

# NETWORK ANALYTICS OF COLLABORATIVE PROBLEM-SOLVING

Simon Kerrigan<sup>1</sup>, Shihui Feng<sup>2</sup>, Rupa Vuthaluru<sup>1</sup>, Dirk Ifenthaler<sup>1,3,\*</sup> and David Gibson<sup>1</sup>  
<sup>1</sup>*Curtin University, Australia*  
<sup>2</sup>*University of Hong Kong, Hong Kong*  
<sup>3</sup>*University of Mannheim, Germany*

## ABSTRACT

Problem-solving and collaboration are regarded as an essential part of 21<sup>st</sup> Century Skills. This study describes a task-focused approach to network analysis of trace data from collaborative problem-solving in a digital learning environment. The analysis framework builds and expands upon previous analyses of social ties as well as discourse analysis and adds new metrics of collaborative learning, problem-solving and personal learning. Using three forms of evidence - actions and use of resources, communications and constructed products - the article outlines and illustrates a framework for characterising individual and team performance on a team project as a basis for documenting individual and team behaviours linked to personal learning, collaboration and team problem solving. This study provides a preliminary demonstration of the effectiveness of network analysis on quantifying and visualizing individual-level and group-level performance in computer-mediated collaborative learning.

## KEYWORDS

Collaborative Problem-Solving, Data Analytics, Network Analysis, Challenge-Based Learning

## 1. INTRODUCTION

*Problems* vary in terms of their structure. Jonassen (1997) classifies problems on a continuum from well-structured to ill-structured. Well-structured problems have a well-defined initial state, a known goal state or solution, and a constrained set of known procedures for solving a class of problems. In contrast, the solutions to ill-structured problems are neither predictable nor convergent because they often possess aspects that are unknown. Additionally, they possess multiple solutions or solution strategies or often no solutions at all (Funke, 2012). Jonassen (2011) reiterates that structuredness of a problem often overlaps with complexity: Ill-structured problems tend to be more complex, especially those emerging from everyday practice, whereas most well-structured problems tend to be less complex. The complexity of a problem is determined by the number of functions, or variables it involves; the degree of connectivity among these variables; the type of functional relationships between these properties; and the stability of the properties of the problem over time (Funke, 1991). Simple problems are composed of few variables, while ill-structured problems may include many variables that may interact in unpredictable ways. When the conditions of a problem change, a person must continuously adapt his or her understanding of the problem while searching for new solutions, because the old solutions may no longer be viable. Static problems are those in which the factors are stable over time while ill-structured problems tend to be more dynamic (Seel, Ifenthaler, & Pimay-Dummer, 2009). Hence, in order to successfully solve complex and ill-structured problems, the person involved in problem-solving must be able to view and simulate the dynamic problem system in its entirety imagining the events that would take place if a particular action were to be performed (Eseryel, Ifenthaler, & Ge, 2013). It has been argued convincingly that all games serve as situated problem-solving environments, in which players are immersed in a culture and way of thinking (Eseryel, Ge, Ifenthaler, & Law, 2011; Gee, 2003).

\*Corresponding author: dirk@ifenthaler.info

*Collaboration* is an essential part in most working environments because it encompasses different views, multiple skills, diverse experiences, analytical judgments, and rich knowledge. Common characteristics of definitions of a collaborative team include at least two involved individuals, common objectives, shared responsibility and interdependence as well as optimal performance (Ifenthaler, 2014). Empirical research shows that through the use of combined resources, teams can successfully handle problems that otherwise would be too complex for a single individual (Badke-Schaub, Neumann, & Lauche, 2011; Cannon-Bowers & Salas, 2001). Digital learning environments, and especially games, designed for team performance, often are characterised by integrated, media-rich contexts with multiple layers of interaction with peers as well as computational resources, which provides a foundation for authentic performance of individual and team-based problem-solving processes with attendant opportunities for unobtrusive observation and documentation of strategies, tools, communications, intentional actions and artefacts (Clarke-Midura, Code, Dede, Mayrath, & Zap, 2012).

A *network* (or graph) is constructed from a set of *vertices* whose relationships are represented by *edges*. Basics of graph theory are necessary to describe the properties of such a network (Diestel, 2000). Various measures from network or graph theory have been applied to assess individual as well as team problem representations and, in addition, to track the development of problem-solving over time (Clariana, 2010). Appropriate structural measures include (a) *number of vertices*, (b) *number of edges*, (c) *connectedness*, (d) *ruggedness*, (e) *diameter*, (f) *number of cycles*, or (g) *average degree of vertices* (Ifenthaler, 2010).

Both, problem-solving and collaboration are regarded as an essential part of 21<sup>st</sup> Century Skills (Griffin, McGaw, & Care, 2012). In this article, we briefly define network measures of personal learning, collaboration and problem-solving and integrate them into a trajectory analysis based on a discrete series of network states of team behaviour evolving during collaborative problem-solving. The case-study illustrates a semester-long collaborative problem-solving task where six teams of three students were engaged, leading to high performing and low performing teams being identified.

## 2. DIMENSIONS OF PERSONAL LEARNING, COLLABORATION AND PROBLEM-SOLVING

The domain model of a learning analysis or assessment is a conceptual representation of the key indicators that experts “might see people say, do, or make as evidence, and situations and activities that evoke it—in short, the elements of assessment arguments” (Mislevy, 2011, p13). For the analysis discussed here, the dimensions of the domain model are personal learning, collaboration and problem solving, which have been defined along with evidence indicators (Gibson, Irving, & Seifert, 2018). We will refer to these as the ‘theory-based evidence targets.’

**Personal learning:** acquisition of knowledge (e.g. new insights, capacities for thinking, acting and employing skills) that is evidenced for outside observers as well as an individual’s own reflection and metacognition (Friedrichs & Gibson, 2003). Evidence targets:

- PL1: Sharing experience
- PL2: Expressing and examining diverse concepts
- PL3: Articulating, applying and building understanding
- PL4: Communicating new powers and creations

**Collaboration:** coordinated group activity resulting from continuous attempts to construct and maintain a shared conception of a problem (Roschelle & Teasley, 1995). *Evidence targets:*

- C1: Establishing and maintaining shared understanding
- C2: Taking appropriate action to solve the problem
- C3: Establishing and maintaining team organization

**Problem solving:** cognitive processing directed at achieving a goal when no solution method is obvious (Mayer & Wittrock, 1996). *Evidence targets:*

- PS1: Exploring and understanding
- PS2: Representing and formulating
- PS3: Planning and executing
- PS4: Monitoring and reflecting

### 3. RESEARCH QUESTIONS

The focus of exploratory data analysis in this research is to determine the challenges and potential of fine-grained time-sensitive analyses of collaborative problem-solving tasks to inform an understanding of the structural, correlational and causal relationships of students achieving learning outcomes. In particular, to what extent can network analyses and related measures assist in the characterisation and prediction of learning processes and learning outcomes (Ifenthaler, 2010)? Guiding the research are five research questions concerning how network analysis can assist in characterising learning in a collaborative problem-solving context:

1. Task Participation – who does what to help the team accomplish its objectives, how team members relate to and divide up the task, and which task activities and outcomes involved which team members?
2. Attention to Feedback for Improvement – how do teams differ in resiliency and the percentage of feedback used to improve, how do teams differ in the type of feedback requested and received
3. Completion Paths – how do teams differ with respect to time to completion, what variability do they exhibit in starting and ending times, and sequence of tasks
4. Use of Time – how do teams differ in their use of time during a long-term project with 24-7 access, which subtasks take the teams more time than others, how do the teams differ in overall time
5. Learning Outcomes – the extent of coverage of outcomes per team member, quality and amount of evidence of achievement of outcomes

### 4. METHOD

#### 4.1 Participants and Context

Participants in the study were  $N = 18$  students in their last year of high school enrolled in a semester-long Vocational Education and Training programme (VET) leading to a certificate in Business Practice with a focus on Health and Workplace Safety and Social Media in Communication. VET programs provide students with learning experiences that are often tailored towards workplace experience, or niche subject content that is not covered in a traditional high school syllabus. Students self-formed into five teams of 3 or 4 members and chose an organisation that they wished to represent in a business scenario. The main task was to research the company and deliver a social media communications plan that effectively educated the company's employees on workplace health and safety legislation rights and responsibilities. The assignment was structured through a series of 17 primary tasks and 76 sub-tasks, referred to as artefacts, which included research, written and design-based work.

The teacher created the project framework, including the design of tasks and sub-tasks linked to learning outcomes, in the Challenge platform, a web-based, mobile-ready application platform for active digital learning experiences and event-level data collection (Gibson & Jakl, 2015). Challenge integrates with Cisco WebEx Teams (<https://www.webex.com/downloads.html>) to provide each team with telecommunications capability for working globally, including a whiteboard, file sharing and teleconference facilities automatically organised by the Challenge platform into the main deliverables in the curriculum design.

The students in this study used the platform to form teams, upload files, chat with team members and complete the assigned tasks for the project. The analyses presented here are based on data collected from student's interactions with the platform, in particular, the creation and submission of artefacts and other inputs required by the tasks, communications among the team members and with the teacher about how to organize the work, and the instructor-judged quality of the team's product as well as the team's self-evaluation of their project. Data for the research team's analyses were collected from log-files and evidence stored on the platform (e.g., uploaded files and the content of page interactions, chat discussions, and written responses to prompts). Analyses and findings of the research team were validated by inspection and protocol review by the instructor as well as by cross-validation of multiple measures presented below.

## 4.2 Data Handling and Analytics

The data used in this study for exploring group collaborative problem-solving was collected from Challenge platform and Webex Teams platforms, merged into one dataset, which was straightforward given the similarities in data structure. Raw transcripts (communications) and trace data (actions and artefacts) were capable of being downloaded at any time for any time frame. Each time that a team member (user) contributed towards an assignment artefact, an interaction transaction was captured by the Challenge platform. Data collected from the Challenge platform included (a) timestamp converted to local time, (b) the user responsible for the interaction, (c) the team of the user, (d) the task they were working on, (e) the artefact they were working on, (f) the content they provided to this artefact and (g) the status of the interaction noted as *visible* (current state), *archived* (saved previous edited version) or *published* (submitted as final state) content. In addition, the communication data among team members was also collected from the Webex Teams platform, including (a) timestamp converted to local time, (b) the user posting the message, (c) the team of the user, and (d) the message content posted to the group. A manually edited column for linking events to tasks and artefacts was manually added and applied to messages in the WebEx Teams data where a student or teacher directly and unambiguously referred to a specific assignment task or artefact. This link allowed analysts to measure the effectiveness and response time to teacher feedback.

The networks modelling the interactions between individual students and artefacts were constructed for analysing individual participation and shared contribution in group collaboration. Two sets of nodes in the network include individual students and task artefacts. The links in the networks represent the interactions between agents and artefacts. There is no link within the same set of nodes (e.g. students to students or artefacts to artefacts) in the network. Bipartite networks, a technique that has been widely used to present the affiliation relationship in social problems, such as personal recommendation (Zhou, Ren, Medo, & Zhang, 2007), were constructed with this approach and used for measuring individual-level and group-level network structures of group collaboration,. Since we are interested in studying the interactive relationships between students and sub-tasks in group collaboration, the number of samples of data per team is the product of the total number of sub-tasks and number of students (e.g. several hundred samples per team). Strength of connections between a student node and an artefact node is considered as a weighted line that summarizes effort (e.g. time and number of interactions) and indicates the relative contribution of a student to an artefact. In addition, three levels of distributions were created to represent artefacts where one, two, or three people had interacted. Implications of the naturally occurring task distributions (e.g. for setting empirical probabilities in future studies) for assessment and social network analysis of collaborative problem solving are under preparation.

## 5. RESULTS

The page limit of the CELDA conference does not allow to present all findings of the above mentioned research questions. The CELDA presentation, however, will include the full coverage of research questions and related findings.

### 5.1 Task Participation

In collaborative learning, group members ideally need to complete key assigned artefacts together in order to achieve the identified learning outcomes. For example, a team cannot acquire or demonstrate any state of collaboration (C1, C2, C3) if they work independently and do not share their work with each other. The visualization of the bipartite networks for a high performing (HP) and low performing (LP) team is presented below (see Figures 1 and 2). The node sets of team members (e.g., person agents – red nodes) are presented in relationship to artefacts classified as 1-person (green nodes), 2-person (yellow nodes) and 3-person (blue nodes) artefacts. In the high performance (HP) team, members worked on more 2-person artefacts compared to the low performance (LP) team. In addition, there was no 3-person artefact in the LP team, in spite of the fact that the relevant team evaluation artefact required all group members to participate.

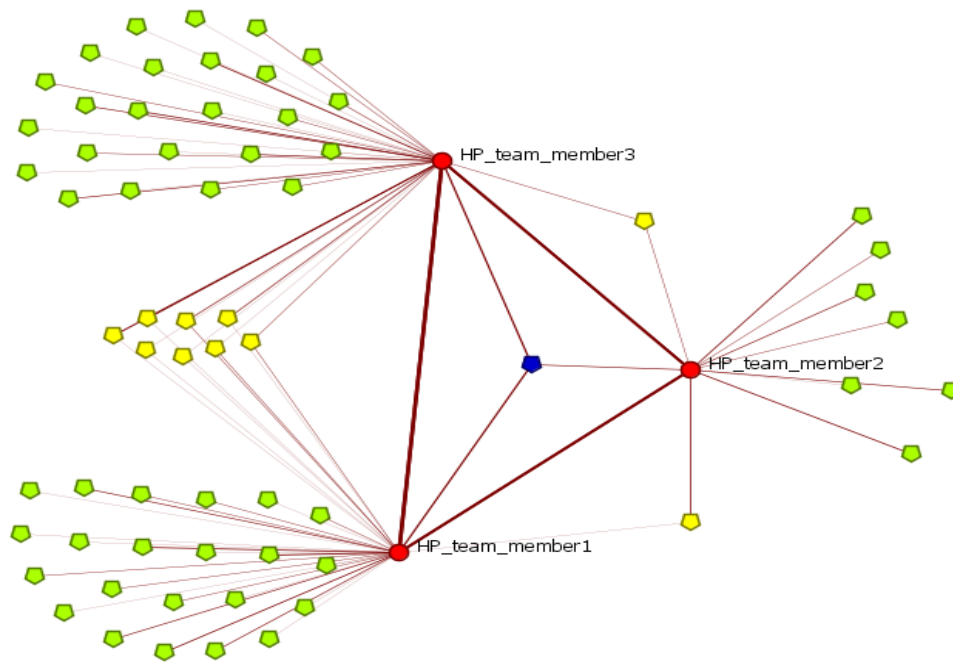


Figure 1. Individual and team task participation in the HP team; semester summary

Examining the extent to which team members worked together on tasks compared to work done on their own, the HP team showed a relatively balanced participation distribution of artefact creation by two members and fewer contributions by a third team member. Incidentally, this appears to be new objective evidence of interpersonal status hierarchies within social expectations states theory (Berger, Cohen, & Zelditch, 1966). Importantly, there were ten instances where HP team members worked on 2-person artefacts and all team members took part in paired production activity. In addition, the HP team self-evaluation included participation by all members. The automatically documented evidence from the Challenge platform is thus strongly linked to the theoretical framework of personal learning in a collaborative problem-solving context.

In comparison, the lowest performing team exhibited a spread of individual workloads among team members but created only 4 artefacts in pairs (Figure 2). There was no instance of all team members working together on an artefact. Team self-evaluation, for example, was ‘filled out’ by only one team member. This suggests that while individual members took some appropriate actions to solve the problem (C2), there is a lack of evidence of an effort to establish a shared understanding (C1) and maintain team organisation (C3), and this is reflected in the team’s overall low performance.

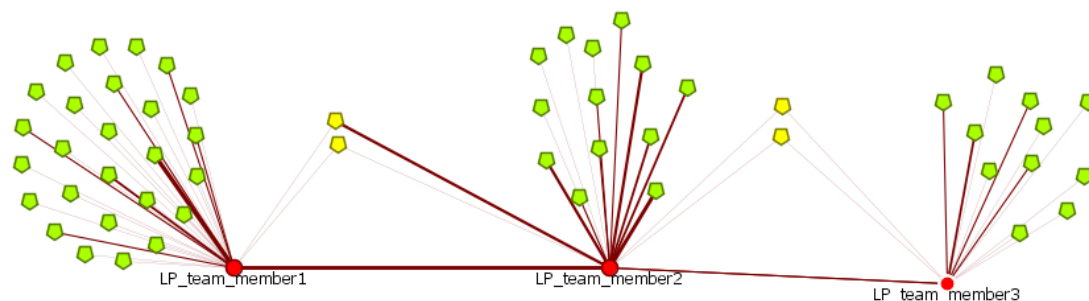


Figure 2. Individual and team task participation in the LP team; semester summary

These network graphs of task participation and distribution (Figure 1 and 2) are summary pictures of the semester-long project, so are missing important dynamic information, which we discuss below in the time series analyses.

## 5.2 Completion Paths

The tasks in Challenge were displayed in a listed order for teams to complete but students were free to start and finish tasks in whatever order they wanted. This provided an opportunity to analyse whether teams differed in their approach to taking appropriate action and planning and executing (C2 and PS3). The general trend in all groups was to start tasks in the order provided by the framework, but the teams exhibited much more variability in the order of completed tasks.

The data concerning completion was computed based on the last time an artefact was touched by any team member, thus capturing the order of any final check by the team. We did not consider the interval from first touch to last touch the actual time on task, because the team could have conducted a last-minute final look at everything. Instead, a time ordered list was created and sequenced with each team’s task interactions (any time the task page was opened or the artefact was edited or uploaded) throughout the project duration. An average duration and sequence order was calculated for each task and subtask, to identify where along the project completion path most of the team’s work occurred. Sorting the tasks by their average sequence value produced an order in which teams started and completed work on the various subtasks (see Figure 3).

The HP team evidenced one of the highest correlations to the benchmark ordering of task in both start and completion order. This suggests that the high performing team methodically approached their work (C3) which may have assisted them in being a high performing team. The LP team on the other hand, exhibited more deviation from the suggested structure in both task start and completion time, and appeared to be less methodical in how they went about completing the tasks; evidenced also by their depth and timing of responses to instructor feedback (C2) and the timing of team member participation (C1). The LP team’s ordering caused some tasks to be completed ‘out of logical order.’ For example, the LP team completed some of the research tasks (Social Media Sites Research) after some of the design-based tasks that were supposed to be research-based (Social Media Summary) suggesting poor organisation and lack of cohesion among the group. This lack of structure (C1) and team cohesiveness (C3) might be a part of why the LP team struggled to create a high-quality final product.

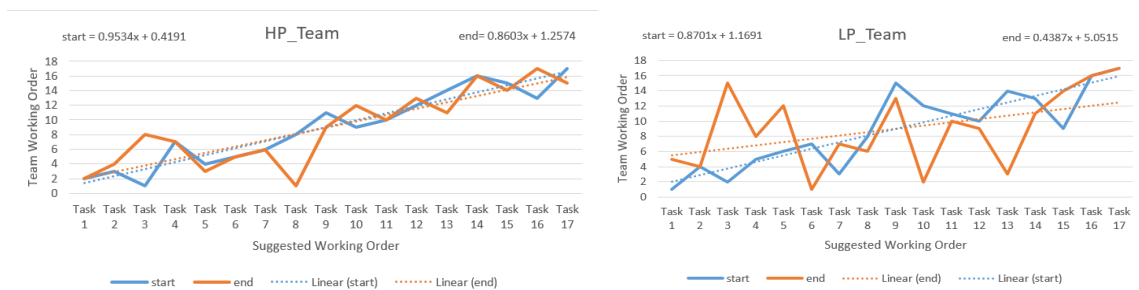


Figure 3. Completion paths of the high performing (HP) versus low performing (LP) team

## 6. DISCUSSION AND CONCLUSION

Network analysis and graph theory have proven to be an appropriate analysis approach for educational applications. Pathfinder and combined techniques (Durso & Coggins, 1990; Schvaneveldt, 1990) provide a reliable representation of knowledge structures and analysis of learning by using pairwise similarity ratings among concepts to create networks. These networks are based on proximity data among entities and are determined by calculating the proximities that best fit within the network. Additionally, graph theory can be applied to almost every area of educational diagnostics. Picard (1980) introduced a promising approach for the design and analysis of questionnaires using graph theory. Furthermore, graph theory has been successfully applied for instructional planning (Hsia, Shie, & Chen, 2008) and evaluation purposes (Xenos & Papadopoulos, 2007).

The current study shows that network-based analyses provide an objective way to represent and evaluate individual participation and contribution during collaborative problem-solving. Network analysis was also found useful for examining the intensity of team-level collaboration by utilising the density property of a bipartite network consisting of agents (team members) and artefacts (team tasks and work products). Furthermore, the analysis of dynamic team networks revealed the periodic changes of individual engagement and group coordination during each stage of a long-term team project, which provided information that in the future could be by instructors to deliver timely intervention and guidance (Ifenthaler, Gibson, & Dobozy, 2018).

Limitations of the study are the low number of participants and teams (even with hundreds of samples over time per team), limited external validity of the findings because the case was limited to one classroom), and labour-intensive manual data processing and analysis. Future research will focus on larger sample sizes and automated analysis techniques.

## ACKNOWLEDGEMENT

This research is supported by Curtin University's UNESCO Chair of Data Science in Higher Education Learning and Teaching (<https://research.curtin.edu.au/unesco/>).

## REFERENCES

- Badke-Schaub, P., Neumann, A., & Lauche, K. (2011). An observation-based method for measuring the sharedness of mental models in teams. In M. Boos, M. O. Kolbe, P. M. Kappeler, & T. Ellwart (Eds.), *Coordination in human and primate groups* (pp. 177–197). Berlin: Springer.
- Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior*, 22(2), 195–202.
- Clariana, R. B. (2010). Deriving individual and group knowledge structure from network diagrams and from essays. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 117–130). New York: Springer.
- Clarke-Midura, J., Code, J., Dede, C., Mayrath, M., & Zap, N. (2012). Thinking outside the bubble: Virtual performance assessments for measuring complex learning. In M. Mayrath, J. Clarke-Midura, D. Robinson, & G. Schraw (Eds.), *Technology-based assessments for 21st century skills* (pp. 125–148). Charlotte, NC: Information Age Publishers.
- Diestel, R. (2000). *Graph theory*. New York, NY: Springer.
- Durso, F. T., & Coggins, K. A. (1990). Graphs in social and psychological sciences: Empirical contributions to Pathfinder. In R. W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization* (pp. 31–51). Norwood, NJ: Ablex Publishing Corporation.
- Eseryel, D., Ge, X., Ifenthaler, D., & Law, V. (2011). Dynamic modeling as cognitive regulation scaffold for complex problem solving skill acquisition in an educational massively multiplayer online game environment. *Journal of Educational Computing Research*, 45(3), 265–287.
- Eseryel, D., Ifenthaler, D., & Ge, X. (2013). Validation study of a method for assessing complex ill-structured problem solving by using causal representations. *Educational Technology Research and Development*, 61(3), 443–463. doi:10.1007/s11423-013-9297-2
- Funke, J. (1991). Solving complex problems: Exploration and control of complex problems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185–222). Hillsdale, NJ: Lawrence Erlbaum.
- Funke, J. (2012). Complex problem solving. In N. M. Seel (Ed.), *The encyclopedia of the sciences of learning* (Vol. 3, pp. 682–685). New York, NY: Springer.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave-Macmillan.
- Gibson, D. C., & Jackl, P. (2015). Theoretical considerations for game-based e-learning analytics. In T. Reiners & L. Wood (Eds.), *Gamification in education and business* (pp. 403–416). New York, NY: Springer.
- Griffin, P., McGaw, B., & Care, E. (Eds.). (2012). *Assessment and teaching of 21st century skills*. Amsterdam: Springer.
- Hsia, T. C., Shie, A. J., & Chen, L. C. (2008). Course planning of extension education to meet market demand by using data mining techniques – an example of Chinkuo technology university in Taiwan. *Expert Systems with Applications*, 34(1), 596–602.

- Ifenthaler, D. (2010). Scope of graphical indices in educational diagnostics. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 213–234). New York, NY: Springer.
- Ifenthaler, D. (2014). Toward automated computer-based visualization and assessment of team-based performance. *Journal of Educational Psychology, 106*(3), 651-665. doi:10.1037/a0035505
- Ifenthaler, D., Gibson, D. C., & Dobozy, E. (2018). Informing learning design through analytics: Applying network graph analysis. *Australasian Journal of Educational Technology, 34*(2), 117–132. doi:10.14742/ajet.3767
- Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development, 45*(1), 65-94.
- Jonassen, D. H. (2011). *Learning to solve problems. A handbook for designing problem-solving learning environments*. New York: Routledge.
- Picard, C. F. (1980). *Graphs and questionnaires*. Amsterdam: North-Holland Publishing Company.
- Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex Publishing Corporation.
- Seel, N. M., Ifenthaler, D., & Pirnay-Dummer, P. (2009). Mental models and problem solving: Technological solutions for measurement and assessment of the development of expertise. In P. Blumschein, W. Hung, D. H. Jonassen, & J. Strobel (Eds.), *Model-based approaches to learning: Using systems models and simulations to improve understanding and problem solving in complex domains* (pp. 17-40). Rotterdam: Sense Publishers.
- Xenos, M., & Papadopoulos, T. (2007). Computer aided evaluation of higher education tutors' performance. *Studies in Educational Evaluation, 33*(2), 175-196.