dovidio JF, Gurtman MB, Tyler RB (1990) Us and them: social categorization and the process of intergroup bias. J Personal Soc Psychol 59:475–486
Pereira L, Moreira F (2005) Majority-vote model on random graphs. Phys Rev E Stat Nonlinear Soft Matter Phys 71:016123
Pfandler A, Rümmele S, Szeider S (2013) Backdoors to abduction. In: Proceedings of the 23rd international joint conference on artificial intelligence (IJCAI), pp 1046–1052
Pierce CS (1931) Elements of logic. In: Hartshorn C et al (eds) Collected papers of Charles sanders pierce. Harvard University Press, Harvard
Pilny A, Poole MS, Reichelmann A, Klein B (2017) A structurational group decision-making perspective on the commons dilemma: results from an online public goods game. J Appl Commun Res 45(4):413–428
Plutzer E, Zipp J (1996) Identity politics, partisanship, and voting for women candidates. Publ Opin Q 60(1):30–57. https://doi.org/10.1086/297738
Polletta F, Jasper JM (2001) Collective identity and social movements. Ann Rev Sociol 27:283–305
Qin M, Jin D, He D, Gabrys B, Musial K (2017) Adaptive community detection incorporating topology and content in social networks. In: Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017, ASONAM ’17, ACM, New York, pp 675–682
Ren Y, Cedeno-Mieles V, Hu Z, Deng X, Adiga A, Barrett C, Ekanayake S, Goode BJ, Korkmaz G, Kuhlman CJ, Machi D, Marathe MV, Ramakrishnan N, Ravi SS, Saraf P, Self N (2018) Generative modeling of human behavior and social interactions using abductive analysis. In: 2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp 413–420
Rousseau D, van der Veen AM (2005) The emergence of a shared identity: an agent-based computer simulation of idea diffusion. J Confl Resolut 49(5):686–712
Russell DG, Sarason IG (1965) Test anxiety, sex, and experimental conditions in relation to anagram solution. J Personal Soc Psychol 1:493–496
Salganik MJ, Watts DJ (2009) Web-based experiments for the study of collective social dynamics in cultural markets. topiCS 1(3):439–468
Sarason IG (1973a) Test anxiety and cognitive modeling. J Personal Soc Psychol 28:58–61
Sarason IG (1973b) Test anxiety and social influence. J Personal 41:261–271
Schelling TC (2006) Micromotives and macrobehavio, revised edn. W. W. Norton & Company, New York
Schubert TW, Otten S (2002) Overlap of self, ingroup, and outgroup: pictorial measures of self-categorization. Self Identity 1(4):353–376
Schweitzer M, Ordonez L, Dumaz B (2004) Goal-setting as a motivator of unethical behavior. Acad Manag J 47:422–433
Sethi R, Somanathan E (2006) A simple model of collective action. Econ Dev Cultural Change 54(3):725–747
Shanahan M (2005) Perception as abduction: turning sensor data into meaningful representation. Cognit Sci 29:103–134
Shank DB, Kashima Y, Saber S, Gale T, Kirley M (2015) Dilemma of dilemmas: how collective and individual perspectives can clarify the size dilemma in voluntary linear public goods dilemmas. PLoS One 10:1–19
Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27(3):379–423
Silke A (2008) Holy warriors: exploring the psychological processes of jihadi radicalization. Eur J Criminol 5(1):99–123
Singla P, Mooney RJ (2011) Abductive markov logic for plan recognition. In: Proceedings of the twenty-fifth AAAI conference on artificial intelligence
Smith K, Feher O, Culbertson J (2017) The influence of word-order harmony on structural priming in artificial languages. In: Proceedings of the 39th annual meeting of the cognitive science society, CogSci 2017, London, 16–29 July 2017
Snow D (2001) Collective identity and expressive forms. In: Smelser NJ, Baltes PB (eds) International encyclopedia of the social and behavioral sciences. Elsevier, Amsterdam, pp 2212–2219
Snow D, McAdams D (2000) Identity work processes in the context of social movements: clarifying the identity/movement nexus. In: Stryker S, Owens T, White RW (eds) Self, identity and social movements. University of Minneapolis Press, Minneapolis, pp 2212–2219
Stones CR (1983) Self-determination and attribution of responsibility: another look. Psychol Rep 53:391–394
Stout CT, Kretschmer K, Ruppanner L (2017) Gender linked fate, race/ethnicity, and the marriage gap in American politics. Polit Res Q 70(3):509–522
Stryker S (1980) Symbolic interactionism: a social structural version. Benjamin/Cummings, San Francisco
Suri S, Watts DJ (2011) Cooperation and contagion in web-based, networked public goods experiments. PLoS One 6:1–18
Sutton C, McCallum A (2011) An introduction to conditional random fields. Found Trends Mach Learn 4(4):267–373
Swann WB, Gomez A, Seyle DC, Morales JF, Huici C (2009) Identity fusion: the interplay of personal and social identities in extreme group behavior. J Personal Soc Psychol 96:955–1011
Swann WB, Gomez A, Dovidio JF, Hart S, Jetten J (2010a) Dying and killing for one’s group: Identity fusion moderates responses to intergroup versions of the trolley problem. Psychol Sci 21:1176–1183
Swann WB, Gomez A, Huici C, Morales JF (2010b) Identity fusion and self-sacrifice: Arousal as a catalyst of pro-group fighting, dying, and helping behavior. J Personal Soc Psychol 99:824–841
Swann WB, Gomez A, Buhrmester MD, Lopez-Rodriguez L, Jimenez J, Vazquez A (2014) Contemplating the ultimate sacrifice: identity fusion channels pro-group affect, cognition, and moral decision making. J Personal Soc Psychol 106:713–727
Swanson MS (2015) Composing democracy: Collective identity formation in small group composition. Ph.D. thesis, University of Washington, Washington
Tajfel H (1974) Social identity and intergroup behavior. Soc Sci Inf 13:65–93
Tanaka Y, Iwata T, Kurashima T, Toda H, Ueda N (2018) Estimating latent people flow without tracking individuals. In: Proceedings of the twenty-seventh international joint conference on artificial

intelligence, IJCAI-18, international joint conferences on artificial intelligence organization, pp 3556–3563
Tarrow S (2010) Mentalities, political cultures, and collective action frames: constructing meanings through action. Sociol Compass 4:174–202
Tatum BD (2003) Why are all the black kids sitting together in the cafeteria? And other conversations about race, 2nd edn. Basic Books, New York
Taylor V, Whittier NE (1992) Collective identity in social movement communities: lesbian feminist mobilization. In: Morris AD, Mueller CM (eds) Frontiers of social movement theory. Yale University Press, New Haven, pp 104–130
Thagard P (1989) Explanatory coherence. Behav Brain Sci 12:435–502
Timmermans S (1999) Social death as a self-fulfilling prophecy: David Sudnow's 'passing on' revisited. Sociol Q 39:453–472
Timmermans S, Tavory I (2012) Theory construction in qualitative research: from grounded theory to abductive analysis. Sociol Theory 30:167–186
Tran T, Lee K (2016) Understanding citizen reactions and ebola-related information propagation on social media. In: Proceedings of the 2016 IEEE/ACM international conference on advances in social networks analysis and mining, ASONAM '16, IEEE Press, Piscataway, pp 106–111
Tresselt ME (1968) Reexamination of anagram problem solving. Trans N Y Acad Sci 30:1112–1119
van Dijk TA (2000) Ideology: a multidisciplinary approach. Sage, Thousand Oaks
van Maanen PP, van der Vecht B (2013) An agent-based approach to modeling online social influence. In: Proceedings of the 2013 IEEE/ACM international conference on advances in social networks analysis and mining, ASONAM '13, ACM, New York, pp 600–607
van Zomeren M, Postmes T, Spears R (2008) Toward an integrative social identity model of collective action: a quantitative research synthesis of three socio-psychological perspectives. Psychol Bull 134(4):504
Vance RJ, Colella A (1990) Effects of two types of feedback on goal acceptance. J Appl Psychol 75:68–77
Vanderhaegen F, Caulier P (2011) A multi-viewpoint system to support abductive reasoning. Inf Sci 181:5349–5363
Verkuyten M, Yildiz AA (2007) National (dis)identification and ethnic and religious identity: a study among Turkish-Dutch Muslims. Personal Soc Psychol Bull 33(10):1448–1462
Vryan KD, Adler EA, Adler P (2003) Identity. In: Reynolds LT, Herman-Kinney NJ (eds) Handbook of symbolic interactionism. Altamira Press, Lanham, pp 367–390
Warren MW, Thomson WJ (1969) Anagram solution as a function of transition probabilities and solution word frequency. Psychon Sci 17(6):333–334
Wei-Kleiner F, Dragisic Z, Lambrix P (2014) Abduction framework for repairing incomplete el ontologies: complexity results and algorithms (extended version). In: Proceedings of the 28th AAAI conference on artificial intelligence
Wendt A (1994) Collective identity formation and the international state. Am Polit Sci Rev 88(2):384–396
Worchel S, Andreoli VA, Folger R (1977) Intergroup cooperation and intergroup attraction: the effect of previous interaction and outcome of combined effort. J Exp Soc Psychol 13:131–140
Zhang H, Vorobeychik Y, Letchford J, Lakkaraju K (2016) Data-driven agent-based modeling, with application to rooftop solar adoption. Auton Agents Multi-Agent Syst 30:1023–1049
Zhang H, Vorobeychik Y (2019) Empirically grounded agent-based models of innovation diffusion: a critical review. Artif Intell Rev 52:707–741

Affiliations

Vanessa Cedeno-Mieles[2,4,6] · Zhihao Hu[5] · Yihui Ren[7] · Xinwei Deng[5] · Abhijin Adiga[1] · Christopher Barrett[1,3] · Noshir Contractor[9] · Saliya Ekanayake[12] · Joshua M. Epstein[10] · Brian J. Goode[2] · Gizem Korkmaz[1] · Chris J. Kuhlman[1] · Dustin Machi[1] · Michael W. Macy[11] · Madhav V. Marathe[1,3] · Naren Ramakrishnan[4,8] · S. S. Ravi[1] · Parang Saraf[8] · Nathan Self[8]

Zhihao Hu
huzhihao@vt.edu

Yihui Ren
yihui.ren@gmail.com

Xinwei Deng
xdeng@vt.edu

Abhijin Adiga
abhijin@virginia.edu

Christopher Barrett
cbarrett@virginia.edu

Noshir Contractor
ncontractor@gmail.com

Saliya Ekanayake
esaliya@lbl.gov

Joshua M. Epstein
je65@nyu.edu

Brian J. Goode
bgoode@vt.edu

Gizem Korkmaz
gkorkmaz@virginia.edu

Chris J. Kuhlman
cjk8gx@virginia.edu

Dustin Machi
dmachi@virginia.edu

Michael W. Macy
mwm14@cornell.edu

Madhav V. Marathe
mmarathe@virginia.edu

Naren Ramakrishnan
naren@cs.vt.edu

S. S. Ravi
ssravi0@gmail.com

Parang Saraf
parang@cs.vt.edu

Nathan Self
nwself@vt.edu

[1] Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, VA, USA

[2] Biocomplexity Institute, Virginia Tech, Blacksburg, VA, USA

[3] Department of Computer Science, University of Virginia, Charlottesville, VA, USA

[4] Department of Computer Science, Virginia Tech, Blacksburg, VA, USA

[5] Department of Statistics, Virginia Tech, Blacksburg, VA, USA

[6] Escuela Superior Politécnica del Litoral, ESPOL, Guayaquil, Ecuador

[7] Computational Science Initiative, Brookhaven National Laboratory, Upton, NY, USA

[8] Discovery Analytics Center, Virginia Tech, Blacksburg, VA, USA

[9] Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, USA

[10] Department of Epidemiology, New York University, New York, NY, USA

[11] Department of Sociology, Cornell University, Ithaca, NY, USA

[12] Lawrence Berkeley National Laboratory, Berkeley, CA, USA