

Modeling and Visualizing Team Performance using Epistemic Network Analysis

Zachari Swiecki¹, A. R. Ruis¹, David Williamson Shaffer^{1,2}

University of Wisconsin-Madison¹, Aalborg University Copenhagen²

INTRODUCTION

A key goal of The Generalized Intelligent Framework for Tutoring (GIFT) project is to scaffold the development of tutoring scenarios that support team-training and assessment (Sinatra, 2018). As Ruis and colleagues (Ruis, Hampton, Goldberg, & Shaffer, 2018) argue, a critical component of team training and assessment is the ability to *model* team performance. In this paper, we describe a specific approach to modeling team performance at both the individual and team levels that prior work suggests is more valid than extant approaches. Moreover, we argue that modeling team performance is not sufficient on its own—given the goals of team tutors, we also need visualizations that effectively summarize team performance and provide actionable information at both the team level and the individual level. Here, we describe the design of a team-tutoring *dashboard* that would allow tutors to monitor individual and team performance in real-time. This system could inform assessment or guide the delivery of feedback either in real-time or after the tutoring scenario is complete.

TEAM PERFORMANCE

When individuals on teams solve problems, their processes include (a) actions toward accomplishing a task and (b) actions toward managing the processes of collaboration. Thus, 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*, or 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. Moreover, team processes have an important *temporal* dimension: team processes unfold in time and are interpreted with respect to the immediately preceding actions—or *recent temporal context*—not the entire history of team interaction (Reimann, 2009; Suthers & Desiato, 2012).

This complexity suggests that valid models of team performance at the *team* level should account for relationships among the recent contributions of the team, and valid models at the *individual* level should account for relationships between a given individual's contributions and the recent contributions of the rest of the team.

Modeling Team Performance

Despite these suggestions, many extant modeling approaches still employ *coding-and-counting* (Chi, 1997; Suthers, 2006). At the team level, this involves aggregating behavioral markers or codes over the entire history of a team task or scenario, ignoring temporal aspects of team processes. Similarly, coding-and-counting at the individual level ignores temporality, and because it separates the processes of individuals from the processes of the team, it also ignores the interactive and interdependent aspects of individual contributions (Csanadi, Eagan, Shaffer, Kollar, & Fischer, 2018). More nuanced analyses are often conducted with techniques that model frequent sequential events in data, such as *sequential pattern*

mining; however, similar to coding-and-counting, such techniques can only model individuals irrespective of the team (Swiecki, Lian, Ruis, & Shaffer, in press [A]).

An alternative approach that can account for these critical aspects of team processes at the team and individual levels is *epistemic network analysis* (ENA) (Shaffer, 2017; Shaffer, Collier, & Ruis, 2016; Shaffer & Ruis, 2017). Specifically, ENA models team activity by identifying categories of action, communication, cognition, and other relevant features and characterizing them with appropriate coding schemes into smaller sets of domain-relevant nodes. The weights of the connections among network nodes (i.e., the association structure of key elements in the domain) are then computed and visualized. Critically, ENA models team actions and interactions in such a way that it is possible to *extract information about each team member's contributions to team performance*.

ENA uses statistical and visualization techniques to enable comparison of the salient properties of different networks, including networks generated by different teams or by teams at different points in time, teams in different spatial locations, or teams engaged in different activities. These salient properties are modeled not just in terms of the general structure of the networks, but ENA also extracts properties relevant to the actual content of the network.

In other words, ENA can analyze (a) what teams are doing, (b) how they are thinking, (c) what role individuals are playing in team performance, and (d) how teams compare to one another in the context of real problem solving. Moreover, prior work using ENA to model the performance of U.S. military teams—a key domain of interest for GIFT—has shown that ENA has both statistical and interpretive advantages compared to coding-and-counting and sequential pattern mining (Swiecki et al., in press [A]; Swiecki, Ruis, Farrell, & Shaffer, in press [B]).

TEAM-TUTORING DASHBOARD

While models such as those produced by ENA are useful tools for examining team performance, they are designed primarily for researchers. As such, their affordances are not necessarily aligned with the goals of other audiences that have an interest in understanding team performance, such as tutors or the teams themselves (Swiecki & Shaffer, 2018). For example, an important goal of researchers is to advance their understanding of phenomena or make predictions, and they are trained to understand and use complex models and visualizations to help them do so. Tutors, on the other hand, need to assess performance and guide interventions, and they may lack the training required to effectively use complex models and visualizations. In turn, tutors need tools that quickly highlight the teams or individuals that need their attention the most, while also providing them information that can guide their interventions.

As a first step toward integrating such a system with GIFT, we have created preliminary designs of a team-tutoring dashboard. This dashboard uses simplified ENA models to provide actionable information on the performance of teams and individuals. These designs are based on prior work in which we successfully designed, built, and implemented an ENA-driven team performance dashboard in a simulation-based learning environment (Herder et al., 2018). The ENA models presented in the designs are based on prior work by Swiecki and colleagues (in press [A], in press [B]). We summarize the data and relevant models from this work in more detail below.

ENA Models of Team Performance

As part of the Tactical Decision Making Under Stress project, sixteen teams participated in training scenarios to test the impact of a new decision-support system on team performance in the context of air

defense warfare (Johnston, Poirier, & Smith-Jentsch, 1998). During the scenarios, teams needed to detect and identify ships and aircraft (referred to as *tracks*), assess whether they were threats, and decide how to respond. Each team consisted of six members who held either a leadership role, such as the Commanding Officer (CO), or a support role, such as the Electronic Warfare Supervisor (EWS). The dataset consists of transcripts of team communications and performance scores for each team.

To create the ENA models, we developed and validated an automated coding scheme that captured the critical aspects of the team task. After coding, we used ENA to create models at the team and individual level. The ENA algorithm uses a sliding window to construct a network 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. In other words, ENA defines a connection between codes as their co-occurrence within a specific number of turns of talk. To create networks for each unit of analysis, ENA aggregates the networks associated with their turns of talk. In this way, ENA can model the network of connections that each team or individual makes between concepts and actions while taking into account the recent actions of others (Siebert-Evenstone et al., 2017).

Two coordinated representations are produced for each team or individual network: an ENA score and a weighted network graph. ENA uses a dimensional reduction via spectral value decomposition (SVD) to create an ENA score for each team or individual that summarizes their network of connections. These scores give their location in the *ENA space* created by the dimensional reduction. Typically, this dimensional reduction maximizes the variance accounted for by each dimension. However, ENA can also combine SVD with a hyperplane projection such that the first dimension maximizes the variance between the means of two subpopulations—for example, high and low performing teams—present in the data.

The nodes of the weighted network graphs correspond to codes, and the edges are proportional to the relative frequency of connection between two codes. The positions of the network graph nodes are fixed across networks, and their positions are determined by an optimization algorithm that minimizes the difference between the ENA scores and their corresponding network centroids. This relationship implies that ENA scores toward the extremes of a dimension have network graphs with strong connections between nodes located on the extremes. As a result, dimensions in this ENA space distinguish teams in terms of connections between codes whose nodes are located at the extremes. In addition, ENA can produce network difference graphs which subtract the edge weights of two networks to show the connections that are strongest in one network relative to another.

At the team level, our analysis suggested that high performing teams made frequent connections between *tactical information*, such as track behavior and track detection, and *tactical actions* such as combat orders. Low performing teams made relatively frequent connections to seeking information, suggesting that they had difficulty maintaining situational awareness (Figure 1, left).

At the individual level, our analysis suggested connections for leadership and support roles that distinguished those on high and low performing teams. Connections that distinguished individuals in leadership roles were very similar to those that distinguished high from low performing teams, so we do not describe them in detail here. Individuals in support roles on high performing teams made frequent connections to status updates, suggesting that they played a critical role in updating the team on the evolving tactical situation. Individuals in support roles on low performing teams made more frequent connections to seeking information which suggests that they were focused on repairing the team's understanding of the tactical situation (Figure 1, right).

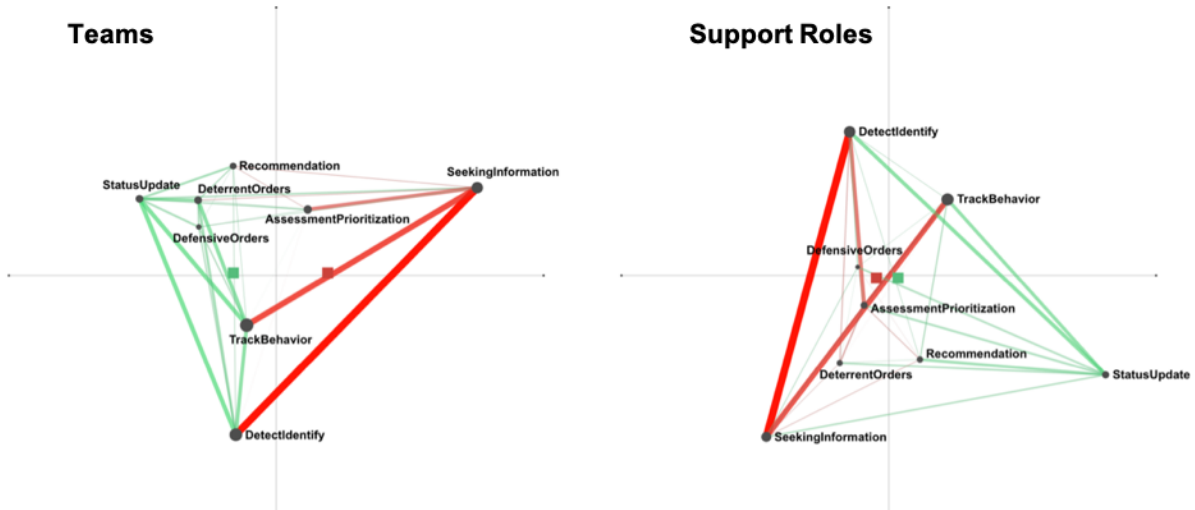


Figure 1. ENA Models of Team Performance: ENA difference network between high performing teams (left, green) and low performing teams (left, red). ENA difference network between individuals in support roles on high performing teams (right, green) and individuals in support roles on low performing teams (right, red). Difference network for individuals in leadership roles not shown due to similarity with team network.

In the next section, we describe our process for integrating ENA models in to a team-tutoring dashboard using these data and results as an example.

Dashboard Designs

The proposed dashboard features a performance *overview* of all teams and individuals and the ability to *drill down* to more specific information about the performance of a team or individual. In the performance overview (Figure 2), each row represents a different individual in a particular role (e.g., CO, EWS, etc.) grouped by their team; each column represents a different training scenario. For a given scenario, high performing individuals are indicated by a green circle, average performers are shown in yellow, and low performers in red. Team members with no activity are represented by an empty circle, and those with no relevant activity by a grey circle. High, average, and low performance indicators are determined by thresholds on the distribution of ENA scores from either the leadership or support ENA model. For these designs, thresholds were set at the first and third quartile of the distributions, but in the general case, they would be customizable.

This overview has several affordances for team tutors. First, it provides a quick reference for how teams or individuals are performing, and thus directs attention to those who may need an intervention. Second, it presents the performance of a team or individual in the context of others, which facilitates comparisons. Third, the horizontal axis allows tutors to track performance over time to examine and compare trends. Finally, the vertical axis allows tutors to track performance across a given scenario to examine whether that scenario is more or less difficult than others. Such affordances are important because they scaffold decisions about whether an intervention is necessary and what kind of intervention to provide. For example, a CO who has high performance across all scenarios but one would likely need a different intervention than a CO whose performance was more variable over time.



Figure 2. Performance Overview

By clicking a team or individual, tutors can drill down to see more detailed information about their performance. For example, Figure 3 shows a simplified *network model* of one team's performance at the end of a training scenario. Green connections between codes are characteristic of high performing teams; red connections are characteristic of low-performing teams. In other words, high-performing teams will have a higher frequency of green connections relative to red.

To create this simplified network model, we selected the connections from the ENA models described above that explained the most variance between high and low performing teams in the dataset—that is, the connections at the extremes of the first dimension in the ENA space. Unlike the models described above, the node placement of the simplified networks is designed for easy comprehension, with nodes connected by green (i.e., indicative of high performance) connections placed at the top of the display. In addition, this drill down view shows the team *activity* represented in the network model—in this case, the coded team transcript. Turns of talk in the activity record with black circles have a code present in the turn. Below the network is a *description* of the connections present in the network (See Figure 5 for more details). Note that placeholder text is used for the activity record and network description in all designs except those in Figure 5.

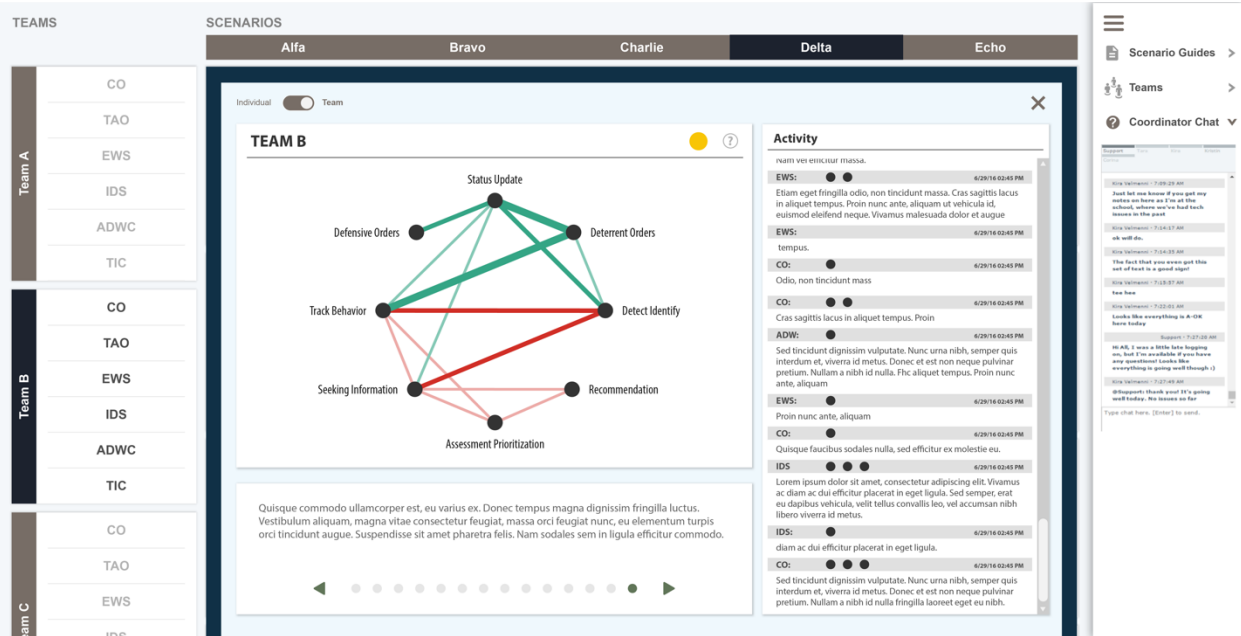


Figure 3. Team Network Model

Tutors can access similar drill downs for each individual on the team. For example, Figure 4 shows the network visualization of this team's EWS, who holds a support role, at the end of a training scenario. Green connections are characteristic of high-performing individuals in support roles; red connections are characteristic of low performing individuals in support roles. To create simplified network visualizations for individuals in either leadership or support roles, we selected the connections from the ENA models described above that explained the most variance between individuals in those roles who were on high performing versus low performing teams. Node placements match the positions of the team networks to maintain visual consistency between individuals and teams.

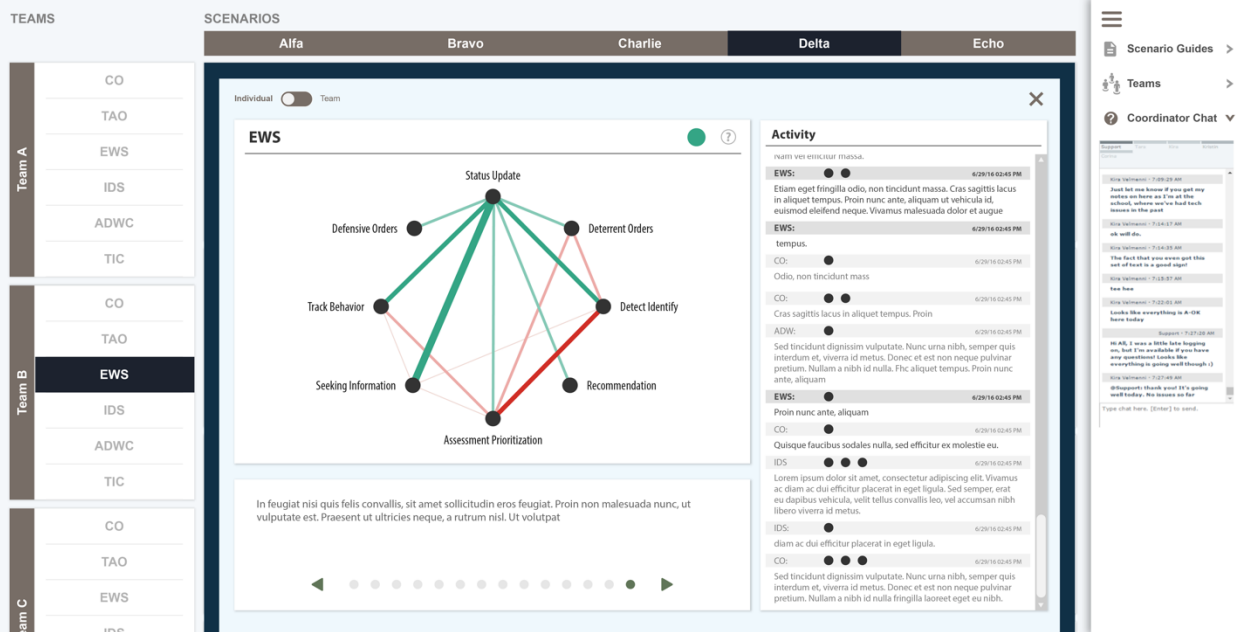


Figure 4. Network Model

Tutors can use the arrows in the description below the network model to step through the scenario in time and review each connection (and the activity contributing to the connection) made by the individual or team. For example, in Figure 5, we can see the first connection made by the team's CO in this training scenario. Here, the CO is responding to tactical information from the Tactical Action Officer (TAO) with an order.

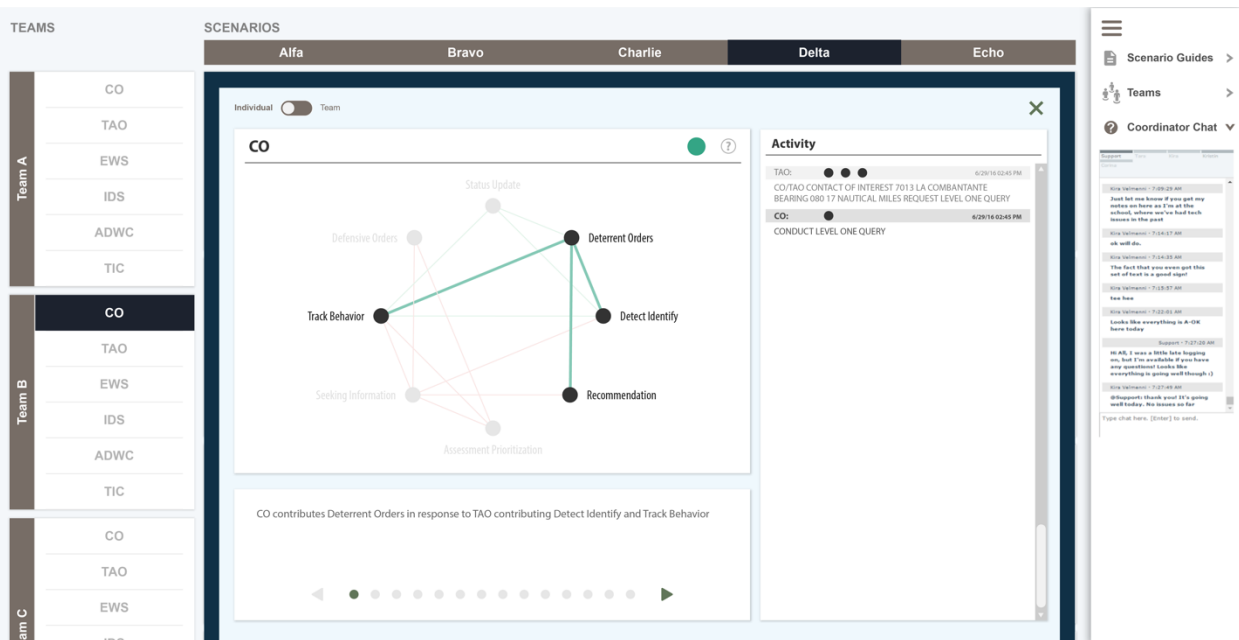


Figure 5. Network Review

In addition to stepping through the network model, tutors can examine the activity contributing to a connection by clicking the connection in the network model. As shown in Figure 6, clicking a connection in the network highlights the most recent activity in which the connection occurred.

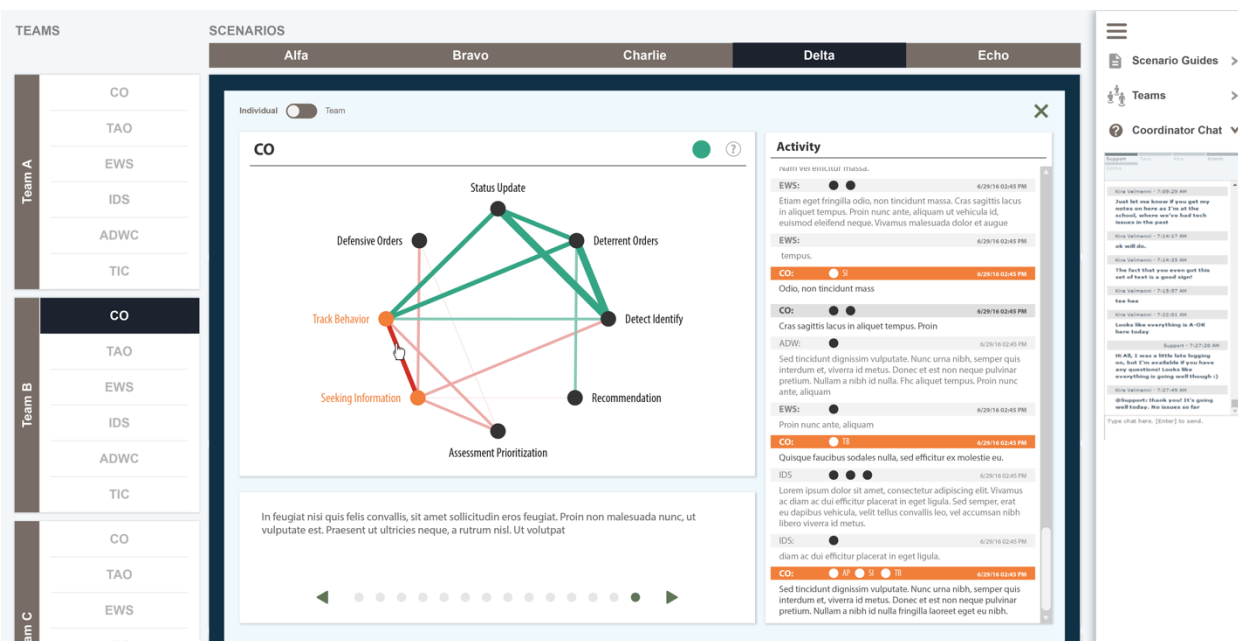


Figure 6. Connection Inspection

Similarly, as shown in Figure 7, tutors can also click a segment of activity in the activity record to highlight any connections that may have occurred within the recent temporal context of that segment.

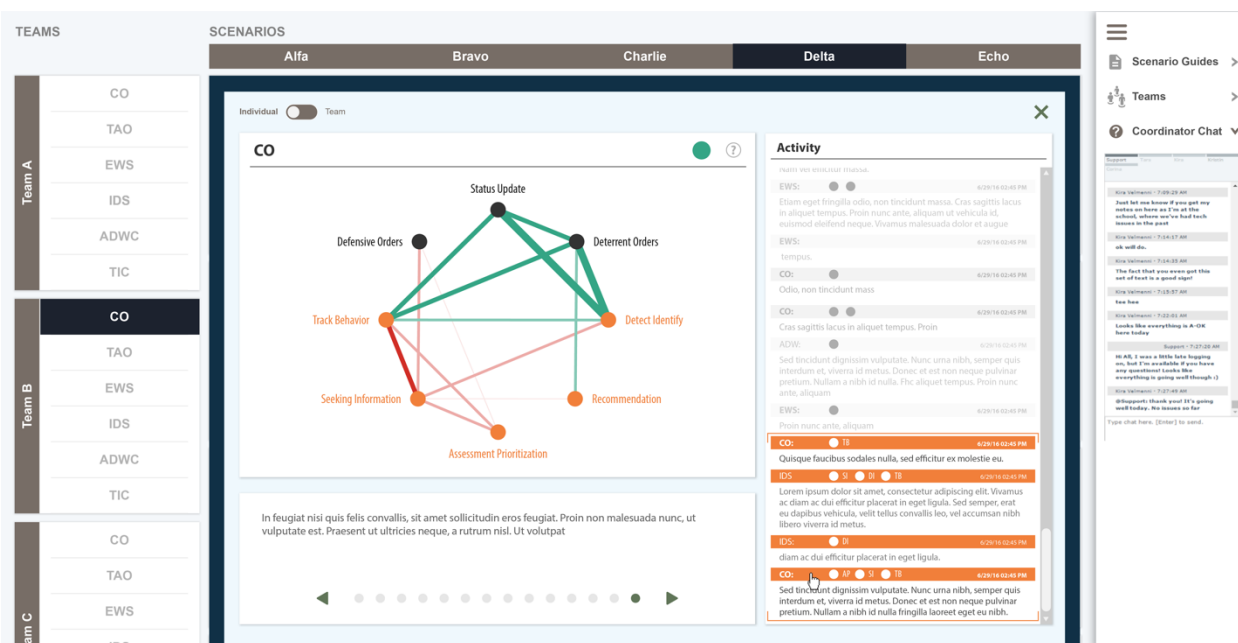


Figure 7. Activity Inspection

Features such as stepping through the network model and investigating connections and their corresponding activity can facilitate after-action reviews by the tutor with teams or specific individuals.

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

The work presented here suggests two recommendations for GIFT. First, team-tutoring assessments in GIFT should include models, such as ENA, that account for the interactive, interdependent, and temporal nature of team processes at both the team and individual levels. Second, in order to be successful, such assessments should include actionable visualizations that help tutors monitor, assess, and provide feedback on team performance. The designs described above are one proposal for such an approach.

While these designs used specific data from a particular domain for illustrative purposes, the approach is agnostic to both the kind of data collected and the domain from which it comes. The only constraints are that the data consists of a machine readable record of ordered events, which may be talk, gestures, mouse-clicks, or other actions, and that there exists a reliable automated coding scheme for the data. In cases where the data is not in a textual format, such as raw audio or video recordings, the system would need a means of converting the data in real-time into a format which could be automatically coded. Moreover, while the network models described here were data driven, it is also possible for authors to specify a priori the connections that distinguish high and low performance for teams and individuals.

Our future work will include adapting the dashboard designs to different team structures, such as squads or platoons, and mapping the components of these designs to the inputs of *domain knowledge files* that manage assessment and pedagogical requests for teams or individuals during a GIFT-managed scenario.

ACKNOWLEDGEMENTS

This work was funded in part by the National Science Foundation (DRL-1661036, DRL-1713110), the U.S. Army Research Laboratory (W911NF-18-2-0039), the Wisconsin Alumni Research Foundation, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.

REFERENCES

- Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6(3), 271–315.
- Clark, H. H. (1996). *Using language*. Cambridge university press.
- Csanadi, A., Eagan, B., Shaffer, D. W., Kollar, I., & Fischer, F. (in press). When coding-and-counting is not enough: Using Epistemic Network Analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning*.
- Herder, T., Swiecki, Z., Fougat, S. S., Tamborg, A. L., Allsopp, B. B., Shaffer, D. W., & Misfeldt, M. (2018). Supporting teacher's intervention in student's virtual collaboration using a network-based model. *Proceedings of the International Conference on Learning Analytics*, 21–25. Sydney, Australia.
- Johnston, J. H., Poirier, J., & Smith-Jentsch, K. A. (1998). Decision making under stress: Creating a research methodology. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 39–59). Washington, D.C.: American Psychological Association.

- Reimann, P. (2009). Time is precious: Variable- and event-centered approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 239–257. <https://doi.org/10.1007/s11412-009-9070-z>
- Ruis, A. R., Hampton, A. J., Goldberg, B. S., & Shaffer, D. W. (2018). Modeling processes of enculturation in team training. In R. Sottialre, A. Graesser, & A. M. Sinatra (Eds.), *Design Recommendations for Intelligent Tutoring System: Volume 6 - Team Tutoring* (pp. 45–51). Orlando, FL: U.S. Army Research Laboratory.
- Shaffer, D. W. (2017). *Quantitative ethnography*. Madison, WI: Cathcart Press.
- Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, 3(3), 9–45.
- Shaffer, D. W., & Ruis, A. R. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 175–187). Society for Learning Analytics Research.
- 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. <https://doi.org/10.18608/jla.2017.43.7>
- Sinatra, A. M. (2018). Team Models in the Generalized Intelligent Framework for Tutoring: 2018 Update. In R. Sottialre (Ed.), *Proceedings of the Sixth Annual GIFT Users Symposium* (pp. 157–161). Orlando, FL.
- Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning: New York*, 1(3), 315–337. <http://dx.doi.org.ezproxy.library.wisc.edu/10.1007/s11412-006-9660-y>
- Suthers, D. D., & Desiato, C. (2012). Exposing chat features through analysis of uptake between contributions. *System Science (HICSS), 2012 45th Hawaii International Conference On*, 3368–3377. IEEE.
- Swiecki, Z., Lian, Z., Ruis, A. R., & Shaffer, D. W. (in press) [A]. Does order matter? Investigating sequential and cote temporal models of collaboration. *Proceedings of the 13th International Conference of Computer-Supported Collaborative Learning*. Presented at the Lyon, France. Lyon, France.
- Swiecki, Z., Ruis, A. R., Farrell, C., & Shaffer, D. W. (in press) [B]. Assessing individual contributions to collaborative problem solving: A network analysis approach. *Computers in Human Behavior*.
- Swiecki, Z., & Shaffer, D. W. (2018). Toward a taxonomy of team performance visualization tools. In J. Kay & R. Luckin (Eds.), *Rethinking Learning in the Digital Age: Making the Learning Sciences Count: Vol. III* (pp. 144–151).

ABOUT THE AUTHORS

Zachari Swiecki is a Learning Sciences Ph.D. candidate at the University of Wisconsin–Madison. His work focuses on modeling and visualizing collaborative problem solving.

Andrew R. Ruis is a researcher at the University of Wisconsin–Madison. His work focuses on modeling deliberation and decision making in healthcare and educational contexts.

David Williamson Shaffer is the Vilas Distinguished Achievement Professor of Learning Science at the University of Wisconsin–Madison. His work focuses learning analytics, with an emphasis on models of collaborative problem solving.