Measuring the Impact of Interdependence on Individuals during Collaborative Problem-Solving

Zachari Swiecki[1]

Abstract
Collaboration analytics often focuses on assessing and monitoring individuals during collaborative problem-solving (CPS). A defining feature of CPS is the interdependence that exists between individuals when they work together — that is, how they respond to and influence one another over time. While models that account for the impact of interdependence at the individual level of analysis (interdependent models) exist, they are often highly complex. This complexity makes them potentially difficult to use in assessments and systems that need to be explainable for educators, learners, and other researchers. Measures of the impact of interdependence at the individual level of analysis could inform decisions as to whether interdependent models should be used, or whether simpler models will suffice. Such measures could also be used to investigate specific questions about interdependence in collaborative settings. In this paper, I present a novel method of measuring the impact of interdependence on individuals using epistemic network analysis. To provide evidence of the validity of the measure, I compare it to qualitative findings that describe the impact of interdependence on individuals participating in team training scenarios. To demonstrate the value of the measure, I use it to assess the impact of interdependence in these data overall and to test hypotheses regarding the collaborative task design. My results suggest that the measure can distinguish between individuals who have been impacted by interdependence differently, that interdependence is impactful in these data overall, and that aspects of the task design may have affected how some individuals were impacted by interdependence.

Notes for Practice

- Collaboration analytics often focuses on assessing and monitoring individuals during collaborative problem-solving.

- As part of these efforts, researchers, educators, and learners interact with models and representations of collaborative processes.

- Because interdependence is a critical feature of collaborative problem-solving, models that account for interdependence are often necessary. However, these models are complex, making them potentially difficult for researchers, educators, and learners to use.

- To make decisions regarding when interdependent models are necessary, we need a way to measure the impact of interdependence.

- This paper describes the development and testing of a novel quantitative measure of the impact of interdependence that can be used to guide modelling decisions and address specific research questions about collaborative contexts.

---

[1] *Email: zach.swiecki@monash.edu Address: Monash University, Faculty of Information Technology, 29 Ancora Imparo Way, Clayton Campus, Wellington Road, Clayton VIC 3800, Australia. ORCID ID: https://orcid.org/0000-0002-7414-5507*

1. Introduction

Collaborative problem-solving (CPS), the sociocognitive process by which individuals work together to achieve some goal, has long been studied in the learning sciences and psychology. In the former, CPS is a key mechanism for learning. During CPS, people learn by externalizing their thinking (Webb, 2008) and negotiating meaning (Miyake & Kirschner, 2014). In the latter, it

is viewed as an essential process that benefits society (Mathieu, Hollenbeck, van Knippenberg, & Ilgen, 2017). In both, CPS is thought to be a critical skill related to success in the 21st century that should be studied, taught, and assessed (Griffin & Care, 2014). In response, there has been a push to develop assessments of CPS by organizations such as the Organisation for Economic Co-operation and Development (OECD, 2017) and the National Assessment of Educational Progress (Fiore et al., 2017). More recently, researchers have sought to apply advances in learning analytics to the study of CPS via *collaboration analytics* (Martinez-Maldonado, Kay, Buckingham Shum, & Yacef, 2019), with the goal of leveraging collaboration data, machine learning, statistics, and interfaces to automatically assess and support collaborative process and learning.

Crucial to assessments of both CPS and collaboration analytics are models that describe the processes of groups and teams. Two general approaches have developed for modelling collaboration in terms of the units of analysis being studied. Researchers have focused on the group or team as the unit of analysis — for example, Johnston, Poirier, and Smith-Jentsch (1998) developed observation protocols to assess the performance of military teams on dimensions such as communication quality and leadership — and they have focused on individuals as the unit of analysis — for example, Stylianou-Georgiou, Pananstiou, and Puntambekar (2011) measured individuals' metacognitive strategies before and after CPS scenarios.

While it may seem more natural to consider the team when examining CPS, there are several reasons for choosing to focus on individuals. As Reimann (2009) argues, important changes in learning and behaviour occur at the individual level. Modelling these changes could inform feedback for individuals, improving their performance in team settings. Furthermore, individuals on teams may be assigned to or come to hold roles — coherent and expected sets of behaviours (Forsyth, 2009). Roles mean that there are different expectations for different individuals on the team. Therefore, a team lens would ignore these differences. Finally, and most practically, there are always more individuals than teams, meaning that focusing on individuals gives researchers more units of analysis to work with, affording certain statistical advantages.

When researchers focus on individuals in collaborative settings, the team is often treated as a contextual variable that needs to be controlled for in order to isolate the processes and performance of individuals (Stahl, Law, Cress, & Ludvigsen, 2014). Such models obfuscate any details about the moment-to-moment interactions between the individual and their team, how individuals may depend on one another for information, and how they may influence one another over time. In other words, they ignore *interdependence*. However, models have been developed that can account for these relationships. These *interdependent models* seek not to isolate individuals from their team but rather to describe how individuals think and act in relation to them.

While interdependent models have been found to differ significantly from *independent* models — that is, those that treat the process of individuals as isolated events by not accounting for how they relate to prior team behaviour — and to align better with qualitative descriptions of collaboration (Swiecki, Ruis, Farrell, & Shaffer, 2020), they tend to be more complex. In turn, they may be more difficult to interpret and apply. Decisions between models that are more valid as opposed to simpler are of the utmost practical importance. On the one hand, models are used to evaluate and train individuals in a variety of contexts. If these models are not explainable, that is, if educators, teams, and other researchers cannot understand these models because they are too complex, their utility is limited (Rummel, Walker, & Aleven, 2016). On the other hand, models are also used as part of assessments that help determine whether individuals are qualified for educational or professional opportunities. In these cases, validity is crucial.

This trade-off between model complexity and validity suggests that measures that describe the impact of interdependence on individuals could be valuable. For example, such measures could help researchers determine when interdependent models are necessary, or when simpler, independent models will suffice. In addition, measuring the impact of interdependence on individuals may be useful for investigating collaboration more generally. For example, many collaborative learning tasks are designed to promote interdependence. One of the most well known designs for doing so is the jigsaw (Aronson, 1978), which deliberately exposes individuals to different information or resources. The jigsaw design implies that individuals will need to coordinate information and behaviour to achieve some goal; however, it does not guarantee it. A measure that describes the impact of interdependence on individuals during the collaborative task could be a useful way to assess the effectiveness of the task design and inform redesign. Current methods exist for measuring the impact of interdependence, but these methods are limited by the facets of interdependence they account for, their reliance on specific kinds of data, or their statistical convenience.

In sum, collaboration analytics often focuses on assessing and monitoring individuals during CPS. As part of these efforts, researchers, educators, and learners interact with models and representations of collaborative processes. Because interdependence is a critical feature of CPS, models that account for interdependence are often necessary. However, these models are complex, making them potentially difficult to use and explain. To make decisions regarding when interdependent models are necessary, we need a way to measure the impact of interdependence. In this paper, I provide a proof of concept for a novel quantitative measure of the impact of interdependence that can be used to guide modelling decisions and address specific research questions about collaborative contexts.

To develop the measure, I examine team discourse data collected from military training scenarios. This same dataset has been used before to examine the differences between independent and interdependent models (Swiecki et al., 2020). This paper extends that prior work by using these differences to derive a novel quantitative measure of the impact that interdependence may have on individuals in collaborative settings. To provide evidence of the validity of the measure, I compare it to qualitative findings that describe the impact of interdependence on individuals participating in the team training scenarios. To demonstrate

the value of the measure, I use it to assess the impact of interdependence in these data overall and to test hypotheses regarding the collaborative task design.

My results suggest that the measure can distinguish between individuals who have been impacted by interdependence differently, that interdependence is impactful in these data overall, and that aspects of the task design may have affected the impact of interdependence on some individuals. Together, the results demonstrate a measure of the impact of interdependence that can be used to investigate whether interdependent models are necessary and how task designs affect collaborative processes.

2. Theory

2.1 Interdependence

Interdependence can be described in terms of the tasks teams undertake, the behaviours and information they coordinate, and how their interactions cluster in time. Interdependence is impactful in the sense that it results in shared understandings, as well as stable cognitive and behavioural states that develop over time and influence individual actions and decisions. Hutchins's (1995) analysis of navigation teams in the US Navy illustrates these facets of interdependence in a CPS setting.

2.1.1 Navigation Teams: An Example

In addition to using Global Positioning Systems, navigation teams in the US Navy manually calculate and monitor the position of their ship at sea. In open water, teams calculate, or *fix*, the ship's location relatively infrequently; the course of the ship tends to be steady and, due to the lack of obstacles, there is a high margin for error. In fact, in open water, a team is not even required; the position of the ship can be fixed at a relaxed pace by a single officer.

At port, however, things are different. Due to the mass of the ship and the confines of the environment, changes in trajectory need to be planned far in advance and the fix needs to be taken at frequent intervals. In these situations, the demands of the task are such that a team of officers is needed.

To fix the ship's position, a team member known as the *bearing recorder* sends a message to one or more *pelorus operators*. The job of the pelorus operators is to take measurements of the angle between the ship's forward line of motion and a given set of landmarks, otherwise known as *bearings*. The tool used to take bearings is called, naturally, the *pelorus*, and taking bearing measurements is called *marking*. Once a bearing is marked, the pelorus operators relay the measurements to the bearing recorder, who, in turn, records them for use by the *plotter* — the individual responsible for fixing the position of the ship on the navigational chart and projecting its future positions. After plotting the fix, the plotter restarts the cycle by choosing new landmarks and directing the bearing recorder to request new marks from the pelorus operators.

2.1.2 Kinds of Interdependence

The navigation example makes clear that the interactions that make up the performance of this navigation team are not merely a collection of isolated actions. Rather, in accomplishing its work, the navigation team is interdependent in at least four ways:

1. *Task*. Van de Ven, Delbecq, and Koenig Jr. (1976) describe task interdependence in terms of the sequencing of tasks. This form of interdependence occurs when the completion of some or all of the subcomponents of a task depends on the completion of the other subcomponents. For example, the plotter cannot plot without first receiving the bearings from the recorder, who cannot record until the bearings are relayed by the pelorus operators, who cannot take the bearings until the plotter chooses the original landmarks.

2. *Behavioural*. DeChurch and Mesmer-Magnus (2010) argue that interdependence can also be defined in terms of what is being coordinated during team activity. Teams have behavioural interdependence when they have to coordinate physical actions. For example, navigation teams need to coordinate actions such as marking and plotting bearings.

3. *Informational*. DeChurch and Mesmer-Magnus's (2010) alternative to behavioural interdependence is informational interdependence, which occurs when team members have to coordinate the knowledge or opinions possessed by different members of the team. In addition to coordinating actions, navigation teams have to coordinate specific information, like bearing measurements and landmark choices.

4. *Temporal*. Halpin and von Davier (2017) argue that interdependence can be defined in terms of the timing of team events. Temporal interdependence, or temporal clustering as they call it, occurs when the actions of some individual(s) on the team make actions by others more or less likely in the near future. For example, the call to mark by the bearing recorder makes it very likely that the pelorus operators will begin taking bearings soon after.

Both information and behaviour can be construed more broadly as *discourse* (Gee, 1999), or the things people say, do, and think that identify them as part of a particular culture, say navigation teams. Taking this view, behaviour and information are ontologically consistent and can be accounted for simultaneously in analyses of collaboration. In what follows, I use the term

interdependence to summarize behavioural, informational, and temporal interdependence and separate the term into temporal versus *discursive* (i.e., behaviour/information) interdependence only when needed.

As several researchers argue (Kozlowski & Ilgen, 2006; Marks, Mathieu, & Zaccaro, 2001), interdependence is made possible by, and gives rise to, stable behavioural, cognitive, and motivational states through repeated interaction. Team cognitive structures, such as shared mental models, are examples of emergent states (DeChurch & Mesmer-Magnus, 2010). As they are formed, these states come to influence subsequent team interactions. This influence can be described by what Clark (1996) calls the *common ground* — the set of shared knowledge and experiences that exist between people when they interact. The contents of the common ground influence subsequent actions and how those actions are interpreted (Dillenbourg, 1999). Common ground and similar ideas, such as shared understandings (Miyake & Kirschner, 2014) and joint problem spaces (Roschelle, 1992), have been used to explain how team processes unfold.

While emergent states and common ground are impactful, empirical evidence suggests that their impact may not span the entire history of team interaction. For example, Shaffer (2017), following Suthers and Desiato (2012), argues that discourse is interpreted with respect to the *recent temporal context*, or the set of interactions that occurred within a recent span of time. Building on this work, Ruis, Siebert-Evenstone, Pozen, Eagan, and Shaffer (2019) found that individuals in collaborative discussions tended to respond to — and were thus influenced by — prior contributions that occurred within relatively short spans of talk. Similarly, Halpin and von Davier (2017) found that individuals in a collaborative setting had actions that were temporally clustered — and as they argued, depended on one another — only within short spans of time.

The descriptions above support the idea that the microgenetic (Chinn & Sherin, 2014), moment-by-moment interactions of individuals during collaboration create a context in which these individuals are impacted by interdependence over time. In turn, they suggest that interdependence is a critical feature of CPS at the individual level of analysis.

2.2 Modelling Interdependence

Several extant modelling techniques account for the nature of interdependence at the individual level of analysis in different ways[1]. Due to theoretical descriptions of CPS as a combination of social and cognitive dimensions (see, e.g., OECD (2017)), I have chosen to discuss only those methods that are explicitly sensitive to both the social (who interacts with whom) and cognitive (what they interact about) structures of the setting.

2.2.1 Transactive Coding

Prior to analysis with statistical or machine learning techniques, collaborative discourse data are often *coded*. Coding involves deciding on salient concepts or behaviours that researchers wish to identify and specifying rules by which the codes are evidenced in the data. The application of the rules may be done manually or automatically via an algorithm.

Several techniques use coded data. One technique, broadly referred to as *coding-and-counting*, has been widely used to investigate collaborative process (Chi, 1997). Coding-and-counting models count the qualitative codes, scores, or observer ratings attributed to individuals and use these counts to make inferences about collaborative processes or outcomes. This technique models collaboration using the frequency of individuals' contributions without reference to the contributions of others or when they occurred. While popular, it has been criticized because it is not sensitive to the microgenetic interactions that unfold between individuals (Csanadi, Eagan, Kollar, Shaffer, & Fischer, 2018).

One extension of coding-and-counting seeks to address this issue by only coding discourse that is related to or dependent on other discourse in a particular way. A typical example of such an approach is *transactive coding*. Transactivity refers to the ways in which individuals in collaborative situations respond to and build upon one another's text or speech contributions to different degrees (Wang, Wen, & Rosé, 2017). For example, a contribution may exhibit no transactivity if it does not refer to a prior contribution in any way; a contribution may exhibit high transactivity if it directly replies to a prior contribution, summarizes the content of the contribution, and adds new content. Critically, coding the contributions of individuals for degrees of transactivity requires attending to the relationships that exist between an individual's contribution and the contributions of others, and thus seeks to account for interdependence at the individual level of analysis.

2.2.2 Dynamic Bayesian Networks

Bayesian networks are machine learning techniques that describe the conditional probabilities between related states or events, which can be expressed in terms of codes applied to collaborative discourse data or other measured features of the data (Witten, Frank, Hall, & Pal, 2016). Traditional Bayesian networks operate under a Markov assumption that subsequent states or events are only influenced by the current state or event. Dynamic Bayesian networks (DBNs) relax this assumption and allow for the investigation of relationships between states or events at arbitrary time lags. With such networks, the relationship of the state of an individual at a given time to the states of other individuals at different times can be accounted for. For example, Gweon, Jain, McDonough, Raj, and Rosé (2013) used DBNs to estimate the prevalence of transactive discourse during a collaborative discussion.

---

[1] There are, of course, other techniques that could be used to investigate interdependence. The broad field of discourse analysis provides several examples. The intent of the list discussed here is not be exhaustive but to highlight well-known methods that could potentially be automated and integrated into collaboration analytics systems.

2.2.3 Cohesion Network Analysis

Cohesion refers to the presence of explicit lexical, grammatical, or semantic cues in textual discourse that indicate connections among ideas. Multiple approaches have been used to measure cohesion, such as the frequency of cue words and phrases, referring expressions, semantic overlap, and semantic similarity (Dascalu, McNamara, Trausan-Matu, & Allen, 2018). Cohesion network analysis (CNA) uses automated measures of cohesion to generate network models that represent interactions between individuals in collaborative settings. In contrast to social network analysis, edges in cohesion graphs indicate not only that particular individuals interacted but also the extent to which their contributions were semantically similar.

2.2.4 Traces

The traces framework (Suthers, 2017) uses four network representations to represent the relationships between individuals in collaborative settings. First, *contingency graphs* are generated that identify how events are related. Contingencies may be defined due to temporal proximity, spatial proximity, address, or reply states, originating from the same actor, or overlapping in content, as measured by overlapping word stems between contributions. These graphs are then aggregated to create uptake graphs. Similar to transactivity, *uptake* refers to the extent to which new contributions take up or include content from prior contributions. Uptake graphs group various contingencies into a single edge. Session graphs cluster uptake graphs to show relationships within and between larger spans of activity. And, finally, sociograms (social networks) are created that show the strength of the relationships between individuals in collaborative settings via the sum of the strength of the uptake between their contributions.

2.2.5 Group Communication Analysis

Group communication analysis (GCA) builds on cohesion and contingency analyses by using automated computational linguistics and analyses of sequential interactions to characterize the CPS processes of individuals (Dowell, Nixon, & Graesser, 2019). GCA combines several sociocognitive metrics, such as participation, internal cohesion, responsivity, social impact, newness, and communication density, to identify coherent sets of behaviours that individuals enact in CPS settings. More specifically, GCA uses cross-correlation measures at different lag values to identify how related the social (e.g., participation levels) and cognitive (e.g., measured by latent semantic analysis (Landauer, Foltz, & Laham, 1998)) interactions of individual teammates are.

2.2.6 Epistemic Network Analysis

Epistemic network analysis (ENA) (Shaffer, Collier, & Ruis, 2016) creates networks that describe interactions among individuals' discourse in collaborative settings. Prior to analysis with ENA, discourse is coded for semantically meaningful concepts in the domain, typically derived from a grounded qualitative analysis of the data (Glaser & Strauss, 1967). ENA models the relationships between these concepts by measuring the co-occurrence rates of codes within a moving window over the data. In this way, ENA can model how an individual's discourse is connected to — that is, dependent on — the discourse of others over time (Siebert-Evenstone et al., 2017).

2.3 Validity and Complexity

Prior work by Swiecki and colleagues (2020) found that *interdependent models* — those that account for the microgenetic nature of interdependence at the individual level like those described above — may more accurately describe collaborative processes than *independent models* — those that do not account for the relationships that unfold over time between individuals during collaboration. While interdependent models are potentially more accurate, they have a cost: they may require more complex mathematical and algorithmic techniques and use more complex visualizations than other methods. This complexity may make them more difficult to understand and use for researchers, educators, and teams (Rummel et al., 2016).

For example, compare and contrast interdependent models with traditional coding-and-counting — a commonly applied independent model. While coding-and-counting ignores interdependence, the technique is relatively simple. To apply it, one only needs to identify the categories of interest and count their frequency. In contrast, in addition to identifying categories of interest, interdependent models require careful decisions about data segmentation and what counts as a relationship between contributions. And, they tend to use more sophisticated computational techniques, such as lagged similarity scores, semantic overlap, conditional probability distributions, or normalized co-occurrence counts.

Model complexity, or parsimony, has long been a criterion used to evaluate models. In statistics, for instance, parsimony refers to explaining a given phenomenon using the fewest parameters possible. This notion is fundamental to model performance metrics, such as the Akaike information criterion (Burnham & Anderson, 2004), that balance model fit and model simplicity. Parsimony is useful because it reduces the chance of the model being too specific to a single context, that is, overfit (Witten et al., 2016).

While there are statistical benefits to parsimonious models, they have conceptual benefits as well. Parsimonious models tend to be easier to interpret and, thus, easier to explain and use outside (and often inside) research contexts. Imagine explaining kinematics to a novice. A complex, though more valid, model to use would be the quantum-mechanical model, which describes

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

79

the motions and interactions of particles in terms of waves (not to mention dead (?) cats)[2]. A simpler and, at times, still valid model would be Newtonian mechanics, which has the advantage of dealing with particle motion at the macro scale without considering waves.

While more complex models may be better aligned with the construct of interest, their use implies that a trade-off between validity and complexity may exist: in our efforts to use more valid models, we may introduce a level of complexity that makes those models difficult to apply and understand. In the context of modelling individuals in CPS settings, a similar trade-off can be seen in the use of multi-level analysis (MLA).

2.3.1 MLA

MLA techniques, specifically, hierarchical linear models (HLMs) (Snijders & Bosker, 2012), are statistical techniques designed to account for dependent observations. In general, techniques like HLM are more complex mathematically and conceptually than standard linear techniques. Despite their complexity, MLA techniques are often needed to make valid inferences about individuals in team settings (Stylianou-Georgiou et al., 2011).

As Janssen and colleagues (2011) describe it, in the MLA framework, interdependence arises because of the hierarchical nature of data collected from collaborative settings. For example, individuals are nested into teams, which in turn may be nested into classrooms that are themselves nested into different schools. Nesting means that the dependent variables we want to measure for individuals in collaborative contexts, such as outcome or process measures, are interdependent in the sense that individuals on a team, for example, influence one another in ways that make them either more similar to one another or more different from one another than we would otherwise expect. This means that our observations may not be independent, a key assumption of many statistical techniques.

Similar to interdependent models, the nature of collaborative settings suggests that using HLMs is always more appropriate than using standard linear models. In the same way that the hypothetical presence of interdependence suggests that interdependent models should be used to understand CPS, nesting in data suggests MLA techniques should be used. However, in many cases, researchers make the explicit choice not to use HLMs. Like interdependent models, HLMs are conceptually and mathematically more complex than their alternative, but unlike interdependent models, the use of HLM methods is mature enough to have concrete and measurable criteria with which to inform decisions about their use. One such measure is the intraclass-correlation coefficient (ICC). The higher the ICC, the more interdependent the team. A statistical test can be performed to investigate whether a given ICC value is statistically significant. If the test is significant, it suggests that interdependence is present and should not be ignored. Thus, there are situations in which nesting and interdependence theoretically occur in the data but can be safely ignored statistically. In turn, if a researcher finds a non-significant ICC value, they may opt to use a less complex modelling technique.

2.4 Measuring the Impact of Interdependence

The ICC demonstrates that there exist criteria for deciding whether to use HLMs or simpler techniques. Unfortunately, the ICC is not appropriate for determining the impact of interdependence on individuals as described here. HLMs are designed to account for interdependence by controlling for the individual and team effects on the dependent variable(s) of interest. HLMs (and other techniques, like analysis of covariance) *isolate* the contributions of individuals and teams; they treat teams as a contextual variable that should be controlled for. In turn, team effects are treated as if they have a continuous influence over time. But as theoretical constructs like recent temporal context and temporal interdependence suggest, the influence of the team on the individual is dynamic. In contrast, interdependent models seek to *describe* the microgenetic relationships that exist between individuals and their teams. Thus, the goals of the two approaches are distinct and, in fact, complementary: interdependent models can be used to measure individual variables that account for the relationships between individuals and their teammates; HLMs can be used to make comparisons between those variables in a statistically valid manner.

At least two methods exist for measuring the impact of interdependence at the individual level that are sensitive to the microgenetic nature of CPS. However, both have important limitations.

2.4.1 Lag Sequential Analysis

Lag sequential analysis (LsA) (Bakeman & Gottman, 1997) is a technique that describes the transition probabilities of events or states. For example, Kapur (2011) used the technique to compare the processes of teams that solved either well-formed or ill-formed problems. LsA allows for the calculation of transition probabilities at different lag values and incorporates statistical tests for determining whether a given transition is significant. In this way, statistical tests of transitions at different lag values can be computed and, if found significant, can indicate that interdependence is impactful. Unfortunately, in a given analysis, the number of salient transitions at different lag values can be high, and determining the impact of interdependence requires significance tests for each transition (Olson, Herbsleb, & Rueter, 1994). As such, it is important to control for family-wise error during these tests. As the number of tests increases, it becomes more difficult to find significant results. This makes using LsA statistically inconvenient for testing the impact of interdependence.

---

[2]See Schrödinger (1935).

2.4.2 Point Process Models

Point process models use information about the temporal clustering of team actions to measure interdependence. For example, Halpin and von Davier (2017) used point process models to measure the interdependence between teammates on a basketball team. Based on the clustering of event times, the technique determines time frames within which individuals are dependent on others and determines whether this dependence is statistically significant. While this technique is a powerful tool for measuring the impact of interdependence, at least in current implementations, it relies solely on event times to detect temporal interdependence and ignores discursive interdependence. This means that it does not account for a defining aspect of interdependence and may be inappropriate for collaborative contexts that unfold over larger spans of time and where participants are free to interact at their leisure, such as discussion boards. In situations like these, timing of contributions seems less important than their content.

2.4.3 What We Need

Interdependence is defined in terms of temporal clustering of activity and the coordination of information and behaviour. It is expressed in the microgenetic interactions between individuals. And as the example of navigation teams illustrates, these interactions may occur in a variety of modes beyond just speech or text.

Together, these characteristics of interdependence and the limitations of existing measures suggest the features needed for a valid and useful measure of the impact of interdependence. In particular, such a measure should be (a) sensitive to both discursive and temporal interdependence; (b) sensitive to microgenetic interactions between individuals, that is, how they occur and change over time; (c) data agnostic, because interdependence can be expressed in terms of phenomena that go beyond speech or text to include ways of behaving, signalling, and manipulating tools; and (d) statistically convenient, meaning that it should not require multiple tests to determine whether the measure is significant.

Such a measure has at least two uses. First, as described above, it could be used to make principled decisions about when interdependent models are strictly necessary. These decisions may impact the explainability of results that are reported by researchers or collaboration analytics systems that support educators and learners.

Second, it could also be used to address specific questions about collaboration. For instance, many collaborative tasks in the real world exhibit task interdependence. One example is the navigation task described above, but there are, of course, other examples, including engineering design, medical care, and decision-making. In such contexts, a measure of the impact of interdependence on individuals could be used to investigate whether individuals are actually impacted by interdependence in ways suggested by the task. For example, in the navigation task, the measure could be used to evaluate whether information and behaviour are actually being coordinated among teammates.

In educational contexts, collaborative learning tasks may be designed to promote interdependence. For example, the jigsaw paradigm (Aronson, 1978) is designed to create differential expertise among group members. An intended effect of these differences is to foster information sharing and negotiation among individuals. A measure of the impact of interdependence could be used to determine whether task designs such as these are effective.

In both real-world and educational collaboration, tasks may have individuals that occupy specific roles. Roles imply an asymmetry in knowledge and status within the group (Dillenbourg, 1999). This asymmetry implies that interdependence should exist — individuals will need to coordinate information and behaviour to succeed because no one has, or few people have, complete knowledge of the task or access to all resources. A measure of the impact of interdependence could provide evidence as to whether expected roles are being followed, which, in turn, could inform pedagogical interventions.

In sum, interdependence between individuals is likely in most CPS tasks, suggesting that interdependent models should be used when analyzing these contexts. However, interdependent models are often complex, making them potentially difficult to use in collaboration analytics systems or assessments that support educators and teams. While interdependence theoretically exists in CPS settings, the example of MLA suggests that measures can be used to determine whether the impact of interdependence is great enough to warrant an interdependent model. Such a measure could help to inform modelling decisions that researchers make (which ultimately affect educators and teams), and it could be used to investigate specific questions regarding CPS contexts. While some measures of the impact of interdependence at the individual level exist, they have important limitations. This work aims to introduce a measure that addresses these limitations.

2.5 Research Questions

One way to measure the impact of interdependence is to compare the results of interdependent and independent models for the same individuals using a similarity metric. Similarity metrics quantify the similarity (or difference) between outcomes of interest. Several of these metrics exist, including correlations, loss functions, and cosine similarity (Witten et al., 2016). In the context of interdependence, such metrics can be used to compare interdependent and independent models and thus derive a measure of the impact of interdependence. If this measure indicates that interdependent and independent models are highly similar or exactly the same, this would suggest that interdependence is not impactful. However, if it indicates that differences between the models do exist, it may suggest that interdependence is impactful. To make this determination, we can apply a significance test to the measure to investigate whether the difference is statistically meaningful. In this way, the similarity

between interdependent and independent models is a measure that suggests both whether interdependence between individuals is impactful and the extent to which it is.

The discussion of models of interdependence in Section 2.2 suggests that there are several models that may be used to derive a measure of the impact of interdependence. Whichever model is chosen should be able to produce a measure that meets the criteria of being (a) sensitive to discursive and temporal interdependence, (b) sensitive to microgenetic interactions, (c) data agnostic, and (d) statistically convenient.

Here, I chose to derive the measure using ENA because it meets these criteria in the following ways. ENA identifies connections within coded discourse only if those connections occur within a prespecified span of time defined by a moving window over the data. Because the codes may refer to either information that is passed or actions that are taken within relatively short spans of time, it can account for both discursive and temporal interdependence. The use of the moving window also means that ENA is sensitive to the moment-by-moment interaction between individuals and how they change over time. ENA is also data agnostic. The kind of data or the medium from which they are collected does not preclude the analyses. The only requirement is that the data be coded. And, finally, ENA produces network representations that can be used to conduct holistic tests between model results. In other words, the entire network can be tested for statistical significance in addition to, or instead of, testing specific connections within the network.

While other models could potentially be used to derive a measure of the impact of interdependence[3], I also chose to use ENA in this study for two additional reasons. First, as described in the Methods section below, the way in which ENA identifies co-occurrences means that independent and interdependent models can be created for the same data simply by varying the segmentation of the data being analyzed. This means that ENA can be used to compare interdependent and independent models using the same modelling framework, thus reducing potentially confounding variables introduced by using different techniques.

Second, ENA's use of discourse codes and network visualizations means that its results may be more explainable, and therefore more useful in collaborative analytic systems employed in the wild. Other techniques, such as GCA, lack an integrated visualization, while cohesion networks and contingency graphs, for example, define connections via abstract constructs like uptake and semantic similarity. In contrast, ENA defines connections via discourse codes that are, most often, developed from a grounded qualitative analysis of the data. These codes may be defined in terms of *emic* concepts — those that reflect how members of the community being studied understand their own actions — meaning that they may be more easily understood by non-research audiences (Shaffer, 2017).

2.5.1 Novel Contributions

To test the proposed measure using ENA, I analyzed data from the Tactical Decision Making Under Stress Project (TADMUS) (Johnston et al., 1998). The TADMUS data were collected as part of a study that assessed US Navy air defence warfare teams during training. Teams participated in a medium-fidelity simulation in which they were tasked with managing potentially hostile threats during littoral warfare scenarios. Specifically, they were tasked with collecting information on local aircraft and ships, assessing their threat level, and taking appropriate action to defend their ship.

ENA has been used to investigate interdependence using these data in a prior study. In particular, the study by Swiecki and colleagues (2020) compared an interdependent model created with ENA and an independent model created by combining coding-and-counting with principal components analysis. They found that an interdependent model captured important aspects of individuals' collaborative process that the independent model missed.

The current study extends this prior work in at least three important ways. First, the prior work focused on demonstrating how ENA as a modelling technique could account for the impact that interdependence may have on individuals in collaborative settings. The current study focuses on developing, testing, and demonstrating the value of a quantitative measure of that impact. Put another way, the prior work provided some evidence that interdependence can be impactful; the current work describes a way to *measure* that impact. Thus, this work has a different — and more specific — focus.

Second, to demonstrate the importance of accounting for interdependence, the prior work compared an interdependent ENA model with an independent coding-and-counting model. Here, to derive a quantitative measure of the impact of interdependence, I compare interdependent and independent models that are both based on ENA. This streamlines the construction of the measure and removes possible confounds that could arise from comparing the results of fundamentally different modelling approaches. Thus, the methodological approach taken in this work is distinct.

2.5.2 RQ 1

While a measure of the impact of interdependence may be useful, to ensure that the measure works as intended, we need evidence of its validity. Evaluating the validity of a measure involves finding empirical evidence that supports the relationship between the measure and the theory informing it (Kane, 2006). In practice, this evidence can be obtained by comparing the measure to other (hopefully) valid interpretations of the data. The quantitative ethnography framework (Shaffer, 2017) outlines a process for making such comparisons via *closing the interpretive loop*. Closing the loop entails comparing quantitative measures of phenomena to qualitative interpretations of those same phenomena. It is a process for determining whether

---

[3]The implications of using models other than ENA are considered in the Discussion section of the paper.

the quantitative measure is actually describing the phenomena of interest. In this sense, closing the loop is similar to data triangulation, which seeks to corroborate interpretations via multiple sources. In turn, the first research question I address is as follows:

RQ1: *Does the measure align with individuals qualitatively determined to be impacted differently by interdependence?*

2.5.3 RQ 2

As described in more detail in the Methods and Results sections, the air defence warfare task is, theoretically, highly interdependent. Individual team members must work together to pass and manage information and make decisions under strenuous conditions. As such, the derived measure should suggest that interdependence is impactful. My second research question addresses this hypothesis:

RQ 2: *Does the measure suggest that interdependence is impactful in this dataset overall?*

2.5.4 RQ 3

Features of the collaborative task suggest that the impact of interdependence might vary within the data. Specifically, individuals in this context were assigned to particular roles. Two individuals held command roles, while the remaining four held supporting roles. This role structure suggests that information and behaviour needed to be coordinated, but not necessarily equally across roles. Because those in command roles needed to coordinate information provided by those in support roles to make tactical decisions, they may have been more impacted by interdependence.

Furthermore, teams in this context were divided into two conditions. Commanders in the experimental condition had access to a decision-support system (DSS) designed to support their awareness of the situation. The DSS provided information directly to the commanders that was normally the purview of supporting members of the team. The direct availability of information to the commanders that is typically provided and re-provided by supporting members of the team suggests that individuals in the experimental condition might be less impacted (act more more independently) than those in the control condition. My last research question addresses these hypotheses regarding the role structure and experimental design:

RQ 3: *Does the measure suggest that individuals in command roles are more impacted by interdependence than those in support roles and that those in the experimental condition are less impacted than those in the control condition?*

3. Methods

3.1 Data

Simulated scenarios with 94 participants in 16 teams were used to test the impact of a DSS and teamwork training on team performance. During the scenarios, teams performed the *detect-to-engage sequence*, which entailed using computer-based tools (*watch stations*) to detect and identify ships and aircraft in the vicinity (referred to as *tracks*), assess whether they were threats, and decide whether to respond with warnings or combat orders.

Each team consisted of six officers assigned to particular roles. Two officers held command roles, and four held supporting roles. The command roles were the commanding officer (CO) and the tactical action officer (TAO). They were responsible for making tactical decisions, such as when to warn or engage tracks. Support roles were the identification supervisor (IDS), the air defence warfare coordinator (ADWC), the tactical information coordinator (TIC), and the electronic warfare supervisor (EWS). They were responsible for reporting critical information to commanders, such as the detection of threats; their identification — for example, whether the threat was a jet, ship, or helicopter; and their behaviour — for example, their speed and location relative to the warship. Supporting officers were also responsible for communicating with threats; sending warnings; and, in some cases, recommending tactical actions.

The teams were divided into two conditions, with eight teams in each condition. Teams in the control condition had access to standard watch stations. Teams in the experimental condition received additional training on critical thinking, teamwork, and communication skills, and had access to watch stations upgraded with the DSS. The DSS provided information about track identification, behaviour, and threat levels, as well as a history of actions taken toward the tracks. As shown in Figure 1, the DSS consisted of five visual displays:

1. The Track Summary Module summarized data related to a selected track, including type, threat assessment, bearing, range, speed, altitude, and electronic emissions.

2. The Track Profile Module provided critical information about a selected track, including current position of the track in relation to the local geography, track altitude, speed, and the range of the weapons carried by the track.

3. The Response Manager Module served as a prompt for upcoming actions and a memory aid for actions already taken, such as deterrent or defensive orders.

4. The Basis for Assessment Module presented historical track data as evidence for classification as either hostile or friendly.

5. The Track Priority Module listed tracks in order of potential threat as determined by an algorithm that combined measures of speed, altitude, direction, and assumed armament.
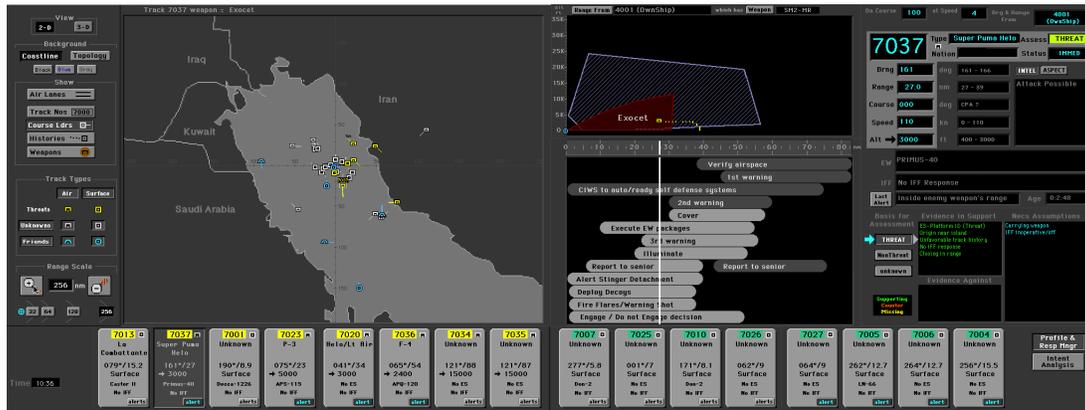

Figure 1. DSS. Track Summary (top right), Track Profile (left), Response Manager (middle right), Basis for Assessment (bottom right), Track Priority (bottom).

During the training scenarios, team members communicated via an open-channel communication system. The dataset from the study consists of transcripts of team discussions that were segmented for further analysis by turn of talk for a total of 12,027 turns.

3.2 Coding

I analyzed the transcripts using the discourse codes described in Table 1 below. These codes were developed, validated, and applied to the data as part of the prior work by Swiecki and colleagues (2020). As described there, the data were coded using regular expression–based classifiers. Turns of talk that matched one or more regular expressions associated with a code were labelled for that code. The coding scheme was validated using Cohen's kappa and Shaffer's rho (Marquart, Swiecki, Eagan, & Shaffer, 2018) as part of a process that compared the ratings of two human raters and the automated classifiers. The automated classifiers for the codes in Table 1 were validated using the standard threshold for kappa (> 0.65) and a rho threshold of < 0.05. All pairwise combinations of raters (humans and automated classifier) achieved kappa > 0.83 and rho $(0.65) < 0.05$.

3.3 RQ 1

Does the measure align with individuals qualitatively determined to be impacted differently by interdependence?

To address this question, I conducted a qualitative analysis of the individuals in the study using the coded transcript data. Next, I developed and applied a measure of the impact of interdependence based on ENA. Finally, I compared the results of the qualitative analysis, the impact measure, and ENA network diagrams to close the interpretive loop.

3.3.1 Qualitative Analysis

Through the qualitative analysis, I sought to identify how individuals in this context were impacted by interdependence. To do so, I examined the coded transcript data for evidence of how information and behaviour were being coordinated (discursive interdependence) and how contributions from individuals might prompt contributions from others (temporal interdependence).

3.3.2 The Impact Measure

Next, I derived a measure of the impact of interdependence by comparing the results of interdependent and independent ENA models on the same data. The main output of ENA is a network for each unit of analysis that represents the relative frequencies of the connections between discourse codes they formed. Given a set of coded discourse data, the ENA algorithm segments the data into *lines* — the smallest segments of discourse to be analyzed — and *conversations*, which define spans of lines that are potentially related to one another in a meaningful way. For example, conversations could be defined as all of the lines from a particular day or a particular team.

The aim of ENA is to measure the associations or *connections* between discourse moves. To do so, it measures the co-occurrence of codes within a given set of lines (within a given conversation) called the *stanza window*. Co-occurrences are counted when a code occurs in the last line of the window, defined as the *referring line*, and a different code occurs either within the referring line or within preceding lines of the window. Depending on the conversation segmentation and units of analysis chosen, lines within the window could be from a single unit of analysis or multiple units of analysis. The size of the stanza window can vary depending on the nature of the data and the phenomena of interest. ENA moves a stanza window over each line and counts whether a given co-occurrence is present in that line or between the line and the preceding lines in the window.

Table 1. Codes, definitions, and examples

Code	Definition	Example
Detection	Talk about radar detection of a track or the identification of a track, (e.g., vessel type).	IR/EW NEW BEARING, BEARING 078 APQ120 CORRELATES TRACK 7036 POSSIBLE F-4
Track behaviour	Talk about kinematic data about a track or a track's location.	AIR/IDS TRACK NUMBER 7021 DROP IN ALTITUDE TO 18 THOUSAND FEET
Assessment	Talk about whether a track is friendly or hostile or the threat level of a track, or indicate tracks of interest.	TRACKS OF INTEREST 7013 LEVEL 5 7037 LEVEL 5 7007 LEVEL 4 TRACK 7020 LEVEL 5 AND 7036 LEVEL 5
Status updates	Talk about procedural information, for example, track responses, or talk about tactical actions taken by the team.	TAO ID, STILL NO RESPONSE FROM TRACK 37, POSSIBLE PUMA HELO
Seeking information	Ask questions regarding track behaviour, identification, or status.	TAO CO, WE'VE UPGRADED THEM TO LEVEL 7 RIGHT?
Recommendation	Recommend or request tactical actions.	AIR/TIC RECOMMEND LEVEL THREE ON TRACK 7016 7022
Deterrent orders	Give orders meant to warn or deter tracks.	TIC AIR, CONDUCT LEVEL 2 WARNING ON 7037
Defensive orders	Give orders to prepare defences or engage hostile tracks.	TAO/CO COVER 7016 WITH BIRDS

The above process results in a matrix of co-occurrence counts for each line of the data, showing the connections that were formed by each line. Next, for each unit, the algorithm sums the connections in their lines and creates a vector for each unit in high-dimensional space that represents their co-occurrence counts for each pair of codes (excluding pairs of identical codes). These vectors are normalized to the unit sphere to highlight differences in their content rather than their magnitude.

This method of identifying co-occurrences means that independent and interdependent models of collaborative discourse can be created for the same data simply by varying the stanza window size. To create an independent model, the window is set to one line; to create an interdependent model, the window is set to more than one line. The output of both models is a normalized vector in the same high-dimensional space that for each unit represents the connections it made in their discourse. The models differ with respect to whether those connections only occurred within their own single lines of discourse or could occur within a window that included lines of discourse from themselves and others.

To create the independent and interdependent models for the data, I used the rENA package for R (Marquart, Swiecki, Collier, et al., 2018). I selected individual team members as the units of analysis. I defined lines as turns of talk, conversations as the lines from a given team and training scenario, and codes as those in Table 1. The independent model used a stanza window size of one line; the interdependent model used a stanza window size of five (one referring and four prior) lines. This window size was chosen based on the qualitative analysis of the data previously conducted by Swiecki and colleagues (2020).

The output of each model was a vector in the same high-dimensional space ($n = 28$ dimensions) that for each unit represented the connections it made. To derive a measure of the impact of interdependence on individuals in the data, I calculated the correlation[4] between the independent and interdependent vectors for each unit of analysis. These vectors represent the normalized co-occurrences for each unit of analysis in the same high-dimensional space. This correlation value represents how similar the independent and interdependent results are for a given unit of analysis. Strong positive correlations suggest that the models are similar, indicating a low impact of interdependence, while smaller correlations suggest that the models are different, indicating a high impact of interdependence.

3.3.3 Closing the Loop
To assess the convergence between the impact measure and the qualitative findings, I triangulated the qualitative results, the impact measure, and differences between network representations from the independent and interdependent models.

As described above, ENA produces normalized vectors for each unit of analysis that indicate the relative frequency of co-occurrences between codes in their data. To visualize these vectors, ENA produces weighted network graphs for each unit of analysis. The nodes of these graphs correspond to the codes, and the weighted edge between any two nodes corresponds

---

[4]For two non-zero vectors, the correlation between them is equal to their cosine similarity, or the cosine of the angle between them.

to the relative frequency of co-occurrence between those codes. Edges that are thicker and more saturated indicate higher frequency. Any two networks can be compared by subtracting their edge weights to produce a *difference graph* that shows which connections are stronger in one network than in the other.

Typically, these network visualizations are projected into the metric space created by performing spectral value decomposition on the matrix of co-occurrence vectors (where each row corresponds to a unit of analysis and each column is a co-occurrence between codes). Network nodes are placed in the space via an optimization routine that minimizes the distance between the centres of mass of the networks and their corresponding points in the space. In this study, I was primarily interested in comparing the networks of particular individuals produced from either the independent or the interdependent ENA model, and not the metric spaces produced by either model. Therefore, for this study, the choice of metric space, and thus the positions of the network nodes, are arbitrary. To facilitate comparisons between networks, I chose the space that produced networks that were easier to interpret — their nodes being reasonably distant from one another.

3.4 RQ 2

Does the measure suggest that interdependence is impactful in this dataset overall?

To address this research question, I tested whether interdependence was statistically impactful overall for the individuals in this dataset. To do so, I calculated 95% confidence intervals around the impact measure I. Due to the nested structure of the data (individuals being nested withing teams), I used a linear mixed model (LMM) that regressed the impact measure on an overall intercept, as well as a random intercept for each team according to the following formula, where U_{0j} is a random effect for teams:

$$I_{ij} = \gamma_{00} + U_{0j} + R_{ij} \tag{1}$$

This model tests whether the intercept — the overall mean — is a significant predictor of the expected value of the impact measure while controlling for the nesting of individuals within teams. The confidence intervals around the mean impact measure can be used to determine whether the interdependent and independent models differed significantly according to whether the interval contains a value of interest. Here, I chose a value[5] of 0.8. Because the impact measures are correlations, prior to running the LMM, they were transformed using Fisher's Z transformation to allow for meaningful averages and the calculation of confidence intervals.

3.5 RQ 3

Does the measure suggest that individuals in command roles are more impacted by interdependence than those in support roles and that those in the experimental condition are less impacted than those in the control condition?

To address this research question, I compared the impact measures of the individuals in different roles and different conditions using an LMM. This model regressed the impact measure on variables for role and condition, controlling for team effects according to the following formula:

$$I_{ij} = \gamma_{00} + \gamma_{10}\delta_{ij}^{Support} + \gamma_{20}\delta_{ij}^{Experimental} + U_{0j} + R_{ij} \tag{2}$$

This model tests whether the difference in the impact measure between those in command and those in support roles is statistically significant, controlling for condition and team effects. It also tests whether the difference in the impact measure between those in the experimental and those in the control condition is significant, controlling for role and team effects.

4. Results

4.1 RQ 1

Does the measure align with individuals qualitatively determined to be impacted differently by interdependence?

4.1.1 The Detect-to-Engage Sequence

The goal for each team in this context was to perform the detect-to-engage sequence. This task is described in more detail in Swiecki and colleagues (2020) and Johnston and colleagues (1998), but, in brief, it entails using computer and radar systems to detect and identify potential threats; assess their threat level; monitor their behaviour; and take appropriate actions toward them, such as warning them or engaging them in combat. In order to complete these tasks, teams have to coordinate information about threat detection, behaviour (e.g., speed, altitude, and bearing), assessment, and responses to warnings. Moreover, individuals have to coordinate behaviours such as sending warnings to threats and readying their own ship's defences.

---

[5]This value was chosen because it allows for a more stringent test of the similarity between models. One (a perfect correlation) was not chosen because the correlation coefficient cannot exceed one. Thus, it is relatively easy for correlation values to be significantly different than one.

Below, I provide examples from the qualitative analysis for a commander in the control condition, a commander in the experimental condition, and a supporting team member in the experimental condition to illustrate the impact of interdependence on individuals in this context. Note that in the excerpts below, speakers often begin their turn of talk by addressing the member(s) of the team for whom the message is intended, followed by who was speaking. Thus, "TAO, CO" should be read as "TAO, this is CO."

4.1.2 TAO: Control Condition

The excerpt in Table 2 illustrates typical interactions between a TAO and their teammates in the control condition. In particular, the excerpt shows how this TAO relies on information from other members of the team in order to maintain awareness of the situation and take appropriate actions.

Table 2. TAO control condition

Line	Speaker	Utterance
1	TAO	TAO, THAT'S AFFIRM, I NEED INFORMATION ON 7023.
2	IDS	TAO ID, WE HAVE NO MODES AND CODES ON 23.
3	TAO	TAO, AYE. EW?
4	EWS	NOTHING AT THIS TIME.
5	TAO	BEARS 257?
6	EWS	CORRECT.
7	EWS	NO EW 23
8	IDS	TAO ID, YOU WANT A WARNING LEVEL 1 ON 23?
9	TAO	ID TAO, STAND BY.
10	CO	TAO CO, LETS GO AHEAD AND SEND OUT A WARNING.
11	TAO	TAO, AYE. ID TAO, LEVEL 1 WARNING 7023.

The excerpt begins as the TAO requests information on a new track designated as 7023 (line 1). The IDS responds to this request, indicating that they have no information regarding whether the track is a civilian or a military vessel: "no modes and codes[6] on 23" (line 2).

In the next line, the TAO asks whether any electronic emissions information has been detected that could help to identify the vessel: "EW?" (line 3). The EWS replies that they do not have such information at the present time (line 4). The TAO follows up by seeking information about the bearing of the track (line 5); the EWS confirms the bearing (line 6) and restates that they have no electronic emissions information (line 7). Next, the IDS recommends sending a level 1 warning to the track (line 8). The TAO rejects this recommendation in the next line — "Stand by" (line 9) — but, soon after, the CO cuts in and states that the warning should be sent (line 10). In the final line, the TAO gives the order to the IDS (line 11).

The conversation illustrates that the TAO in this situation depends on information from their teammates. In particular, the TAO requests information about the track three times (lines 1, 3, and 5), and this information is provided by the IDS and EWS (lines 2, 4, 6, and 7). Moreover, instead of ordering the level 1 warning of their own volition, the TAO only sends the order in response to a recommendation from the IDS (line 8) and a direct order from the CO (line 10). Hence, the excerpt suggests that information had to be coordinated between different members of the team in order for the TAO to take appropriate action. In other words, this TAO was impacted by interdependence.

These conclusions are supported by this TAO's interdependent and independent network graphs and their impact measure. Figure 2 shows the relevant network graphs. The interdependent network is relatively dense (most nodes being connected), with its strongest (thickest and most saturated) connections between seeking information and track behaviour and seeking information and detection. The independent network is also dense, with all connections being a similar frequency. However, the difference graph shows that connections to seeking information are stronger in the interdependent model, while most of the remaining connections are stronger for the independent model. In other words, the two networks are quite different. The impact measure for this TAO is 0.16, suggesting that the interdependent and independent models are very different and the impact of interdependence is high for this individual.

4.1.3 TAO: Experimental Condition

The excerpt in Table 3 illustrates typical interactions between a TAO in the experimental condition and their teammates. In contrast to the TAO shown in the prior excerpt, this TAO relies less on information from their teammates, instead contributing their own information and taking action based on their own contributions.

In the first line, the EWS reports the detection of a new track at bearing 235, a high-speed patrol boat referred to as "Boghammer" (line 1). Next, the TIC provides a status update on track 7011, which had been previously detected, indicating

---

[6]Air defence warfare teams use identification, friend or foe codes (IFF) to understand track intent. IFF codes are broadcast by tracks and classified into five categories, or *modes*, that identify the track as civilian or military.

(a) Interdependent network   **(b)** Independent network   **(c)** Difference graph
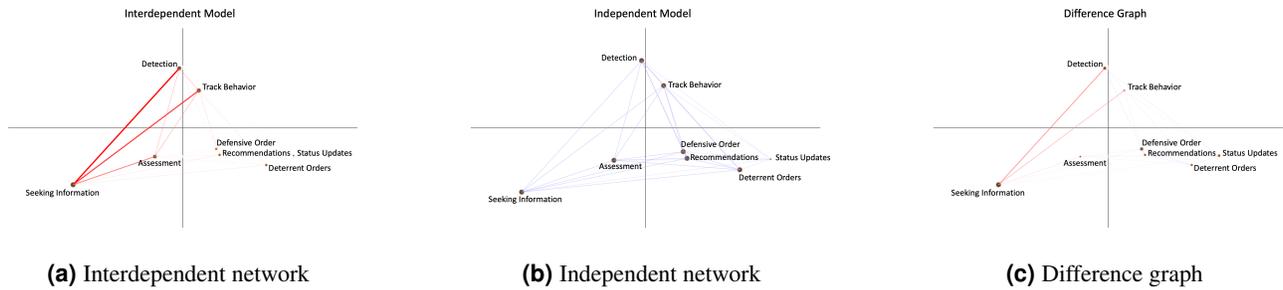
Figure 2. Network comparison for control condition TAO.

Table 3. TAO experimental condition

Line	Speaker	Utterance
1	EWS	TAO/EW NEW EMITTER BOGHAMMER BEARING 235.
2	TIC	TAO THIS IS TIC NEGATIVE RESPONSE LEVEL ONE QUERY TO TRACK 7011 SIR.
3	TAO	EW THIS IS TAO YOUR EMITTER CORRELATES TO TRACK 7016 BEARING 235 AT 5 MILES.
4	EWS	ROGER.
5	TAO	TIC/TAO LEVEL ONE TO 7016 BEARING 235 AT 5 MILES.

that the track did not respond to a level 1 warning (line 2). Continuing to focus on the newly detected track, the TAO restates its bearing and provides critical information about its distance from their own ship "Bearing 235 at 5 miles" (line 3). The EWS acknowledges this information (line 4), and, directly after, the TAO orders a level 1 warning to the track, restating the bearing and distance information (line 5).

Although these individuals were in the experimental condition, they still responded to information from their teammates and, thus, were impacted by interdependence. Here, the TAO responded to the track detection by the EWS in lines 3 and 5. However, the impact of interdependence on this TAO appears here to be less than the impact on the TAO in the control condition. While this TAO responds to the EWS, they repeat track behaviour information provided by the EWS (bearing) and add new track behaviour information (distance) (line 3). Moreover, they reference this same information when giving a deterrent order (line 5). When viewed in isolation, the TAO does not appear much different than when viewed in context; in both cases, the TAO makes connections between deterrent orders and track behaviour.

These conclusions are supported by the TAO's interdependent and independent networks and their impact measure. Figure 3 shows their relevant networks. The interdependent network is relatively dense, with its strongest connection between detection and track behaviour. The independent network is also dense, with all connections having similar strengths. The difference graph indicates that the connection between detection and track behaviour is stronger in the interdependent model. However, the networks are relatively similar overall — with most connections in the subtraction being thin and lightly saturated, suggesting that the two networks had similar edge weights. The impact measure for this TAO is 0.81, suggesting that the impact of interdependence on this individual was relatively small.

(a) Interdependent network   **(b)** Independent network   **(c)** Difference graph
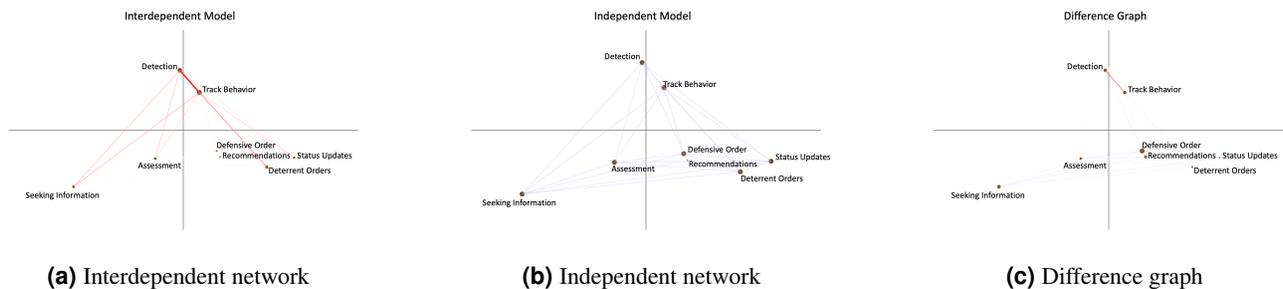
Figure 3. Network comparison for experimental condition TAO.

4.1.4 EWS: Experimental Condition

The excerpt in Table 4 illustrates typical interactions between the EWS, a supporting member of the team, and their teammates. The excerpt illustrates how this EWS maintained consistent behaviour associated with their role despite other information being passed before and after their contributions.

In the first line, the EWS reports that electronic emissions from track 7027, identified as an F-4 jet, have ceased: "Racket ceased" (line 1). The ADWC acknowledges this message (line 2), and, directly after, the IDS reports that track 7027 has descended 900 feet in altitude (line 3). The IDS follows up with information that track 7027 has disappeared from radar completely (line 5). In the final line, the EWS reports the detection of a new track, 7017, at bearing 007 emitting an APS 115 radar signal (line 7).

The excerpt suggests that this EWS maintained consistent behaviour and did not respond directly to information from their teammates[7]. In particular, the EWS first contributed information about track 7027 (line 1). The ADWC and IDS respond to the information, with the IDS adding more information about the track's behaviour (lines 3 and 5). In the final line, the EWS continues to report information about a new track (line 7). The excerpt shows that the EWS remained consistent in their behaviour and that they contributed the same kind of information as those speaking around them (track behaviour information). This suggests that the impact of interdependence on the EWS was relatively small.

Table 4. EWS experimental condition

Line	Speaker	Utterance
1	EWS	AIR/EW RACKET CEASED ON THE F-4 TRACK NUMBER 7027
2	ADWC	AIR AYE
3	IDS	AIR/ID TRACK 7027 COMP-2 THE F-4 HAS NOW DESCENDED TO 900 FEET
4	ADWC	AIR COPY
5	IDS	AIR/ID TRACK 7027 THE F-4 HAS DISAPPEARED OFF THE RADAR
6	ADWC	AIR AYE
7	EWS	AIR/EW NEW RACKET BEARING 007 APS 115 P-3 CORRELATION TRACK 7017

These conclusions are supported by the interdependent and independent network graphs for the EWS and their impact measure. Figure 4 shows the relevant networks for the EWS. The strongest connection in both the independent and interdependent network graphs is between detection and track behaviour. The difference graph shows no connections between any codes save for weak connections between seeking information and detection and seeking information and track behaviour. This indicates that the EWS's interdependent and independent networks have highly similar edge weights. The impact measure for the EWS is 0.95, indicating that their interdependent and independent models are highly correlated and the impact of interdependence on the individual is small.

(a) Interdependent network **(b)** Independent network **(c)** Difference graph
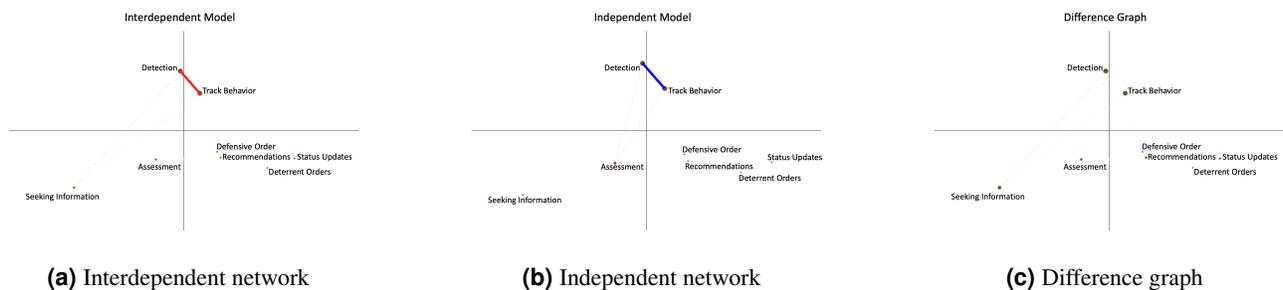
Figure 4. Network comparison for experimental condition EWS.

4.2 RQ 2

Does the measure suggest that interdependence is impactful in this dataset overall?

Table 5 shows the results of the LMM that tested the overall impact measure *I*. The hypothesis test of the intercept estimate tests whether the value is significantly different from zero. Here, however, I am interested in whether the intercept is significantly different from 0.8. The estimate and confidence intervals reported in the table can be converted back to the original correlation scale using the hyperbolic tangent function. Converted to the original scale, the lower bound, mean, and upper bound are [0.51, 0.58, 0.64]. This interval does not contain 0.8, showing that the average impact measure for these data is statistically significant. These results suggest that interdependence, as measured by *I*, has a significant impact on these individuals overall.

---

[7]Of course, my claim is not that this EWS never responded to their teammates, but that they were characterized by a consistent focus on track detection.

Table 5. Test of the overall impact measure I

(Intercept)	0.66***
	(0.05)
(CI)	(0.56–0.75)
AIC	93.06
BIC	101.08
Log Likelihood	−43.53
Num. obs.	107
Num. groups: Team	16
Var: Team (Intercept)	0.02
Var: Residual	0.11

$^{***}p < 0.001;\ ^{**}p < 0.01;\ ^{*}p < 0.05$

4.3 RQ 3

Does the measure suggest that individuals in command roles are more impacted by interdependence than those in support roles and that those in the experimental condition are less impacted than those in the control condition?

Table 6 shows the results of the LMM that tested for differences between individuals in different roles and conditions. Converted back to the original correlation scale, the results show that, on average, individuals in support roles had I values that were 0.11 greater than those in command roles, controlling for condition and team effects, though this value is not significant. The results also show that individuals in the experimental condition, on average, had I values that were 0.25 greater than those in the control condition, controlling for role and team effects. This difference is significant at $p < 0.05$. In other words, individuals in the experimental condition tended to have correlations between independent and interdependent models that were significantly higher. In turn, this result suggests that individuals in the experimental condition were impacted by interdependence significantly less than those in the control condition.

Table 6. Test of role and condition effects

(Intercept)	0.45***
	(0.08)
$\delta^{Support}$	0.11
	(0.08)
$\delta^{Experimental}$	0.26**
	(0.08)
AIC	94.37
BIC	106.92
Log Likelihood	−42.18
Num. obs.	91
Num. groups: Team	16
Var: Team (Intercept)	0.00
Var: Residual	0.13

$^{***}p < 0.001;\ ^{**}p < 0.01;\ ^{*}p < 0.05$

5. Discussion

In this paper, I described and tested a novel measure of the impact of interdependence on individuals in a CPS context. The measure was developed by correlating independent and interdependent model results for the same individuals using ENA. I provided evidence of the validity of this measure by applying it to collaborative discourse data collected during a study of air defence warfare teams in training. Specifically, I triangulated the measure with network model comparisons and a qualitative analysis. My results suggest that the measure can distinguish between individuals who have been differently impacted by interdependence. To demonstrate the utility of the measure, I used it to determine whether interdependence was impactful on individuals in these data overall, and whether the design of the collaborative task contributed to the impact of interdependence. These results suggest four conclusions.

First, the measure is sensitive to differences in the impact of interdependence. According to the qualitative analysis, some individuals were more impacted than others, and the measure aligned with those interpretations. Consequently, it may suffice

to use independent models to describe some individuals (like the EWS), while other individuals (like the TAO in the control condition) are better described by interdependent models.

Second, overall, individuals in this context were significantly impacted by interdependence, and hence an interdependent model is more appropriate for this context in general. This result aligns with the understanding of the air defence warfare task as highly interdependent.

Third, individuals in command roles were more impacted by interdependence than those in support roles, though not significantly. Because commanders rely on information from other teammates to make decisions, this is not surprising. However, the non-significant effect suggests that more data may be needed to support this claim.

And, finally, individuals on teams with access to a DSS were significantly less impacted by interdependence. This result supports the hypothesis that the presence of the DSS, which provided commanders with access to information typically reported by supporting members, served to help individuals in this condition act more independently.

The results of this work suggest that comparisons between independent and interdependent models can be used to derive criteria for determining when the impact of interdependence on individuals is meaningful. Such criteria can be used to evaluate the effects of task structures on collaboration and to decide whether interdependent models (which tend to be more complex than independent models) are necessary. The work described here extends prior investigations of interdependence by developing and testing a measure that is (a) sensitive to discursive and temporal interdependence and the microgenetic nature of collaboration because it is based on co-occurrence counts between discourse codes in a moving window; (b) data agnostic because, although text discourse was used in this study, the method only requires that the discourse (from any mode) be coded; and (c) statistically convenient because it relies on one statistical test regarding the networks rather than one for each network edge.

Of course, the work presented here has several limitations. First, the measure was tested on only one dataset. While this dataset was appropriate for studying interdependence, it is possible that results may vary on different data, particularly if those data were collected from collaborative tasks that were less interdependent. As such, this work constitutes a step forward in the investigation of interdependence, not an end result. Future work will include testing the impact measure on different real-world datasets, as well as simulated collaborative discourse data.

Second, the data examined here were from only one mode: speech. However, we know that collaboration may involve multiple modes of interaction. In fact, these data were collected from a situation in which teams interacted heavily with computer systems. Unfortunately, these interactions were not recorded. Future work will need to explore the impact of interdependence using multi-modal data to provide a more complete understanding of how interdependence impacts individuals.

Third, this study only used one particular method to develop a measure of the impact of interdependence, ENA. The intent of this work is not to argue that ENA is the only method that could be used, but that the impact of interdependence can be measured by comparing interdependent and independent models and that ENA is a particularly useful technique for doing so. While ENA meets the discussed criteria for developing an effective measure of the impact of interdependence, it is not alone in this respect — other methods could be used. In particular, DBNs are a promising alternative approach. One could measure the impact of interdependence via this technique by comparing networks that include relationships between a given individual's contributions and the contributions of others to simpler Bayesian networks that only consider the current state or the prior states of that individual.

Both ENA and DBNs are agnostic to the type of data analyzed. Techniques that rely on cohesion metrics, however, such as CNA and GCA, operate exclusively on spoken or textual discourse, limiting their utility. Nevertheless, these methods could be used to investigate the impact of interdependence using only speech or text data. For example, GCA could be used to compare internal cohesion results to responsivity results, which measure the cohesion of an individual's contributions with others. If responsivity is found to be significantly different than internal cohesion, interdependence may be impactful.

Such comparisons would be insightful; however, the relationship between cohesion and interdependence is different from the relationship suggested by the ENA-based metric. Cohesion-based metrics describe relationships in terms of how similar contributions are to one another. And, in some cases, the decision as to whether relationships are meaningful is based on the level of cohesion — for example, in CNA, links are only represented if a given level of cohesion (similarity) is present. If used to determine whether interdependence is impactful according to the method described in this paper, such measures would indicate that interdependence is more impactful the more similar an individual's contributions are to those of their teammates. In contrast, the ENA-based measure indicates that interdependence is more impactful the more different the contributions are. This is because ENA does not count connections between the same discourse codes within the moving window, only different codes. It follows, then, that interdependent ENA models are only different from independent ones if individuals tend to make contributions in response to other individuals contributing different kinds of discourse.

Future work will need to examine this difference between cohesion- and non-cohesion–based metrics more closely. The fundamental decision is whether being more similar or more different from your teammates is indicative of the impact of interdependence. This decision may hinge on the type of collaborative task. In highly interdependent tasks with clearly defined roles, it seems more likely that interdependence is expressed in terms of the differences between teammates — individuals depend on one another for information and resources that they do not possess. It may be less important for the team to cohere than for the right people to get the right information. In less interdependent tasks without explicit roles, such as general

collaborative discussions, interdependence may be better expressed in terms of the similarities between individuals — whether they are taking up aspects of one another's discourse.

6. Conclusion

Despite the above limitations, this work provides a proof of concept for a novel measure of the impact of interdependence on individuals. The evidence suggests that the measure can discriminate between those who have been more impacted by interdependence and those who have been less impacted. The measure builds on other techniques for investigating interdependence, such as the intra-class correlation, lag sequential analysis, and point process modelling. It is distinguished by its sensitivity to discursive and temporal interdependence and microgenetic interactions, its flexibility regarding the data used, and its statistical convenience. Furthermore, the results suggest that this measure can be used to address specific questions about collaborative settings. While the measure was developed using ENA, the general approach — statistically comparing the results of interdependent and independent models of the same individuals — can be conducted with other models that account for interdependence.

The work presented here makes several contributions to the study of CPS. First, it provides a technique for examining interdependence that, unlike point process models, is sensitive to more than just temporal interdependence. Thus, it is applicable to a broader range of collaborative settings. Second, the measure can be used to determine whether interdependent models are strictly necessary: if interdependent and independent models of the same individuals do not significantly differ, the independent model may suffice. Because interdependent models tend to be more complex than independent ones, the measure can help researchers make principled decisions regarding models that affect the explainability of their results to other audiences. These decisions may have import for existing and future assessments of CPS and collaboration analytics systems that present model results to educators and learners. Finally, the measure can be used to test hypotheses about collaborative situations, such as whether the task design affects how interdependence impacts individuals or whether individuals in particular roles are behaving in expected ways. Such tests may further our understanding of collaboration and lead to refinements in collaborative task design.

Declaration of Conflicting Interest

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

References

Aronson, E. (1978). *The Jigsaw Classroom*. Newbury Park, CA: Sage.

Bakeman, R., & Gottman, J. M. (1997). *Observing Interaction: An Introduction to Sequential Analysis*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511527685

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, *6*(3), 271–315. https://doi.org/10.1207/s15327809jls0603_1

Chinn, C. A., & Sherin, B. (2014). Microgenetic methods. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 171–190). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.012

Clark, H. H. (1996). *Using Language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620539

Csanadi, A., Eagan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). 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*, *13*(4), 419–438. https://doi.org/10.1007/s11412-018-9292-z

Dascalu, M., McNamara, D. S., Trausan-Matu, S., & Allen, L. K. (2018, April). Cohesion network analysis of CSCL participation. *Behavior Research Methods*, *50*(2), 604–619. https://doi.org/10.3758/s13428-017-0888-4

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, *95*(1), 32–53. https://doi.org/10.1037/a0017328

Dillenbourg, P. (1999). *Collaborative Learning: Cognitive and Computational Approaches. Advances in Learning and Instruction Series*. ERIC. https://doi.org/10.1016/S0360-1315(00)00011-7

Dowell, N. M. M., Nixon, T. M., & Graesser, A. C. (2019, June). Group communication analysis: A computational linguistics approach for detecting sociocognitive roles in multiparty interactions. *Behavior Research Methods*, *51*(3), 1007–1041. https://doi.org/10.3758/s13428-018-1102-z

Fiore, S. M., Graesser, A. C., Greiff, S., Griffin, P., Gong, B., Kyllonen, P. C., ... von Davier, A. (2017). Collaborative problem solving: Considerations for the national assessment of educational progress. Retrieved from https://orbilu.uni.lu/handle/10993/31897

Forsyth, D. R. (2009). *Group Dynamics*. Cengage Learning.

Gee, J. P. (1999). *An Introduction to Discourse Analysis: Theory and Method*. London, UK: Routledge. https://doi.org/10.4324/9780203005675

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine de Gruyter. https://doi.org/10.4324/9780203793206-1

Griffin, P., & Care, E. (2014). *Assessment and Teaching of 21st Century Skills: Methods and Approach*. Springer. https://doi.org/10.1007/978-94-017-9395-7

Gweon, G., Jain, M., McDonough, J., Raj, B., & Rosé, C. P. (2013). Measuring prevalence of other-oriented transactive contributions using an automated measure of speech style accommodation. *International Journal of Computer-Supported Collaborative Learning*, *8*(2), 245–265. https://doi.org/10.1007/s11412-013-9172-5

Halpin, P. F., & von Davier, A. A. (2017). Modeling collaboration using point processes. In A. von Davier, M. Zhu, & P. Kyllonen (Eds.), *Innovative Assessment of Collaboration* (pp. 233–247). Springer. https://doi.org/10.1007/978-3-319-33261-1_15

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA, USA: MIT Press. https://doi.org/10.1023/A:1008285430235

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2011). Multilevel analysis in CSCL research. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing Interactions in CSCL* (pp. 187–205). Springer. https://doi.org/10.1007/978-1-4419-7710-6_9

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. https://doi.org/10.1037/10278-002

Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17–64). Westport, CT: American Council on Education and Praeger.

Kapur, M. (2011, March). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer-Supported Collaborative Learning*, *6*(1), 39–56. https://doi.org/10.1007/s11412-011-9109-9

Kozlowski, S. W., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, *7*(3), 77–124. https://doi.org/10.1111/j.1529-1006.2006.00030.x

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259–284. https://doi.org/10.1080/01638539809545028

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *The Academy of Management Review*, *26*(3), 356. https://doi.org/10.2307/259182

Marquart, C. L., Swiecki, Z., Collier, W., Eagan, B., Woodward, R., & Shaffer, D. W. (2018). *rENA: Epistemic Network Analysis*. Retrieved from https://cran.r-project.org/web/packages/rENA/index.html

Marquart, C. L., Swiecki, Z., Eagan, B., & Shaffer, D. W. (2018). *ncodeR*. Retrieved from https://cran.r-project.org/web/packages/ncodeR/ncodeR.pdf

Martinez-Maldonado, R., Kay, J., Buckingham Shum, S., & Yacef, K. (2019). Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data. *Human–Computer Interaction*, *34*(1), 1–50. https://doi.org/10.1080/07370024.2017.1338956

Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the *Journal of Applied Psychology*. *Journal of Applied Psychology*, *102*(3), 452–467. https://doi.org/10.1037/apl0000128

Miyake, N., & Kirschner, P. A. (2014). The social and interactive dimensions of collaborative learning. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 418–438). New York: Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.026

OECD. (2017). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving, revised edition*. Paris: PISA, OECD Publishing. https://doi.org/10.1787/9789264281820-en

Olson, G. M., Herbsleb, J. D., & Rueter, H. H. (1994). Characterizing the sequential structure of interactive behaviors through statistical and grammatical techniques. *Human-Computer Interaction*, *9*(3-4), 427–472. Retrieved from https://dl.acm.org/doi/10.1207/s15327051hci0903%25264_6

Reimann, P. (2009). Time is precious: Variable- and event-centred 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

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, *2*(3), 235–276. https://doi.org/10.1207/s15327809jls0203_1

Ruis, A. R., Siebert-Evenstone, A. L., Pozen, R., Eagan, B. R., & Shaffer, D. W. (2019). Finding common ground: A method for measuring recent temporal context in analyses of complex, collaborative thinking. In K. Lund, G. P. Niccolai, C. Hmelo-Silver, G. Gwon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings: 13th International Conference on Computer Supported Collaborative Learning* (CSCL '19), 17–21 June 2019, Lyon, France (Vol. 1, pp. 136–143). Retrieved from https://repository.isls.org/handle/1/1559

Rummel, N., Walker, E., & Aleven, V. (2016, June). Different futures of adaptive collaborative learning support. *International Journal of Artificial Intelligence in Education*, *26*(2), 784–795. https://doi.org/10.1007/s40593-016-0102-3

Schrödinger, E. (1935). Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, *23*(48), 807–812. https://doi.org/10.1007/BF01491987

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. https://doi.org/10.18608/jla.2016.33.3

Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (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

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks, CA: Sage.

Stahl, G., Law, N., Cress, U., & Ludvigsen, S. (2014). Analyzing roles of individuals in small-group collaboration processes. *International Journal of Computer-Supported Collaborative Learning*, *9*(4), 365–370. https://doi.org/10.1007/s11412-014-9204-9

Stylianou-Georgiou, A., Pananastiou, E., & Puntambekar, S. (2011). Analyzing collaborative processes and learning from hypertext through hierarchical linear modeling. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing Interactions in CSCL* (pp. 145–159). Boston, MA: Springer. https://doi.org/10.1007/978-1-4419-7710-6_7

Suthers, D. D. (2017). Multilevel analysis of activity and actors in heterogeneous networked learning environments. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of Learning Analytics* (1st ed., pp. 189–197). Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17.016

Suthers, D. D., & Desiato, C. (2012). Exposing chat features through analysis of uptake between contributions. In *Proceedings of the Hawaii International Conference on System Sciences* (HICSS '12), 4–7 January 2012, Honolulu, HI, USA (pp. 3368–3377). IEEE. https://doi.org/10.1109/HICSS.2012.274

Swiecki, Z., Ruis, A. R., Farrell, C., & Shaffer, D. W. (2020). Assessing individual contributions to collaborative problem solving: A network analysis approach. *Computers in Human Behavior*, *104*. https://doi.org/10.1016/j.chb.2019.01.009

Van de Ven, A. H., Delbecq, A. L., & Koenig Jr., R. (1976). Determinants of coordination modes within organizations. *American Sociological Review*, *41*(2), 322–338. https://doi.org/10.2307/2094477

Wang, X., Wen, M., & Rosé, C. (2017). Contrasting explicit and implicit scaffolding for transactive exchange in team oriented project based learning. In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Making a Difference: Prioritizing Equity and Access in CSCL, 12th International Conference on Computer Supported Collaborative Learning* (CSCL '17), 18–22 June 2017, Philadelphia, PA, USA (Vol. 1, pp. 25–32). International Society of the Learning Sciences. Retrieved from https://repository.isls.org/handle/1/274

Webb, N. M. (2008). Teacher Practices and Small-Group Dynamics in Cooperative Learning Classrooms. In R. Gillies, A. Ashman, & J. Terwel (Eds.), *The teacher's role in implementing cooperative learning in the classroom* (pp. 201–221). Springer. https://doi.org/10.1007/978-0-387-70892-8_10

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann. https://doi.org/10.1016/C2009-0-19715-5