Collaborative and Individual Scientific Reasoning of Pre-Service Teachers: New Insights Through Epistemic Network Analysis (ENA)

Andras Csanadi, Ludwig Maximilian University of Munich, andras.csanadi@psy.lmu.de
Brendan Eagan, University of Wisconsin-Madison, beagan@wisc.edu
David Shaffer, University of Wisconsin-Madison, dws@education.wisc.edu
Ingo Kollar, University of Augsburg, ingo.kollar@phil.uni-augsburg.de
Frank Fischer, Ludwig Maximilian University of Munich, frank.fischer@psy.lmu.de

Abstract: When assessing scientific reasoning both (1) modeling connections in the discourse and (2) doing so at an appropriate grain size can be challenging for researchers. Our study suggests combining a novel theoretical (Fischer et al., 2014) and a novel methodological (Shaffer et al., 2006) framework to respond to these challenges by detecting epistemic networks of scientific reasoning processes in the context of collaborative vs individual problem solving of pre-service teachers. We investigated (1) whether the combination of these frameworks can be fruitfully applied to model scientific reasoning processes and (2) what unit of analysis researchers or instructors should choose to answer questions of interest. One no ve l aspect of our study is that we compared epistemic networks in case of collaborative vs individual reasoning processes. Our results show that (1) epistemic networks of scientific reasoning can reliably capture reasoning processes when comparing collaborative vs individual reasoning; and (2) propositional and potentially larger units might be considere d as "optimal" units of analysis to detect such differences.

Keywords: collaborative problem solving, epistemic network analysis, scientific reasoning

Introduction

Assessment of scientific reasoning in process data is a critical for the development of appropriate learning support. Although many fruitful approaches have been developed for the evaluation of reasoning and argumentation (Brown, Furtak, Timms, Nagashima & Wilson, 2010); general theoretical and methodological frameworks that allow analysis of scientific reasoning patterns on multiple layers (e.g., Chi, 1997) are scarce. Consequently, the selection of grain size at an early stage of the analysis and a resulting dilemma surrounding creation of larger units that allow further interpretation of the data (e.g., Weinberger & Fischer, 2006) often limit the generalizability of findings (Chi, 1997; Stegmann & Fischer, 2011). Also, using a pre-defined selection of a unit of analysis might cause difficulties when a researcher or a tutor would like to be more conclusive about the reasoning processes: simultaneously making qualitative and quantitative assessments. For example, a researcher (or tutor) may be interested in ideas, or codes, at a very fine grained (e.g., propositional) level in order to detect "elementary" units of reasoning processes. Meanwhile, she might be also interested in the connections, or relationships, between these ideas or codes captured at that fine-grained level, in order to asse ss the quality of reasoning processes (Chi, 1997; Weinberger & Fischer, 2006). Moreover, when aggre gating data into larger chunks, what would be an optimal choice? Would combining multiple propositions or defining a larger, e.g. sentence units, lead to better representation of reasoning processes? The present study investigates whether a combination of a novel theoretical framework on scientific reasoning (Fischer et al., 2014) as we ll as a novel methodological approach on modelling reasoners' epistemic networks (Shaffer, 2006) can be meaningfully combined 1) to analyze patterns (epistemic networks) of scientific reasoning and 2) to disambiguate the question on grain size selection and data aggregation when assessing patterns (epistemic networks) of scientific reasoning.

Scientific reasoning and argumentation

There are different theoretical frameworks to conceptualize and analyze scientific reasoning. Many follow a "structural" approach, focusing on the structure of argumentation (see Brown et al., 2010) while others emphasize the role of engagement in scientific reasoning processes (Okada & Simon, 1997). Our work be lo ngs to the latter stream of research understanding scientific reasoning as engagement of individuals or groups in a sequence of epistemic activities (Fischer et al., 2013). According to this model, scientific reasoning involves reasoners identifying an existing problem (Problem identification), articulating questions of how to proceed with their reasoning processes (Questioning), derive possible explanations of the problem (Hypothesis generation), construct artifacts, such as intervention plans, to solve the problem (Generating solutions), generate and collect information (Evidence generation), evaluate that information (Evidence evaluation), engage others in the reasoning process (Communicating & scrutinizing), and draw conclusions (Drawing conclusions). Earlier studies found that both individual and collaborative reasoning in a professional problem solving context can be reliably coded using this framework (Csanadi, Kollar & Fischer, 2016).

Collaborative vs. individual scientific reasoning processes

Collaborative scientific reasoning has the potential to lead individuals to higher engagement in epistemic processes such as hypothesis generation and evidence evaluation compared to reasoning alone (Okada & Simon, 1997; Teasley, 1995). Similarly, more recent findings (Csanadi et al., 2016) showed that when pre-service teachers solved a problem from their future practice as dyads, they engaged more in hypothesis generation (i.e., trying to find an explanation to the problem) but less in generating solutions than individuals did. Nevertheless, this purely frequency-based approach for analysis to count the occurrence of certain codes has clear constraints. Most importantly, it cannot be conclusive enough regarding the patterns of epistemic processes that can characterize collaborative vs individual reasoning. For example, although dyads were found to be more explanatory, indicated by a higher engagement in hypothesizing, whether they did this in a more evidence-based manner (i.e. if they made more connections between hypothesizing and evaluating evidence) remained unclear. Being able to identify such connections or patterns in the data is, therefore, important for assessing quality aspects of scientific reasoning.

Selection of grain size and data aggregation to capture patterns of reasoning

To assess and compare reasoners with respect to the patterns of the epistemic activities they engage in, researchers should find answers to two related questions. First, what is an appropriate grain size (i.e., unit of analysis) and second, how should coded data be aggregated in order to gain a deeper understanding of the quality and features of the reasoning processes. Many researchers emphasize that data segmentation should be a separate and preceding step to coding (Chi, 1997; Strijbos, Martens, Prins & Jochems, 2006). This would mean that the division of verbal data into chunks that carry meaningful information for further analysis should precede further analyses. However, this early selection of the unit of analysis has its limitations (e.g., Chi, 1997). Especially the use of smaller grain sizes (e.g., propositional unit) allow for a more fine-grained analysis of reasoning processes (e.g., to interpret the relation between independent clauses of compound sentences) and allow for frequency-based analyses. Indeed, many quantitative approaches to the analysis of scientific reasoning processes (e.g., Okada & Simon, 1997) suggest analyzing frequencies of single categories. However, considering that discourse moves are not unrelated to each other, relying on solely frequency-based information of data can lead to missing meaningful patterns of discourse (Cress & Hesse, 2013). At this point an emerging concern of data aggregation (Stegmann & Fischer, 2011), i.e., how the researcher/tutor can make higher level inferences based on data coded at a lower grain size, often generates uncertainty. When looking for relationships between coded units (e.g., propositions), how far these units can fall from each other? Can we meaningfully detect relationships between two neighboring units or does allowing for slightly "longer distance" connections increase explanatory power? A method that allows more adaptable choice of grain size (Siebert-Evenstone et al., 2016), such as considering multiple units of analysis instead of relying on a pre-defined selection in order to model scientific reasoning could help to answer such questions.

Another issue associated with coding-independent segmentation may arise if some codes turn out to be highly frequent ones while others occur relatively rarely. "Uneven" frequency distributions can bias further analyses of the dataset (e.g., Csanadi, Daxenberger, Ghanem, Kollar, Fischer & Gurevych, 2016). For example, high frequency codes might generate many connections with each other while also being related to many other codes. On the other hand, low frequency codes may lack enough connections with other codes to demonstrate the power to discriminate between epistemic networks of different groups (e.g., dyads vs individuals). Thus, in case of modeling reasoning processes, this can mean that some reasoning patterns may emerge as mere artifacts while other connections in the data may remain undetected, and therefore, models of scientific reasoning should account for such limitations.

To summarize, using a hierarchical segmentation procedure and reliance on solely frequency-related information when analyzing scientific reasoning processes and comparing reasoners, leaves open the questions of (1) how to aggregate and identify meaningful larger patterns in the data that can (2) help more validly capture the reasoning performance beyond simply counting the occurrences of single codes.

Epistemic Network Analysis: A method to analyze (multiple scopes of) scientific reasoning

One solution of the abovementioned problems can be to code on multiple levels of granularity (Stegmann & Fischer, 2011). As Chi (1997) notes, this approach has the advantage of leading to more reliable results and interpretations at different levels. Generally speaking, segmentation might be a matter of the researchers' focus of interest (Chi, 1997), the theoretical framework they apply (Clara & Mauri, 2010), the nature of data (e.g. synchronous vs asynchronous discussions) and more. Still, selecting multiple levels of analysis can contribute to more valid interpretations about the data (Chi, 1997; Weinberger & Fischer, 2006) as different lenses may capture different aspects of collaborative learning and reasoning processes.

Epistemic Network Analysis (ENA; Shaffer, 2006) is a method to identify meaningful and quantifiable patterns in discourse/reasoning. It can provide an alternative to the widespread "code and count" approach. ENA moves beyond the traditional frequency-based assessments by examining the structure of the co-occurrence, or connections in coded data. Moreover, compared to other methodological approaches, e.g., sequential analysis (see in Cress & Hesse, 2013), ENA has the novelty of (1) modeling whole networks of connections and (2) it affords both quantitative and qualitative comparisons between different network models.

A main theoretical assumption of ENA is that repeated co-occurrences of two or more codes in the discourse can reveal epistemic networks which characterize an underlying Discourse (Gee, 1999; Collier et al., 2016), e.g., to collaborative (vs. individual) scientific reasoning. To identify a unit of analysis for calculating such co-occurrences, ENA provides an adaptable feature: the *moving stanza window size* (MSWS; Siebert-Evenstone et al., 2016). The term stanza window refers a window or scope within which ENA is searching for connections. This means that a MSWS=1 allows search for connections only between a proposition of reference and its preceding proposition. Therefore, a MSWS=1 results in connections only between neighboring propositions. A MSWS=2, however, allows one further step: it allows connection between a proposition of reference and the two preceding propositions. By changing MSWS from smaller values to larger it is possible to open the "search window" from very narrow context to wider ones. As a result, the researcher or tutor can look for connections not only within propositions (as in case of "coding and counting" approaches) or between neighboring propositions, but even between propositions that are two, three or more steps further from each other in the discourse. In short, it offers the advantage of multiple scopes for analysis. Here we aim to investigate if ENA can reveal some characteristics of collaborative (compared to individual) scientific reasoning processes as well as to articulate what grain sizes should be considered when using ENA for that analysis.

Furthermore, ENA provides the opportunity to quantitatively and qualitatively compare different epistemic network models with each other. Quantitative comparison is possible by using calculated centroids for every epistemic networks generated by ENA. Such centroid values are determined by the strength of connections between nodes in the epistemic network. Nodes are the codes (such as epistemic activities, see below) while the strength of connections between them are generated based on their local co-occurrences (within each stanza window: see above). These centroid values can be used for quantitative analyses. Furthermore, qualitative comparison of epistemic networks is possible using various options for visualization. One option is "Subtracting networks" which means contrasting two network models by subtracting their nodes and connections weights from each other. A resulting "subtracted network" represents the difference between two reasoning networks and therefore, can illustrate what makes dyadic reasoning different from individual reasoning.

Research questions

RQ1: Do collaborative and individual reasoners exhibit different epistemic networks of scientific reasoning while solving a professional problem?

While earlier studies demonstrated differences between collaborative and individual reasoning in terms of their engagement in different epistemic activities (Csanadi et al., 2016; Okada & Simon, 1997), these results were mainly frequency-based. E.g., the researchers compared proportions as well as raw frequencies of engagement in different epistemic activities, such as evaluating evidence or hypothesizing. Thus, an open question is whether dyads also differ from individuals in the patterns of epistemic activities they engage in during scientific reasoning. In this study we address this question using ENA (Shaffer et. al. 2009) to capture meaningful patterns of co-occurrences between epistemic activities (i.e., epistemic networks of scientific reasoning), and to compare dyads with individual reasoners.

Epistemic networks can, however, also be defined based on larger speech units (e.g., across multiple propositions) and we can also implement larger grain sizes beyond analyzing neighboring propositions or within sentences. To fully answer RQ1, therefore, we investigated whether some grain sizes can provide potentially better explanation of patterns in the data than others.

RQ2: Do the epistemic networks we detect investigating RQ1 differ from epistemic networks based on the same data set that has been randomly resorted (i.e. with the same frequency information)?

ENA models co-occurrences of codes, since some codes occur more frequently than others, it is more likely that these highly frequent codes make connections (co-occur) with other codes more often than lower frequency codes. Consequently, ENA may "overestimate" some connections. Therefore, to answer our second research question, we compared ENA results from RQ1 to ENA results obtained from a dataset that contained only frequency information of the original discourse (see below). If the epistemic networks identified in relation to RQ1 cannot be explained merely by the frequency distribution of epistemic activities, the epistemic networks detected in relation to RQ1 should differ from the epistemic networks of the randomly resorted dataset.

Method

The data analyzed in this study is a re-analysis of process data from another study (Csanadi et al., 2016). In the original study N=76 preservice teachers (59 female, M_{Age}=21.22, SD_{Age}=3.98) solved a problem case from their future profession in one of two between-subject conditions: either as individuals (N=16) or as dyads (N=30 dyads). Think aloud and discourse data of their problem solving were first manually segmented into propositional units and then coded for further analysis. The coding scheme of that study was developed based on the framework of scientific reasoning by Fischer et al. (2014). Epistemic activities identified by the framework (see above) were applied (Table 1): Problem identification for an initial attempt to build an understanding of the problem; Questioning for statements or questions triggering further inquiry; Hypothesis generation for developing explanations of the problem; Evidence generation for reference to information or lack of information that could support a claim; Evidence Evaluation to evaluate a claim; Communicating and scrutinizing for planned discussions with others (e.g., in order to find out further information); Drawing conclusions for concluding outcomes of reasoning. Finally, the epistemic activity of "Constructing artefacts" (in Fischer et al., 2014) was operationalized as developing interventions or solution plans, and such propositions were labelled as Generating solutions. Moreover, the codes for Evidence generation and Evidence evaluation were merged into Evidence evaluation. Both segmentation (79.73% of agreement by Coder 1 and 85.09% of agreement by Coder 2) and coding (κ = 0.68) proved to be reliable. We used this dataset (original dataset) to analyze further in our present study.

We used the abovementioned original dataset to answer RQ1. To be able to answer RQ2 we created a randomized dataset in the following way. Using the original dataset within each dyad and individual participants we created a random sequence of the pre-segmented propositions (Csanadi et al., 2016). That meant, the original sequence of propositions were randomized while the relative frequency of propositions was preserved (no propositions were deleted). This new randomized dataset preserved the information of the occurrence of epistemic activities, yet, in a randomized order; containing the information to which individual or dyad the epistemic activities belong to, how frequently they occur, but without any information regarding their sequence in the original dataset.

We used ENA to identify epistemic networks of scientific reasoning in order to answer both RQ1 and RQ2. We built epistemic network models using ENA in four steps. First, we calculated co-occurrences between epistemic activities (MSWS=1, means rotation was applied) for dyads and for individuals. At the same time ENA automatically generated a centroid value for each dyad or individual that served as a numeric representation of their epistemic network and it was included in further analysis to compare dyadic and individual epistemic networks of scientific reasoning. Second, mean, or "average," networks were defined for both the dyadic and the individual reasoning conditions, respectively. Each of these networks visually represented all the connections that participants (dyads or individuals) generated in the given condition. Third, we quantitatively compared epistemic networks for dyads with epistemic networks for individuals by comparing the mean centroid values (calculated in step 1) in the two conditions. Fourth, we subtracted the mean dyadic and mean individual networks from each other (by using the "Subtracting networks" option in ENA). The resulting subtracted networks visualized what connections contributed to the difference between the two reasoning conditions (dyadic vs individual, calculated in step 3).

To be able to fully answer RQ1 regarding grain size, we sequentially set MSWS from 1 to 7, step-by step, performing the same analysis for each stanza window size. The resulting epistemic network models at each MSWS level allowed us quantitative as well as qualitative (visual) comparisons.

To answer RQ2, we used the randomized dataset selecting the same parameters and performing the same analysis as in case of RQ1. We compared the outcomes of this analysis with the ENA results from RQ1.

Table 1: Coding scheme for epistemic activities

Code	Short Description	Example

Problem identification	An attempt to understand the problem.	"So it is about a student, // who has low grades"
Questioning	A question orienting inquiry.	"Ok, so what is the reason for that?"
Hypothesis generation	Explanation of the problem.	"...the reason is her learning method"
Evidence generation	Referring to any information / lack of inf. relevant for the inquiry	"She studies diligently at home"
Evidence evaluation	Evaluation information.	"...you can even exclude the problem of exam nerves"
Generating solutions	Planning an intervention / solution to the problem.	"You should discourage her from using surface strategies"
Communicating & scrutinizing	Planning to engage others.	"You can also talk to the parents"
Drawing conclusions	Concluding the outcomes of the earlier steps of inquiry.	"For me these would be the most important points..."
Non-epistemic	Everything else, e.g. coordination.	"Ok, have you read it through?"

Results

RQ 1: To answer RQ1, as a first step, we compared dyadic and individual networks at the grain size of MSWS=1 which lead to the following results. The mean centroid value for individuals' epistemic networks (*M*=.21,*SD*=.32) was significantly different from the mean centroid value for dyads' epistemic networks (*M*=-.11,*SD*=.21), $t(44)=3.65$, $p<.01$, $d=1.32$. Plotting epistemic networks (Figure 1) further revealed that the central epistemic activity accounting for most of the connections was evidence evaluation. Moreover, in case of dyads evidence evaluation showed more complex network than in case of individuals: for dyads it was connected to hypothesis generation, communicating and scrutinizing, generating solutions and non-epistemic propositions; while in the case of individuals it was only connected to hypothesis generation and generating solutions. Finally, subtracting individual from dyadic networks revealed that in case of individual networks it was solution generation rather than evidence evaluation that played a central role in contrast to dyadic networks where only evidence evaluation showed multiple connections after subtraction.

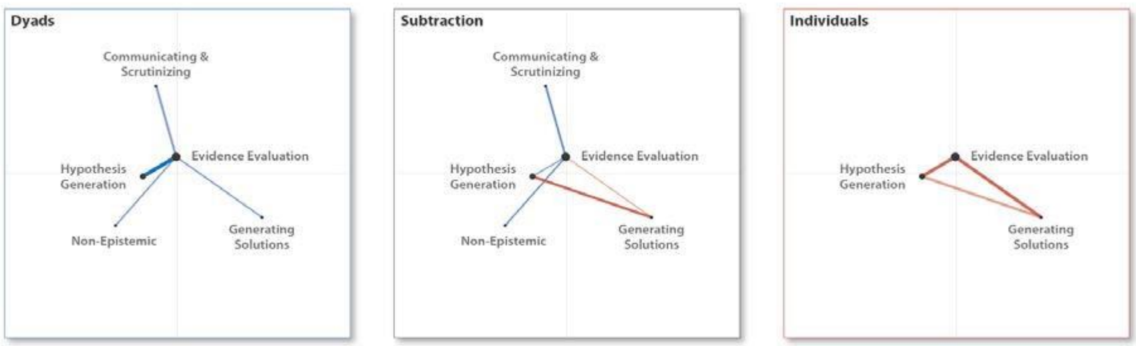

Figure1. Epistemic networks of dyads (blue, left), individuals (red, right) and the difference between their networks (center) using the original dataset.

To completely answer RQ1 and in order to see whether there is an optimal grain size that can best capture the differences between epistemic networks of dyads and individuals, we compared epistemic networks at $1 \leq$ MSWS ≤ 7 levels which led to the following results. All comparisons were statistically significant at least under $p<.01$. Although effect size showed a small increase at every MSWS level, these differences were small: the explained variance increased only by 5.35% ($\Delta R^2=.05$) from MSWS=1 ($R^2=.30$) to MSWS=7 ($R^2=.36$).

Finally, a visual inspection of the epistemic networks conducted at 1 ≤ MSWS ≤ 7 levels suggested highly similar patterns at every MSWS levels (see Figure 1).

RQ 2: Similar to the outcomes of RQ1, when using the randomized dataset, the mean centroid value for individuals' epistemic networks (*M*=.17,*SD*=.26) was significantly different from the mean centroid value for dyads' epistemic networks (*M*=-.09,*SD*=.20), $t(44)=3.35$, $p<.01$, 95%, $d=1.15$. Plotting epistemic networks (Figure 2), however, revealed no visible difference between dyadic and individual networks. Dyadic and individual networks showed identical patterns regarding complexity: connections occurred among the three most frequent epistemic activities: hypothesis generation, solution generation and evidence evaluation. This was in clear contrast with the results of RQ1 where epistemic networks were different for collaborative vs individual reasoning (Figure 1). A further important difference is that Figure 2 does not indicate any central epistemic activity, neither for dyadic and individual nor for the subtracted pattern. Moreover, Figure 2 shows very low level of network complexity for dyads (connections among the highest-frequency activities) compared to Figure 1. Finally, the subtracted network model on Figure 2 consists of only blue lines, indicating that dyads made more connections among the highly frequent codes than individuals.

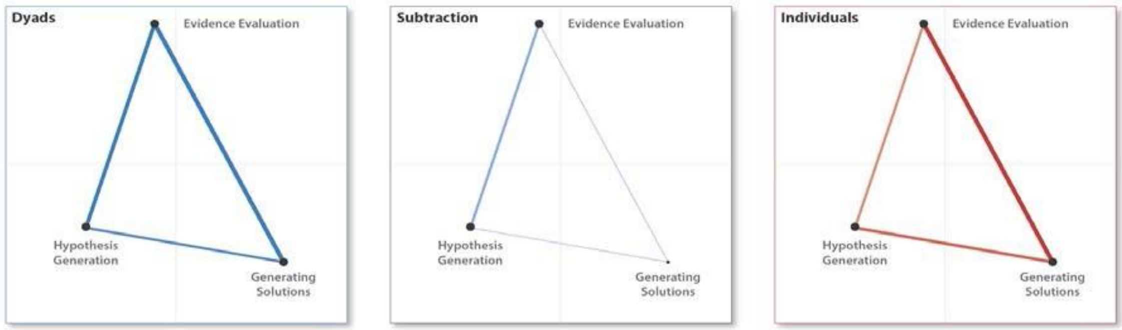

Figure2. Epistemic networks of dyads (blue, left), individuals (red, right) and the difference between their networks (center) using the randomized dataset.

Discussion

The two main aims of our study were (1) to see whether we can aggregate data to capture meaningful patterns (epistemic networks) of scientific reasoning processes regarding collaborative and individual reasoning (RQ1 & RQ2) and (2) to search for an optimal grain size, or unit of analysis, for such aggregation (RQ1). We sought to answer these questions by the application of a novel theoretical framework on scientific reasoning (Fischer et al., 2014) and a novel methodological approach on modelling epistemic networks (Shaffer, 2006).

The outcomes for RQ1 suggest that epistemic networks of scientific reasoning can meaningfully differentiate between collaborative and individual reasoning processes. More specifically, dyads seemed to engage in a more complex manner in scientific reasoning compared to individuals: they made more connections between epistemic activities (specifically, with evidence evaluation). Moreover, while individual reasoning was rather solution-focused; dyadic reasoning seemed to be more evidence-focused. These results are also in accordance with previous frequency-based findings (Csanadi et al., 2016; Okada & Simon, 1997).

To be able to fully answer RQ1 we ran further analyses at different stanza window sizes that resulted in patterns quite similar to those in Figure 1. On the one hand, this suggests the robustness of our findings, on the other, a question of the optimal grain size to detect meaningful patterns of scientific reasoning cannot be conclusively answered. A partial answer is, however, that choosing larger speech unit (e.g., sentences) at a first step may represent reasoning patterns in the data at least closely as well as propositions do. Yet, further empirical research could test (1) whether this is true and if (2) varying stanza window sizes on sentence units would lead to different results. Based on the results of this study and considering the exhaustiveness of hand-coding procedure, however, choosing larger units of analysis that still carry the information needed to model scientific reasoning may be an efficient choice for the researcher/tutor.

The outcomes on RQ2 show that epistemic networks extracted on discourse data (original dataset) are likely to be valid models for the evaluation of reasoning patterns in the data as they are not reducible to the frequency distribution of codes. Furthermore, it is clear that merely frequency-information in the data resulted in only "poor" network models: networks represented solely the most frequent codes and their connections. Additionally, after subtracting those networks the results suggested that dyads made more connections everywhere. These results did not add much explanatory value to the frequency-based outcomes of the earlier findings (Authors, 2016a), which underlines the assumption that ENA conducted on real discourse data can detect meaningful patterns of scientific reasoning.

Finally, the results imply that identifying epistemic processes on the propositional level and aggregating data by conducting epistemic network analysis can offer a powerful way to meaningfully assess scientific reasoning in discourse.

Final conclusions

Our results have further important consequences.

First, the theoretical (Fischer et al., 2014) and the methodological (Shaffer, 2009) frameworks could be fruitfully combined to result in a series of robust analyses of identifying epistemic networks of scientific reasoning.

Second, dyadic vs. individual reasoning networks can be valid models of scientific reasoning in discourse. Yet, we need more empirical research to see if this result holds as well as see the predictive validity of our findings. For example, the extent to which dyads' more extensive connections could potentially predict learning outcomes and whether some connections might play a stronger moderating role in that process, are questions for future research.

Finally, additional analyses that can more directly address the impact of frequency distribution of codes on epistemic networks could also contribute to conclusions regarding the validity of the findings. For example, alternative measures provided by ENA could account for "imbalanced" frequency distribution in the data. Those measures could apply, for example, some weighting method for assigning less weight to higher frequency codes or to connections among higher frequency codes, in order to reduce the chance of detecting artefactual connections due to higher probability of co-occurrence between high-frequency codes. Similarly, if ENA could generate a simple frequency-based epistemic network model (similar to the outcomes on RQ2) and would allow its subtraction from the epistemic network model on the real dataset; that would afford the visualization of reasoning patterns beyond highly frequent connections. Yet, such measures should be implemented with caution: connections captured in the discourse should always represent connections in the Discourse (Gee, 1999).

References

Brown, N. J., Furtak, E. M., Timms, M., Nagashima, S. O., & Wilson, M. (2010). The evidence-based reasoning framework: Assessing scientific reasoning. *Educational Assessment, 15*(3-4), 123-141.

Chi, M. T. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences, 6*(3), 271-315.

Clarà, M., & Mauri, T. (2010). Toward a dialectic relation between the results in CSCL: Three critical methodological aspects of content analysis schemes. *International Journal of Computer-Supported Collaborative Learning, 5*(1), 117-136.

Cress, U., & Hesse, F. W. (2013). Quantitative methods for studying small groups. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. O'Donnell (Eds.). *The international handbook of collaborative learning,* 85-111. New York: Routledge, Taylor & Francis.

Csanadi, A., Kollar, I., Fischer, F. (2016). Scientific reasoning and problem solving in a practical domain: Are two heads better than one? In C. K. Looi, J. L. Polman, U. Cress, & P. Reimann (Eds.), *Transforming learning, empowering learners: The International Conference of the Learning Sciences (ICLS) 2016*, Vol. 1 (pp. 50-57). Singapore: International Society of the Learning Sciences.

Csanadi, A., Daxenberger, J., Ghanem, C., Kollar, I., Fischer, F., & Gurevych, I. (2016a). *Automated text classification to capture scientific reasoning and argumentation processes in different professional problem solving contexts.* Paper presented at the 26th Annual Meeting of the Society for Text & Discourse, Kassel, Germany.

Collier, W., Ruis, A. R., & Shaffer, D. W. (2016). Local versus global connection making in discourse. In C. K. Looi, J. L. Polman, U. Cress, & P. Reimann (Eds.), *Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS)* Vol. 1 (pp. 426-433). Singapore: International Society of the Learning Sciences.

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R.,. . . Eberle, J. (2014). Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research, 5,* 28-45.

Gee, J. P. (1999). *An introduction to discourse analysis.* New York: Routledge.

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *The Journal of the Learning Sciences, 14*(2), 201-241.

Shaffer, D.W. (2006). Epistemic frames for epistemic games. *Computers and Education 46*(3): 223-234.
Shaffer, D. W., Hatfield, D. L., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E. A., … Frank, K. (2009). Epistemic Network Analysis: A prototype for 21st century assessment of learning. *International Journal of Learning and Media, 1*(1), 1–21.
Siebert-Evenstone, A. L., Arastoopour, G., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D.W. (2016). *In search of conversational grain size: Modeling semantic structure using moving stanza windows.* Paper presented at the 12th International Conference of the Learning Sciences, Singapore.
Stegmann, K., & Fischer, F. (2011). Quantifying qualities in collaborative knowledge construction: the analysis of online discussions. In S. Putambekar, G. Erkens, C. Hmelo-Silver (Eds.). *Analyzing Interactions in CSCL: Methods, approaches and issues* (pp. 247-268). Springer US.
Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. (2006). Content analysis: What are they talking about? *Computers & Education, 46*(1), 29-48.
Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & education, 46*(1), 71-95.

Acknowledgments

This research was supported by the Elite Network of Bavaria under Grant K-GS-2012-209. This work was funded in part by the National Science Foundation (DRL-0918409, DRL-0946372, DRL-1247262, DRL-1418288, DUE-0919347, DUE-1225885, EEC-1232656, EEC-1340402, REC-0347000), the MacArthur Foundation, the Spencer Foundation, the Wisconsin Alumni Research Foundation, and the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison. The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.