2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Generative Modeling of Human Behavior and Social Interactions Using Abductive Analysis

Yihui Ren*, Vanessa Cedeno-Mieles* † §, Zhihao Hu‡, Xinwei Deng‡,
Abhijin Adiga*, Christopher Barrett*, Saliya Ekanayake*, Brian J. Goode*, Gizem Korkmaz*, Chris J. Kuhlman*,
Dustin Machi*, Madhav V. Marathe*†, Naren Ramakrishnan¶†, S. S. Ravi*, Parang Saraf¶†, Nathan Self¶†
Biocomplexity Institute*, Discovery Analytics Center¶, Departments of Computer Science† and Statistics‡
Virginia Tech - Blacksburg, VA*
Escuela Superior Politécnica del Litoral, ESPOL - Guayaquil, Ecuador§

Noshir Contractor
Department of Industrial Engineering
and Management Sciences
Northwestern University
Evanston, IL

Joshua Epstein
College of Global Public Health
New York University
New York, NY

Michael W. Macy
Department of Sociology
Cornell University
Ithaca, NY

***Abstract*—Abduction is an inference approach that uses data and observations to identify plausible (and preferably, best) explanations for phenomena. Applications of abduction (e.g., robotics, genetics, image understanding) have largely been devoid of human behavior. Here, we devise and execute an iterative abductive analysis process that is driven by the social sciences: behaviors and interactions among groups of human subjects. One goal is to understand intra-group cooperation and its effect on fostering collective identity. We build an online game platform; perform and analyze controlled laboratory experiments; form hypotheses; build, exercise, and evaluate network-based agent-based models; and evaluate the hypotheses in multiple abductive iterations, improving our understanding as the process unfolds. While the experimental results are of interest, the paper's thrust is methodological, and indeed establishes the potential of iterative abductive looping for the (computational) social sciences.**

I. Introduction

A. Background and Motivation

Abduction is an inference approach that uses data and observations to identify plausible (preferably, best) explanations for phenomena [?]. Abduction has broad application in robotics, genetics, automated systems, and image understanding [?], [?], [?], [?].

However, in contrast to this 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. As a differentiator from previous work, our interests are behaviors and human interactions within networked groups in the social sciences.

In particular, our exemplar is to understand whether a cooperative game can *produce* collective identity (CI) within a group. CI is an individual's cognitive, moral, and emotional connection with a broader community, category, practice, or institution [?]. There are many applications and contexts in which CI is important, including team formation, maintenance, and behavior in organizations, communities, and marginalized groups (e.g., [?], [?], [?]).

Inspired by [?], we use a group anagrams (word construction) game to engender CI, where players work cooperatively to form words by sharing letters, and use the Dynamic Identity Fusion Index (DIFI) score [?] as our measurement of CI. Identity fusion is a feeling of oneness with the group that induces people to tether their feelings of personal agency to the group [?]. Our first novelty is that this is the first work on performing and analyzing *online* experiments, and developing and evaluating agent-based models (ABMs), for the group anagrams game and CI. (The first *face-to-face* experiments were conducted recently [?], with no modeling work done; their setup is somewhat different than ours.) Our experimental findings below constitute our second novelty.

The abductive loop (AL) process is described in Section II, but among its components are experiments and modeling, and we make note of experiment-modeling interactions here. There have been several controlled experimental studies of comparable size to our experiments (e.g., [?], [?], [?]). Also, empirically grounded, data-driven modeling of human behavior is done [?], [?], [?], [?]. 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. Thus, our third novelty is abductive iterations where data drive models, and model predictions drive new experiments in a principled approach. Looping over abductive analyses is relatively rare (see the robotics work [?] as

IEEE/ACM ASONAM 2018, August 28-31, 2018, Barcelona, Spain
978-1-5386-6051-5/18/$31.00

an exception), and the use of abduction and abductive iterations in the social sciences is very rare. Our approach provides an exemplary case of coupling theory development/evaluation with real problems [**?**]: real data guide our theory evaluations.

This work was motivated in part by the DARPA Next Generation Social Science (NGS2) program. Goals of the program include devising methods to identify theories that are and are not applicable for explaining societal events.

B. Contributions

1. Specification and demonstration of iterative abductive analysis process. Using [**?**], [**?**] 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 anagrams 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. For example, our method of model construction from data (see Contribution 2 below) can be used to capture other temporal human action sets among interacting agents.

2. Data-driven networked agent-based models (ABMs) of experiments: design, construction, and evaluation. We design, construct, and evaluate three data-driven ABMs as part of the iterative abductive analysis. We adapt a conditional random fields (CRF) [**?**] modeling approach with four parameters to flexibly incorporate history effects on agent actions that evolve in time. It 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 [**?**]; 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, compare results from multiple ABMs, and evaluate the transition matrices of our ABMs using a statistical approach.

3. New experimental understanding of the formation of collective identity (CI). We discover three novel insights on the formation of CI by coupling the team anagrams game and DIFI score. First, players' DIFI scores increase with increasing numbers of neighbors in the anagrams game. Second, the number of interactions increases as number of neighbors (i.e., a player's network degree) increases from 2 to 4. However, the numbers of interactions, relatively speaking, saturates with further increases in degree. Third, despite this saturation, the DIFI score continues to increase with degree, suggesting complicated interactions among game parameters. Our analysis is a first work on quantifying the formation of CI since little work has been conducted on this subject in the literature. It is important to note that this experimental work (like the modeling work) takes place within the abductive loop framework.

Organization. An overview of the abductive loop process is presented in Section II, providing a framework for the rest of the paper. Related work is in Section III. In Sections IV and V, the experiment is described and modeling is presented, which address the respective components of the abductive loop. This enables a more streamlined description of the abductive loop for CI in Section VI. Section **??** summarizes.

II. Overview of Abductive Loop

Figure 1 illustrates our iterative abductive process, which includes inductive and deductive steps and hypothesis testing. This structure follows that of [**?**], [**?**], which are based on Piercian abduction [**?**], 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 subsequent experimental results, one-step abduction first generates data through experiments or observations. (Abduction uses data to drive the scientific 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 (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 [**?**]; 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.

Several variations on Figure 1 are possible. For example, the modeling, hypothesis, and theory steps may be interchanged, e.g., modeling results may be used to formulate hypotheses and devise theories. Also, as a model matures, a deductive analysis may be executed: (quantitative) model predictions in one iteration are evaluated with new experiments in the next iteration.

Furthermore, our abductive loop process makes modeling a much more prominent feature of the process than in [**?**], [**?**]: in hypothesis evaluation and in "What is Next?," which specifies experiments for the next loop. Moreover, we build generative ABMs, while models are based predominantly on similarity in [**?**]. Consequently, inference plays a large role in our work. We execute this loop in Section VI in evaluating CI.

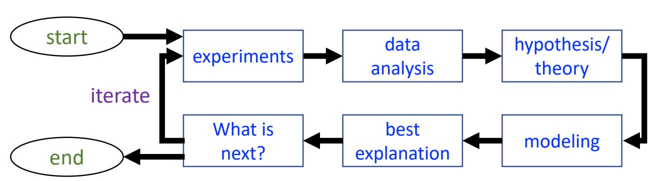


Fig. 1: Steps in our iterative abductive analysis/loop.

III. Related Work

Anagrams and CI Experiments. Over 20 experimental works (e.g., [**?**]) use anagrams games—with *individual* players. The only cooperative team-play of an anagrams game is reported in [?]; their goal, like ours, is to *produce* CI. While this motivated our experiment, there are several differences in procedures and context, e.g., theirs is a face-to-face experiment; ours is online. They measure CI with the proxy of public goods game contributions, while we use DIFI score. There are several in-person experiments studying the *implications* of CI where team members interact (e.g., [?]).

Agent Based Models of Anagrams and CI. The closest work to ours is [?], [?], where ABMs of identity diffusion are presented in which an agent adopts (changes) her type of identity to that of a neighbor with a stronger (higher valued) type of identity. Hence, this is a contagion process much like a voter model. We, in contrast, model the process of producing CI. There are no ABMs (or models of any kind) of group anagrams games, to our knowledge.

Data-Driven Modeling. Data-driven modeling works include [?], [?], [?], [?]. These works cover explore-exploit networked experiments with limited modeling [?]; individual models of single-choice (i.e., one-shot) evacuation decisions [?]; ABM of emotion and information contagions spreading on a network and comparisons with a single event [?]; and ABM of solar panel adoption and comparisons with data in San Diego county [?]. 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.

Abduction and Abductive Loop. Constructive procedures for implementing abductive analyses include [?], [?]. We extend those works by making modeling a first-class process, and by adding the process of what to do in the next iteration. In addition to the applications cited in the Introduction, abduction was used to understand emergency room personnels' efforts to save injured people in terms of "social viability" [?]. Perhaps the work closest to ours is [?] 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.

IV. Game Platform and Experiment (Game)

Web App Platform for Experiments. Owing to space limitations, we provide a very brief overview of the web app game platform that we built. The platform consists of the oTree infrastructure [?] 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 in a 2-phase game. Experiments and data analyses *are part of* the abductive loop of Sections II and VI and Figure 1.

2-Phase Game Description. Phase-1 is the group anagrams (word construction) game, where n players cooperate in sharing letters to form and submit words of length ≥ 3. Communication channels between pairs of agents mean that they can share letters with each other, and this induces a graph on the players. We use random regular graphs of degree k on the n players so that everyone has the same number of neighbors. Over all

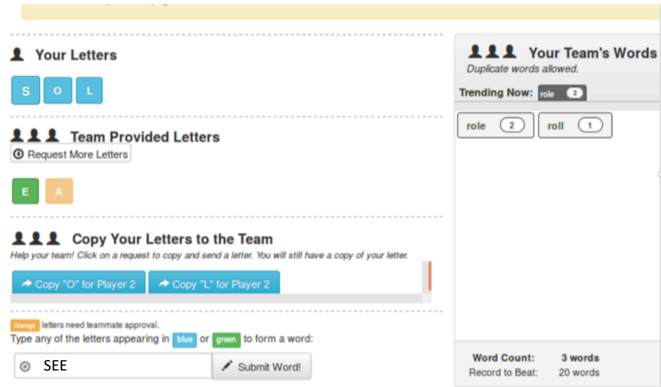


Fig. 2: The anagrams game screen, phase-1, for one player. This player has own letters "S," "O," and "L" and has requested an "E" and "A" from neighbors. The "E" is green, so this player's request has been fulfilled and so "E" can be used in forming words. But the request for "A" is still outstanding so cannot be used in words. Below these letters, it shows that Player 2 has requested "O" and "L" from this player; this player can reply to these requests, if she so chooses. Below that is a box where the player types and submits new words, like "SEE."

abductive loops, experiments are run in groups with nominal values of $10 \leq n \leq 20$ and with regular degree $2 \leq k \leq 8$.

A screen shot of one player's screen at one point in time, is shown in Figure 2. 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. Also, letters can be used any number of times in a word, meaning that if a player forms the word *tat* using a and one t (used twice), the player still has the a and t to form more words. Hence, a player needs to receive a requested letter only once.

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-minute anagrams game duration, players can idle, form words, request letters from their neighbors, and reply to requests.

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 interest to assist their neighbors. After the phase-1 anagrams game, each player is told the total number of words formed by the team, and each player's individual earnings.

The phase-2 DIFI procedure follows immediately after phase-1. Each player executes individually the DIFI procedure [?], to measure the degree to which the player feels part of the team (i.e., associates their identity with that of the team). Each player does this by moving a circle in a browser, representing herself, 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. Plots of data are in Sections V and VI.

V. Agent-Based Models of Anagrams Game and Modeling Results

We present three ABMs of the anagrams game that are used in the abductive loop analyses to follow in Section VI. 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 Section VI. All models, wherein each player is an agent, are data-driven, and hence *inductive inference* is used with data in three ways: to inform model structure, to characterize model parameters, and to estimate parameter values.

In all models, we represent the set V of players and their communication channels E as an undirected graph $G(V, E)$. The game is modeled as a discrete-time stochastic process, where at each time step, a player performs one of the actions from the action set $A = \{a_1, a_2, a_3, a_4\}$, consisting of: (i) a_1: idling (i.e., thinking); (ii) a_2: replying to a neighbor with a requested letter, (iii) a_3: requesting a letter from a neighbor, and (iv) a_4: forming and submitting a word.

ABM M0 is a baseline model that is presented after M1 for ease of exposition. ABMs M1 and M2 model the actions of A, but are generic in that request a_3 and reply a_2 letters 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 will form a word. 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 data.

A. Agent-Based Model M1, and then Baseline Model M0

ABM M1 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; $i, j \in [1..4]$; 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 $m \equiv |A|$ (here, $= 4$) and

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

To make Π dynamic in time and account for history effects, four variables are introduced that evolve in time: 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. Thus, letting $z = (1, z_L, z_W, z_B, z_C)_{(m+1)\times 1}$, we can model π_{ij} as a function of these covariates, i.e., $\pi_{ij} = f_{ij}(z)$. We use multinominal logistic regression to model π_{ij} as

$$\pi_{ij} = \frac{\exp(\boldsymbol{z}'\boldsymbol{\beta}_j^{(i)})}{1 + \sum_{h\neq i} \exp(\boldsymbol{z}'\boldsymbol{\beta}_h^{(i)})}, \quad (2)$$

where $\boldsymbol{\beta}_j^{(i)} = (\beta_{j1}^{(i)}, \ldots, \beta_{j,m+1}^{(i)})'$, $\boldsymbol{\beta}_i^{(i)} = \mathbf{0}$, and prime indicates transpose. For a given i, the parameter set can be expressed as

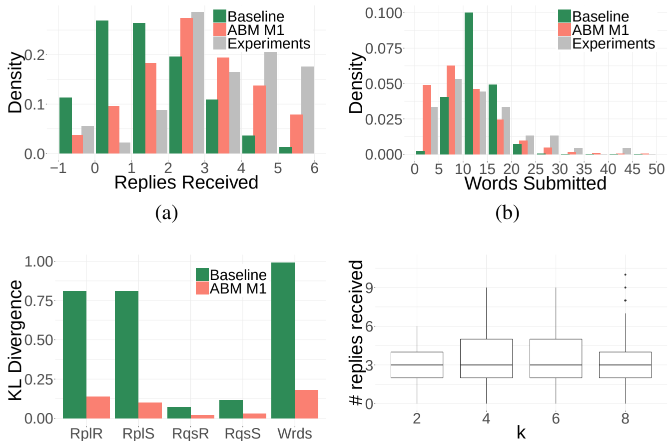
$$\boldsymbol{B}^{(i)} = \begin{pmatrix} \beta_{11}^{(i)} & \beta_{12}^{(i)} & \cdots & \beta_{1,m+1}^{(i)} \\ \beta_{21}^{(i)} & \beta_{22}^{(i)} & \cdots & \beta_{2,m+1}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{41}^{(i)} & \beta_{42}^{(i)} & \cdots & \beta_{4,m+1}^{(i)} \end{pmatrix}. \quad (3)$$

Baseline ABM M0. ABM M0 is a simplification of M1. The transition matrix Π is formed from the data so that the π_{ij} in Equation (1) are *constant*; time-invariant, independent of z.

Inductive Inference. We address the three dimensions of inference stated above. First, the model structure is informed by the $k = 2$ data. Second, and briefly, the parameters used in the feature vector z are justified as follows: $z_L(t)$ captures the idea that the more letters v has in-hand, the more likely the agent is to form words; $z_W(t)$ captures the notion that the more words that have been formed, the larger the vocabulary of the player. $z_B(t)$ captures the notion that the more letter requests that have not been replied to, the more likely v is to reply; and $z_C(t)$ captures the notion that the more time v is idle (thinking), the more likely v will take some other action at the next timestep. Third, parameters in Equation (3) are inferred from the $k = 2$ experimental data using the framework of maximum likelihood estimation for the multinomial distribution.

Results. Throughout, we use k to denote the number of neighbors (degree) of an agent 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*). To measure the performance of our models, we use *KL-divergence* between our model prediction on x and the experimental observation of x, *throughout this manuscript*.

Figure 3 provides model predictions and comparisons with experimental results. The first two plots show distributions of experimental data (in gray) and ABM M1 predictions in red. The green curves are from the baseline model M0. Clearly, ABM M1 is in better agreement with the experimental data compared to M0 in Figures 3a and 3b. From KL-divergence in Figure 3c, 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 graphs with larger $k > 2$, resulting in more interactions. Counterintuitively, as shown in Figure 3d, 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 $k = 2$ to predict the case of $k = 2$ (self-consistency checks), and to predict results of other cases with $k > 2$, as in Figure 3d.



(a) (b) (c) (d)

Fig. 3: ABM M0 and M1 predictions of the $k = 2$ experiments, and other simulation results. (a) Distribution of replies received, and (b) distribution of words formed, each at the end of the 5-minute anagrams game (gray bars) for all $k = 2$ experiments, compared to M1 predictions (red) for 100 simulations of an $n = 10$ player game. A baseline model M0 is shown in green for comparison. (c) KL-divergence values for the baseline (M0) and M1 models across the five parameters of x: lower values are better. These figures show that M1 generates predictions much closer to the experimental data than does M0. (d) 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.

We remark that we also fitted M1 using experimental data with $k = 4$, and consequently made predictions for the case of $k = 2$. We compared the distributions of x between prediction and experimental results using KL-divergence, whose values range from 0.11 to 0.46, indicating good predictions.

B. Agent-Based Model M2

Enhancement from ABM M1: arbitrary network topology. 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 (Section **??** below).

Development for Arbitrary Degree. We build a hierarchical model to incorporate the effect of agent degree k. For different values of k, the parameter coefficients in $\boldsymbol{B}^{(i)}$ in Equation (3) will be 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

$$\beta_{jh}^{(i)}(k) = \alpha_0^{(i,j,h)} + \alpha_1^{(i,j,h)}\xi_l(k) + \alpha_2^{(i,j,h)}\xi_q(k), \qquad (4)$$

where ξ_l and ξ_q are the linear and quadratic functions of the orthogonal basis in terms of k. This formulation provides a means to capture nonlinear effects. We have

$$\boldsymbol{B}^{(i)}(k) = \boldsymbol{C}_0^{(i)} + \boldsymbol{C}_1^{(i)}\xi_l(k) + \boldsymbol{C}_2^{(i)}\xi_q(k), \qquad (5)$$

where

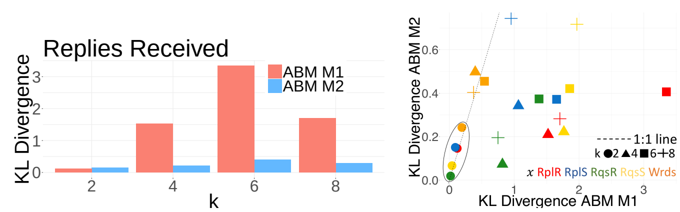
$$\boldsymbol{C}_r^{(i)} = \begin{pmatrix} \alpha_{11}^{(i,r)} & \alpha_{12}^{(i,r)} & \cdots & \alpha_{1,m+1}^{(i,r)} \\ \alpha_{21}^{(i,r)} & \alpha_{22}^{(i,r)} & \cdots & \alpha_{2,m+1}^{(i,r)} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{41}^{(i,r)} & \alpha_{42}^{(i,r)} & \cdots & \alpha_{4,m+1}^{(i,r)} \end{pmatrix}, r = 0, 1, 2, \qquad (6)$$

with $\alpha_{ih}^{(i,r)} = 0$ for any r and h.

Inductive Inference. To estimate the parameters sets $\boldsymbol{C}_0^{(i)}, \boldsymbol{C}_1^{(i)}, \boldsymbol{C}_2^{(i)}$, we use maximum likelihood estimation across the observations for $k = 2$, 4, 6, and 8.

Results. Figure 4 provides comparison results between ABMs M1 (red) and M2 (blue). KL-divergence values for distributions of replies received are shown in Figure 4a.

Although M1 performs well for $k = 2$, M2 betters M1 for $k > 2$, as M2 incorporates experimental data with $2 \leq k \leq 8$. This improvement is consistent among other x variables as shown in Figure 4b.



(a) (b)

Fig. 4: Comparisons of KL-divergence values generated with models ABM M1 and ABM M2. (a) KL-divergence values for ABM M1 (red) and M2 (blue) for distributions of replies received, for experiments with $k = 2$, 4, 6, and 8. M2 gives much better performance, as expected, as it explicitly accounts for agent degree. (b) A scatter plot of KL-divergence for M1 (x-axis) and M2 (y-axis) for 4 k values and 5 x variables (different scales on x,y axes). For $k > 2$, M2 performs better. As expected, M1 and M2 perform equally well (highlighted) for $k = 2$ as M1 is learned from $k = 2$ experimental data.

C. Model Evaluation

To evaluate the goodness of fitting for the proposed hierarchical model in Equation (4), we compare the estimated (model) transition probability matrix $\hat{\Pi} = (\hat{\pi}_{ij})$ for M2 with the empirical (data) transition probability matrix $\tilde{\Pi} = (\tilde{\pi}_{ij})$ under different settings of covariates. Here the empirical transition probability matrices $\tilde{\Pi}$ are obtained under the settings by grouping the value of each covariate with three levels, as described in Table I, to provide comparable numbers of samples across bins. Under each setting, 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 with transition from state i to state j. Consequently, we calculate the empirical probability $\hat{\pi}_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}$. There are 324 settings in total, and 279 of them have valid empirical transition probability matrices. For the estimated transition probability matrix $\hat{\Pi} = (\hat{\pi}_{ij})$, the value

TABLE I: Three bins and ranges of values for the z variables from Section V-A.

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

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

$$RMSE = \sqrt{\frac{1}{4|\mathcal{I}|}\sum_{i\in\mathcal{I}}\sum_{j=1}^{4}(\hat{\pi}_{ij} - \tilde{\pi}_{ij})^2} \quad (7)$$

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

Under each setting of covariates, we define the *Min.Count* as $n_{min} = \min_{i,j}\{n_{ij} : n_{ij} > 0\}$ as the smallest nonzero counts among transitions from state i to state j. It is known that when n_{min} is small, the empirical probability $\tilde{\pi}_{ij}$ is not accurate. Figure 5 shows the scatter plot between the RMSE and n_{min} for the 279 settings. From the figure, the proposed method generally provides an accurate estimation of probability transition matrix in most 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 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.

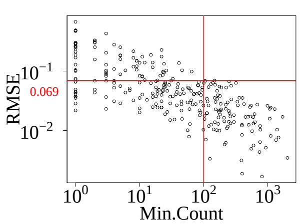


Fig. 5: Scatter plot of RMSE against Min.Count for different settings of covariates in Table I. See Equation (7) for RMSE and text for Min.Count.

VI. Abductive Loop (AL) Analyses and Results

In this section, we present the results of iterative abductive analyses, using the steps in Figure 1. We present two ALs. We also present an abductive tree with more hypotheses, that puts these two loops in context. We note that the experiments (Section IV) and modeling (Section V) are major components of the abductive looping process, and were separated out to make this section more streamlined.

A. Abductive Loop 1 (AL-1)

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

Data Analysis. Data were explored for patterns [**?**]. Time-series of actions in A, per player, show that players tend to request particular letters with the goal of forming specific words: when a requested letter is received, there is often a burst of words formed with that letter. The other extreme, of requesting all letters initially and then figuring out words, is done far less frequently. The time-histories of actions, per player, also led to the model structures of the ABMs. Data for all players, for the five components of x in Section V-A, were combined to produce distributions of numbers of actions, against which models were compared. Measurements of these variable values are selected for correlation with DIFI scores.

Hypothesis/Theory. Hypothesis H_{11}: *In the team-based anagrams game, the sense of 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 [**?**] 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 [**?**] indicates that interactions are important for forming an interdependent 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.

Model. The model of Section V-A was constructed from the time histories of actions of players for experiments with $k = 2$. The results relevant to this iteration are provided in Figure 3. ABM M1 is much better at capturing the dynamics in the experiments than is baseline model M0.

Best Explanation. Results of a linear regression in Table II indicate that hypothesis H_{11} is falsified because Wrds, the number of words formed, is not significant, while RplR, RplS and RqsS are significant. Thus, Social Exchange Theory can be eliminated as 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). 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).

TABLE II: Results of linear regression of variables in x against dependent variable DIFI score, indicating that interactions are more significant than number of words formed in producing CI. These data are generated in AL-1.

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

What is Next? Figure 3d indicates that the model predicts behavior that is invariant with the degrees 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 as using increasing k (i.e., $k > 2$). Thus, the ABM M1 (driven by the data) is guiding what to do next. However, we are not using M1 to predict specific quantitative experimental results, as in a deductive analysis. Rather, M1 is used in a qualitative manner.

B. Abductive Loop 2 (AL-2)

In reality, we executed multiple abductive iterations, studying, in turn, $k = 4$ and then $k = 6, 8$. However, in the interest of space, we combine them into one iteration.

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

Data Analysis. We continued the same types of analyses described in AL-1, but with the added dimension of k.

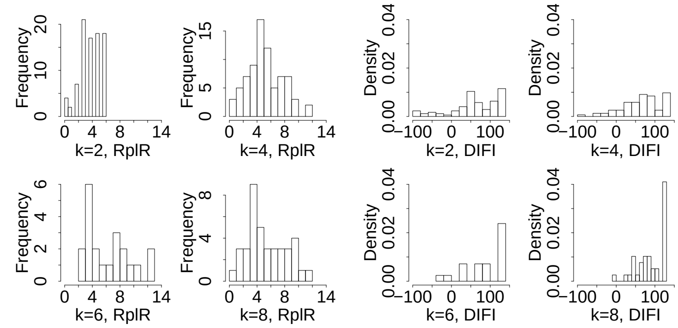
Hypothesis/Theory. Hypothesis H_{22}: *(a) As the number of neighbors increases in the anagrams game, the level of CI increases because there are more interactions. (b) However, beyond four neighbors (equivalently, for more than 12 neighbor letters) there is no benefit of additional neighbors.* The theory of social interactions states that interactions with more neighbors creates more interdependence. Theory of cognitive load [**?**] suggests that cognitive load might be too great at some point, resulting in a player being unable to take advantage of more input.

Model. The model of Section V-B was constructed from the time histories of actions of players, from the combined data from *both* iterations. Model results relevant to this iteration are provided in Figure 4. ABM M2 captures trends in degree k much more effectively than ABM M1, for all parameters of x. Although not shown, ABM M2 shows effects of increasing k: the predicted numbers of actions increases as k increases, which is far different from the predictions in Figure 3d for ABM M1, where k has no effect.

Best Explanation. Figure **??** provides results that address hypotheses H_{22}(a) and H_{22}(b). Note that H_{22}(b) has two interpretations because "benefit" may be in terms of numbers of interactions (from AL-1) or in terms of DIFI score. We first examine H_{22}(b) in terms of numbers of interactions. Figure **??** shows the frequency distributions for replies received, for the four values of k. Focusing on interactions, note the large change in distributions in going from $k = 2$ to $k = 4$, but relatively minor changes for further increases in k. Thus, the saturation in the distributions (and others are similar), supports hypothesis H_{22}(b): the number of neighbors increases, but the number of interactions does not, for $k > 4$. This is consistent with cognitive load theory.

Now we evaluate H_{22}(a) and (b) in terms of DIFI score. Figure **??** shows that as k increases from 2 to 8, the probability density of DIFI scores shifts demonstrably to increasing DIFI. That is, greater numbers of neighbors produce more CI, as measured by the DIFI score. This does not wholly support H_{22}(a); while increasing k does correlate with increasing DIFI score, it is not because of the number of interactions, which does not increase appreciably for $k > 4$. These data falsify H_{22}(b): there is additional benefit, in terms of increased DIFI score, with increasing number of neighbors. The applicability of the theory of social interactions is not clear, but the data suggest that it is the number of different people with whom one interacts that is important, rather than the total number of interactions. More experiments are needed.

What is Next? At this point we halt the iterative abduction process for this paper, although it could continue. In a next iteration, we could (i) try to isolate the effects of number of interactions versus the number of neighbors in different experiments, or (ii) study the effects of different degrees of players and different numbers and qualities of letters initially assigned to players within the same experiment. 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.



(a) (b)

Fig. 6: Statistical analysis correlation results of the anagrams game parameters and DIFI score. **??** Frequency distributions of replies received change markedly from $k = 2$ to $k = 4$, but relatively little for further increasing k. **??** Probability density of DIFI score moves dramatically to larger DIFI score with increasing k. Each of these results is novel—these results are new. All the more novel is the combination of the two: while game measurables saturate (other data besides replies received), the DIFI score does not. These data are generated in AL-2.

C. Abductive Loops: Role of Analyst and Bigger Picture

Two ALs have been demonstrated. Many additional loops are possible, as illustrated in Figure **??**, which depicts several hypotheses, including the two addressed above (in orange). Figure **??** and Table **??** 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 Figure **??** 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.

Fig. 7: Abductive tree representing candidate abductive loops with dependencies. Hypotheses are nodes, and are provided in Table **??**; edges are outcomes of ALs. 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. Note that the actual hypotheses are not specified a priori; they are based on newly-generated data from the abductive iterations per Figure 1.

TABLE III: Candidate hypotheses to be evaluated in abductive iterations of Figure ??. Hypotheses H_{11} and H_{22} are given in the text. H_{43} is the same as H_{22}.

Hypothesis Number	Description
H_{12}	Playing the Phase 1 Anagrams game will produce greater individual DIFI scores than not playing Phase 1.
$H_{21} = H_{32}$	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.
$H_{31} = H_{42} = H_{44}$	Playing the game with players face to face will produce greater individual DIFI scores in Phase 2 (by enabling players to communicate and pick up on visual and verbal cues).
H_{33}	Lesser payouts in the Phase 1 anagrams game means that players do not have enough incentive to engage their neighbors.
H_{41}	Having the Phase 1 game score of another team displayed during Phase 1 will increase CI because it will create a stronger in-group/out-group paradigm.

VII. Summary

We formalize an abductive loop, implement it computationally, and exercise it in an experimental setting (the 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.

ACKNOWLEDGMENT

We thank the reviewers of ASONAM 2018 for providing valuable suggestions. We also thank the computer systems administrators and managers at the Biocomplexity Institute for their help in this and many other works: Dominik Borkowski, William Miles Gentry, Jeremy Johnson, William Marmagas, Douglas McMaster, Kevin Shinpaugh and Robert Wills. This work is supported in part by DARPA Cooperative Agreement D17AC00003 (NGS2), DTRA CNIMS (Contract HDTRA1-11-D-0016-0001), National Science Foundation via grants DGE-1545362, IIS-1633363, and DIBBS ACI-1443054, and Army Research Laboratory under grant W911NF-17-1-0021. The US Government is authorized to reproduce and distribute reprints of this work for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of DARPA, DTRA, NSF, Army Research Laboratory, or the U.S. Government.