Modeling the Contributions of Individuals to Collaborative Problem Solving Using Epistemic Network Analysis

Zachari Swiecki, University of Wisconsin-Madison, swiecki@wisc.edu
David Williamson Shaffer, University of Wisconsin-Madison, Aalborg University-Copenhagen, dws@education.wisc.edu

Abstract: Using data from team training simulations, I will use Epistemic Network Analysis to model collaborative problem solving while accounting for time-dependent relationships between the contributions of an individual and the contributions of their team. I will conduct a qualitative analysis of the data, comparisons between models that either ignore or account for team contributions, and a simulation study that investigates the conditions under which team contributions significantly impact the individual level of analysis. This work will test a measurement approach that is potentially more valid than extant approaches and provide tools for determining whether the interaction between team and individual will meaningfully impact models.

Project background and goals

Collaborative problem solving (CPS) is widely recognized as critical 21st Century Skill. When individuals on teams solve problems, their processes include actions toward accomplishing a task and actions toward managing the processes of collaboration. Team processes are not simply the sum of individual actions; rather, individual actions *interact* with one another, creating a context independent of any single individual. As interactions unfold, they contribute to the *common ground*—the shared knowledge and experience that exists between people when they interact (Clark, 1996); as a result, the discourse of the team is *interdependent*: the actions of one individual impact the actions of others on the team. Importantly, this impact is *temporally* bounded; as Suthers and Desiato (2012) argue, interactions are interpreted with respect to immediately preceding events. These features suggest that valid measurements of CPS should account for the relationships between the contributions of an individual and the recent contributions of the team.

Despite these features of CPS, extant measurement approaches treat individuals as isolated, independent, and atemporal actors. For example, *coding and counting* aggregates discourse codes attributed to individuals over the course of collaboration while ignoring contributions from the team. Other approaches script CPS interactions. For example, the PISA CPS assessment (OECD, 2017) has students interact only with computer agents using a small set of choices. Still other approaches, such as sequential pattern mining (SPM) and social network analysis (SNA), model the complexity of CPS. However, SPM only models individuals irrespective of the team, and SNA only models the structure of team processes (who talks to whom), not the discourse.

An alternative approach is Epistemic Network Analysis (Shaffer, 2017) (ENA). ENA builds network models that describe interactive, interdependent, and temporal phenomena that (unlike SNA) describe interactions among the discourse of an individual and the discourse of others within the context of recent contributions (Siebert-Evenstone et al., 2017). It thus accounts for the interactivity, interdependence, and temporality of CPS at the individual level.

In my dissertation project, I will examine ENA for modeling individual CPS processes. I ask: (1) How do individual contributions relate to team contributions? (2) How does a quantitative model that accounts for the relationships between individual and team contributions compare to models that ignore these relationships? (3) Under what conditions should we account for the relationships between individual and team contributions?

Methods

My data comes from the Tactical Decision Making Under Stress project. Sixteen teams participated in simulated training scenarios to test the impact of a new decision-support system (DSS) on team performance in the context of air defense warfare (Cannon-Bowers & Salas, 1998). During the scenarios, teams needed to detect and identify ships and aircraft, assess whether they were threats, and decide how to respond. Teams in the control condition use standard technology; teams in the experimental condition used the DSS. The dataset consists of transcripts of team communications.

To address my first research question, I will conduct a grounded qualitative analysis informed by existing literature, and code the data with topics relevant to this CPS context.

To address my second research question, I will create three models of individual contributions. Models I and II will *ignore* the contributions of other team members, while Model III will *account for* their contributions.

For Model I, I will calculate the code frequencies for each code for each individual and use Principal Components Analysis (PCA) to create explanatory variables. For Models II and III, I will use ENA. The ENA algorithm uses a sliding window to construct a network model for each turn of talk in the data, showing how codes in the current turn of talk are connected to codes within the recent temporal context. Individual networks are created by aggregating turns of talk for each person. In this way, ENA can model the connections that each individual makes between concepts and actions while accounting for the actions of others. ENA uses a technique similar to PCA to create ENA scores for each individual which summarize their network of connections; I will use these ENA scores as explanatory variables for Models II & III. In Model II, I will create an ENA model that only identifies connections between codes within an individual's own talk. In Model III, I will create an ENA model that identifies connections between an individual's own talk and the talk of other team members. For each model, I will construct a subsequent regression model of individual performance using explanatory variables from ENA or PCA. I will compare these models based on their fidelity to the qualitative analysis, variance explained, and model efficiency measured by the Akaike information criterion (AIC).

To address my third research question, I will generate simulated datasets that vary individual contributions to the common ground by varying the code frequency distribution of the team, and create ENA models for each dataset. At one extreme, the distribution of code frequencies will be completely asymmetric, simulating situations where team members have specific expertise and the team must be interdependent to function. At the other extreme, the distribution of code frequencies will be completely symmetric, simulating situations where team members have identical expertise. I will vary the distribution of code frequencies between these extremes to identify the *threshold* at which contributions of other team members have a significant impact on individual contributions by testing for significant differences between ENA scores at each step.

Expected findings and contributions

My preliminary work has focused on processes used by commanders in the two conditions. I produced a reliable automated coding scheme and conducted a qualitative analysis of the data. I developed two quantitative models of commander contributions, one using a coding and counting approach and one using an ENA approach accounting for team interactions. Only the ENA approach corroborated findings from the qualitative analysis, and the ENA approach found significant differences that coding and counting did not, suggesting that it could be a more valid approach for measuring individual CPS. I will extend these findings by examining all roles on the team, making more sophisticated quantitative comparisons between models, and using simulation studies. These studies will determine whether team contributions are more relevant in datasets with asymmetric code frequency distributions, suggesting that accounting for the interaction between individual and team is critical when assessing CPS for individuals, particularly those in heterogeneous groups.

The implications of this work are primarily methodological: This work will empirically test an approach to measuring individual contributions to CPS that is potentially more valid than extant approaches. Similarly, the proposed simulation study could provide a tool that allows researchers to determine if the interaction between team and individual will meaningfully impact their models. Furthermore, this approach will be replicable across contexts, and in many cases, automatable, meaning that it could be integrated with real-time or after-action assessment systems. As such, this work also has the potential to impact pedagogy and learning in a variety of collaborative contexts.

References

Cannon-Bowers, J.A., Salas, E. (1998). *Making decisions under stress: Implications for individual and team training* (pp. 39–59). Washington, D.C.: American Psychological Association.

Clark, H. H. (1996). *Using language*. Cambridge university press.

OECD. (2017). PISA 2015 collaborative problem‑solving framework. In *PISA* (pp. 131–188). Organisation for Economic Co-operation and Development.

Shaffer, D. W. (2017). *Quantitative ethnography*. Madison, WI: Cathcart Press.

Siebert-Evenstone, A., Arastoopour Irgens, G., Collier, W., Swiecki, Z., Ruis, A. R., & Williamson Shaffer, D. (2017). In Search of Conversational Grain Size: Modelling Semantic Structure Using Moving Stanza Windows. *Journal of Learning Analytics*, *4*(3), 123–139.

Suthers, D. D., & Desiato, C. (2012). Exposing chat features through analysis of uptake between contributions. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 3368–3377). IEEE.