7-2021

Connected design rationale: a model for measuring design learning using epistemic network analysis

Golnaz Arastoopour Irgens

Instructional Science (2021) 49:561–587
https://doi.org/10.1007/s11251-021-09551-8

ORIGINAL RESEARCH

Connected design rationale: a model for measuring design learning using epistemic network analysis

Golnaz Arastoopour Irgens[1]

Received: 1 March 2020 / Accepted: 20 June 2021 / Published online: 16 July 2021

Abstract
Virtual learning environments have the potential to support students' development of design skills in engineering education. However, few approaches exist for modeling and measuring design learning as it emerges in authentic practices, which often includes collaboration. This study merges learning sciences research with engineering design education to develop an approach for modeling and measuring design thinking. I propose a *connected design rationale* model which identifies relationships among design moves and rationale. Results from a qualitative examination of how professional engineers make connections among moves and rationales were used as the foundation to examine students in virtual internships. Using digital collaborative chat data and Epistemic Network Analysis (ENA), the discourse networks of students who had high and low scores in the virtual internship were compared to the discourse patterns of professional engineers to determine if measuring connected design rationale reveals meaningful differences between expert and novice design thinking. The results show a significant difference between high and low-performing students in terms of their patterns of connections and that high-performing students in the virtual internship made connections that were more like experts than low-performing students. Results suggest that a connected design rationale model distinguishes between experts and novices in meaningful ways and can be a robust approach for research in learning sciences and engineering education.

Keywords Engineering education · Virtual internships · Discourse analysis · Qualitative analysis · Expert-novice

Introduction

Due to the techno-industrial changes of this century, engineers in training need to develop the skills to optimize solutions and design products more than ever before. The field of engineering education has embraced this challenge of educating the new-century engineer and has made design a more central component of engineering education, offering design experiences for student engineers early in their academic careers (Atman et al., 2014; Dym

✉ Golnaz Arastoopour Irgens
garasto@clemson.edu

[1] Clemson University, 207 Old Main, Clemson, SC 29631, USA

et al., 2005). One option for effective design learning is through a virtual learning environment in which student novices interact with authentic tools and experts to learn the ways of the practice. Such virtual learning spaces have the potential to help students reflect on their work and develop complex design thinking (Shaffer, 2006; Svarovsky, 2011). However, measuring and evaluating design learning as it emerges during authentic practices, such as collaboration in design teams, is a significant issue and there is lack of consensus on an effective approach (Adams et al., 2011; Bartholomew, 2017).

This study tackles the problem of modeling and measuring design skills by proposing an approach for measuring design thinking in virtual learning environments. This approach frames design work as fundamentally requiring two key skills: (1) making appropriate *design moves*—actions taken during the design process (Schön, 1983) and (2) providing explicit *design rationale*—justifications for chosen design moves (Rittel, 1988). However, learning a practice, such as engineering design, centers on understanding the connections among domain-relevant elements rather than isolated skills and knowledge (DiSessa, 1993; Linn et al., 2013; Shaffer, 2006). Through the integration of learning sciences research with engineering design education, I propose that one critical piece of measuring design thinking centers on the ways in which students understand the connected relationships among design moves and design rationale. This study describes one approach to modeling and measuring the connections that learners make when engaging in collaborative design work in a virtual learning environment.

Theoretical framework

Engineering design practice

Design is the central and defining activity of engineering (Simon, 1996). Schön (1984, 1987, 1988) described the design process as a series of making *design moves*—actions taken during the design process to help the designer reach a final solution. These moves include conducting research on the potential components of a design, modifying a design drawing, and selecting a component for a product. At times, designers imagine or execute moves individually, but often times these moves occur in collaboration with others. Similar to other professions, collaboration in engineering design work involves sharing and developing ideas through creative exploration. However, collaborative design work uniquely involves posing questions and proposing answers to multi-dimensional problems with conflicting components (Lloyd, 2019). When designers execute moves, either collaboratively or individually, they generate and "see" new representations of the design, which transforms their understanding of the design scenario (Gero & Kannengiesser, 2008).

McCall and Burge (2016) claim, however, that designers may not always spontaneously "see" the consequences of a decision or understand why unanticipated consequences exist. This is especially true of novice or student designers. To support such "seeing" and the iterative movement of reflection and action through the design process, designers document explicit justifications for their design moves (Rittel, 1988). Whether through spoken or written language, articulating one's reasoning may reduce overlooking critical aspects of a problem, promote understanding connections to other similar problem scenarios, and facilitate communication among a design team or stakeholders. This notion of design reasoning as articulating an argument is now widely known as *design rationale*. A number of researchers have suggested design rationale provides the fundamental logic of design moves and is the basis for decision-making and actions in engineering and have explored options for documentation (Lee, 1997; Lee & Lai, 1991; Rockwell et al., 2010).

In engineering education, digital engineering notebooks are used as a way for students to document their rationale and as a data collection tool for instructors and researchers (Buchal, 2018; Febrian et al., 2015). For example, Moyne et al. (2018a, b) developed the Design Evaluation and Feedback Tool (DEFT) to assist students in documenting their design choices and reasoning. This tool also assisted instructors and researchers in collecting data and evaluating learning. Their findings showed that students relied on their notebooks to organize their design processes and used them mainly for documentation. Furthermore, Bergsman (2018) argues notebooks with pre-written scaffolds help students in understanding what and how to document their design work. These supports can also guide students to understand that documentation is a complementary practice to making design moves.

Thus, at its core, engineering design practice centers on two practices: (1) the ability to make appropriate design moves, meaning knowing how and when to take appropriate actions during the design process, and (2) the ability to use design rationale, meaning knowing how and when to provide explicit justifications for design moves.

How engineers learn to design

Schön (1987) argues professional practicums offer a space for professionals-in-training to reflect on their work by engaging in authentic tasks. In undergraduate education, one common example of practicums is internships in which students engage in design work at engineering companies. Such work-based learning programs give students an opportunity to work alongside senior practitioners who mentor them through their projects. Students can experience realistic aspects of design work such as working in teams, communicating with clients, and iterating through potential solutions. A real-world internship offers students an opportunity to apply their scientific and technical skills, and learn conflict resolution skills, how to manage job stress, and the consequences of missing deadlines in the workplace (Tener et al., 2001). Disadvantages of real-world internships in engineering education are that the priorities of an internship may be more company-focused than student learning-focused, that students often need to be advanced in their studies rather than being able to participate as first-semester freshmen, and that students typically commit all of their time to the internship and cannot take other classes.

One alternative approach to the real-world internship experience is for researchers and instructors to design virtual internship experiences for students. Such learning environments, when designed to be simulations of internships, prioritize student learning of core design competencies, offer opportunities for students to participate in workplace experiences early in their undergraduate program, and can be time-constrained to fit within a class. The design of simulated and authentic learning environments can be approached in different ways. Hod and Sagy (2019) claim that there are two approaches to school-based authentic learning designs: simulations in which students use developmentally appropriate versions of the tools and practices of the authentic culture and hybrid designs in which students interact with a simulation and with practitioners where the purpose is to advance both the students and the practitioners learning or interests. In the design of simulations, they suggest considering three different cultures at play: (1) the *current culture* contains the cultural practices within the authentically-designed classroom, (2) the *authentic culture* of the profession that the instructor or designer wants the students to enculturate; and (3) the *intended culture* that is the instructor or designer's vision or imagined world of that authentic culture. Findings from their review revealed that the combination of current classroom culture and the intended culture resulted in varying activities and interactions for students.

From a research standpoint, virtual simulations are advantageous because every student action in the simulation can be logged and organized in a central database. This form of data collection provides an opportunity for rich analysis of student design learning at a larger scale than with traditional studies.

Modeling design thinking

Given its complexity as it emerges in real-world and virtual practice, measuring design thinking is a challenge (Adams et al., 2011; Bartholomew, 2017). Learning sciences research suggests that complex thinking is characterized not by isolated pieces of knowledge and skill but by the organization of and relationships among domain-relevant knowledge and skills (DiSessa, 1993; Linn et al., 2013). Shaffer (2006) has characterized complex thinking in terms of a connected *epistemic frame:* ways of knowing, doing, and being that are linked together that are unique to a particular professional culture such as engineering (Arastoopour et al., 2016), urban planning (Bagley & Shaffer, 2009), and journalism (Hatfield, 2015). Identifying with a professional culture means acquiring a particular Discourse (with a capital D)—the ways people talk, read, write, understand, believe, act, and interact that are socially meaningful (Gee, 2011). An epistemic frame, then, is a formal configuration of the Discourse exhibited by those acculturated into a practice. Developing an epistemic frame as a learner means understanding and enacting the relationships among Discourse elements that are characteristic of the community, and thus developing expertise related to ways of knowing, being, and doing within a practice.

In this view, learning the practice of engineering design is not a stepwise procedure of accumulating skills, nor is it merely making moves and providing rationale. Rather, learning how to do design involves recognizing which design moves are linked to which rationale and the complex web of relationships among moves and rationale in design problems. I identify this web of relationships as a *connected design rationale* (Arastoopour Irgens, 2017).

Skilled designers exhibit a connected design rationale when they reflect on a problem and implement and justify the appropriate moves to develop a solution. For instance, if an undergraduate engineering student is designing handlebars for a bicycle, his or her design process can be represented as in Fig. 1 in which some moves and rationale are connected and some are not. In this scenario, the student makes two initial moves: gathering information and documenting the design process. The student justifies documenting the process and gathering information because it will help to better understand the problem. In this example, these two moves may have occurred independently of one another, meaning that the student did not relate these moves to one another. However, the justification of better understanding the problem is connected to each of these moves because the student has provided this rationale for both of these moves and thus, has identified a relationship among this rationale and these two moves. Once these two moves are enacted and the student has a better understanding of the problem, two new potential moves arise: taking a vote among team members or conducting experimental tests to choose a material. Documenting the design process and better understanding the problem are both linked to conducting experimental tests because the student has realized that understanding the problem

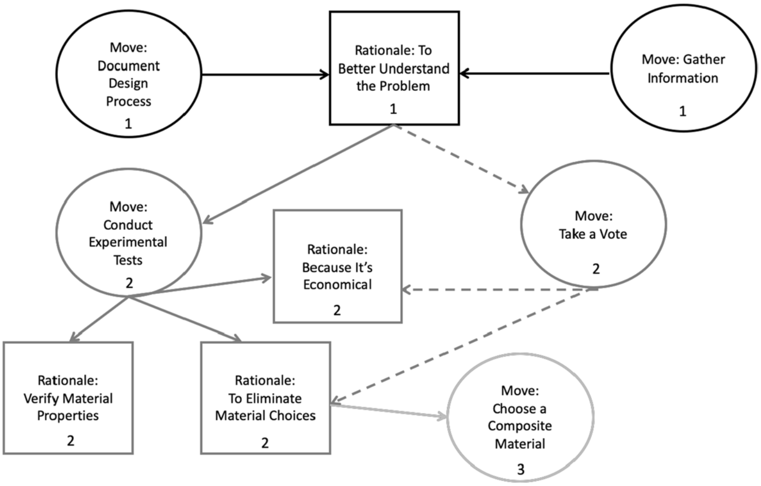


Fig. 1 Example of a student's connected design rationale in a bicycle handlebar design project at one instant in time. Circles represent design moves and rectangles represent rationales. Solid lines represented enacted moves and dotted lines represent imagined or suggested moves

will be further accomplished by conducting experimental tests and that documenting the process should continue when the team conducts such tests.

Ultimately, the team decides to conduct experimental tests and subsequently continue to document the process because these moves have three linked justifications: that experimental testing using a simulation will help the team narrow down their material choices, that it is more economical than other approaches, and that the simulation will help verify critical properties of each material. The other option of taking a vote to determine which material to use was not enacted because there was not a strong enough rationale to enact such a move. The next move would then be to choose the composite material and once again gather more information about the problem. This example reveals a short part of one skilled engineering student's design process as a series of interconnected enacted and imagined moves and rationales as a student team designs a product.

Measuring design learning

Prior work has demonstrated that connections among discourse elements can be modeled in terms of a network using *Epistemic Network Analysis* (ENA), a tool for visualizing and measuring complex thinking and learning (Shaffer, 2017; Shaffer & Ruis, 2017). In the context of design thinking and learning, ENA can measure connected design rationale by modeling a person's network of moves and rationale that have been articulated as discourse either through written documents, conversations, or actions. ENA accomplishes this by measuring the co-occurrences of discourse elements and representing them in weighted network models. Furthermore, ENA enables researchers to compare networks both visually and through summary statistics that reflect the weighted structure of connections. Thus, researchers can use ENA to not only model discourse networks, but also quantitatively compare the discourse networks of various individuals and groups of people.

Because of these affordances, ENA has been used to compare the epistemic frames of mentors and learners (Nash & Shaffer, 2011), the epistemic frames of students in classrooms and practicums (Hatfield, 2015), and in dozens of other learning scenarios (see Eagan et al., 2019). More relatedly, ENA has been used to model engineering design thinking by measuring the quality of discourse among students during the design process (Arastoopour et al., 2014, 2016).

Overview of Study and Research Questions

This study is grounded in a connected design rationale framework for modeling a learner's connected understanding among design moves and rationale in design practice. The analysis investigates how engineering professionals make connections among moves and rationale in their interactions with interns in a real-world company internship program and identifies several common patterns of connections. These patterns of connections from the real-world internship were then identified in engineering student digital data from a virtual internship program, *Nephrotex*. ENA was used to measure connected design rationale by building discourse networks of students' group conversations and a rubric was developed to measure students' individual notebook entries, which served as an outcome measure. Based on the individual outcome measure, networks were divided into high and low performing groups and were compared to the discourse patterns of experts from the real-world internship to determine the validity of using ENA to measure and model connected design rationale.

The research questions in this study are [RQ 1] How do expert engineers make connections among moves and rationale in a professional real-world internship setting? [RQ 2] Are there differences in how low-performing and high-performing students make connections among moves and rationale in a virtual internship? [RQ 3] How do low-performing and high-performing students' connections among moves and rationale in a virtual internship compare to expert engineers' in a real-world setting?

Real-world internship study

Methods

In a prior study, ethnographic data from two expert engineers and four interns was collected over the course of three months from an internship program at an engineering company, GammaCorp, that primarily designs, produces, and distributes high-pressure hydraulics (Arastoopour & Shaffer, 2015). In this prior work, I used the method of *epistemography*, an ethnographic analysis of a professional practicum through the lens of epistemic frames (Bagley & Shaffer, 2010; Hatfield & Shaffer, 2010). This approach focuses on collecting and analyzing data about the participant structures of reflection during the learning process and the epistemological underpinnings of the professional culture that support such structures.

Using the results from this prior epistemography work, as well as interviews, field notes, and recorded observational data, a grounded theory approach was used (Glaser & Strauss, 1967; Strauss & Corbin, 1994, 1998) to develop a set of qualitative categories, or codes, representing specific design moves and rationale (Table 1). These codes served as a basis for the coding scheme in the virtual internship in order to connect to an authentic professional learning environment. In a grounded approach, two core macro-categories were

Table 1 Connected Design Rationale coding scheme categories for real-world internship interviews, field notes, and recorded conversations

Code	Description	Examples from real world internship	Examples from virtual internship
J.Customer & Consultant Requests	Justifying design choices/devices or strategies by stating that they meet or should meet customer/consultant requests	*But I guess if they,* ***if the customer requests*** *like a P392 with a this and this and this, then you want to look at the P392 and then use that as a base*	*Hey, I was thinking if we should base it off of 5 of our consultants,* ***because if I want to test one nanotube for my consultant***
J.Performance Parameters & Requirements	Justifying design choices/devices or strategies by referring to general performance parameters or specific results either from documentation/papers or results from their own testing. The reference to the documentation or performance results does NOT to be explicit	*So, it's basically like having all the different engineers in our group… and seeing* ***if there is anything that can be tweaked to improve the design or maybe make it more cost effective***	*I feel like we should really look into the manufacturing process of Phase Inversion* ***because it seems to keep flux high and it is in the middle when it comes to cost***
J.Communication	Justifying design choices/devices or strategies by referring to facilitating communication efforts among colleagues or among engineers and customers	*Yeah we try to collect as much history so that we have answers* ***so that when we get into the project we have a good understanding of what we're getting into***	*Lets all put our stuff in the shared space* ***so we all can see***
M.Experimental Testing	Setting up an experiment by using a control device or have constants and changing one variable at a time. Or using experimental tools or techniques to understand technical features of a product	*Yeah****, I did [have to do a stress analysis] to a certain extent****… because these lifting I's, they are rated for a vertical 1000 lb and everything was getting for a 45 degree angle*	***I thought we were changing one variable at a time.*** *1 is the control. 2 is a different nanotube percent. 3 and 4 are different surfactants, and 5 is a different process*
M.Making Design Choices	Choosing a specification or characteristic for a prototype or design product	***Or we could use propane***	***We should go with hydrophilic for the third prototype***
M.Asking Questions	Asking questions or referring to the move of asking questions	*But the biggest thing is not being afraid to* ***ask the questions***	***do you guys think cost matters?***

J indicates Justification; M indicates Move identified: moves and justifications. Justifications was used in place of rationale because during the grounded analysis, I discovered that the predominant rationale used was to provide justifications for design moves. Particular forms of moves and justifications were identified through an open coding procedure, and an initial set of codes were developed. After several analytic iterations of removing, adding, and combining codes until I felt saturation was reached, a coding scheme consisting of six codes categorizing moves and justifications that emerged from the discourse was finalized. Each utterance in the interview and observational data was coded for these six codes. Then, I conducted axial coding (Strauss & Corbin, 1998) in which I analyzed connections among the codes and generated a descriptive story using the coded data that focused on connections among moves and rationale.

Results: a qualitative analysis of expert engineers

This section addresses RQ 1: *How do expert engineers make connections among moves and rationale in a professional real-world internship setting?* The examination of the real-world engineering internship focused on two participants: Warren (white, male) and Zara (white, female), who were engineers with 15 and 10-years' experience at the company, respectively. Warren and Zara were chosen because they interacted with and mentored the student interns most often compared to other employees who were observed. The analysis illustrates the patterns of connections expert engineers made among moves and rationale during conversations with interns, and during interviews with a researcher by providing examples from Warren's and Zara's discourse. The most common pattern in the engineers' discourse were specific to the domain of design, such as making design decisions in order to meet performance parameters and making design decisions based on the customer's preferences.

Analysis of Warren's discourse

Throughout the internship program, interns received tasks from the coordinator and lead engineer, Warren, who gave them preliminary information to start the task. In one instance, Warren, guided Brian, an intern, through the design task of a customized tow cart that would house a variety of tools.

When Brian met with Warren to receive guidance on the tow cart design, Warren explained that Brian had to design the cart using a CAD drawing tool. Warren suggested that instead of trying to design the final product, Brian should first "have the really basic design done before getting into too many specifics." After the meeting, Brian returned to his desk to work on designing the tow cart and began by reviewing the sketches that he and Warren had made. After experimenting with the CAD tool for several days, Brian discovered which pieces he could mount together and how they would fit collectively.

Once Brian had a preliminary design, Warren and other engineers met with him to provide more detailed feedback on the design. After reviewing Brian's design, Warren identified some issues. Warren explained that "they [the customer] want this design pump… with all these full controls and… they want storage for these hoses. And they want to be able to lift and drive it around the shipyard." However, in Brian's design, the orientation of the pump resulted in "all these hoses are sticking out in different ways," which blocked the customer's access to the controls on the cart. Warren made connections across moves and rationale when he specified that "the pump had to be oriented a certain way (*move of* *making design decisions)* to have the customer operate everything on this tow cart standing from one point (*rationale based on the customer requests*).” The design did not meet some of the performance requirements or customer requests and the engineers asked that Brian make several changes and present the design to them once the changes had been made.

When the researcher interviewed Warren about giving feedback to the interns on their design projects, he explained, “We [need to] ask, what is the application? What is the customer looking for? (*rationale of customer requests*) And then look to make sure the design meets those requirements (*rationale of performance parameters*)… Be it building tooling or something for the production line (*move of making design decisions*).”

Warren continued to explain that successful interns at GammaCorp asked many questions. He said, “It [asking questions] (*the move of asking questions*) kind of makes them [the interns] step back and rethink that they need to explain or reiterate: here’s what I understand you’re looking for (*rationale of effective communication*).” It was important to ask questions and have effective communication so that interns did not, as Warren put it, “spend all that time and effort on something that’s not needed or wanted.”

The excerpts above illustrate the context within which Warren made connections and highlighted some of the most common connections that were made. For example, Warren often connected the move of asking questions with the rationale of communication and justified design moves based on customer requests and the performance of the product.

Analysis of Zara’s discourse

In this next example, Zara, the expert engineer, worked on a quote for a hydraulic cylinder with Alice, the intern. The interaction between Alice and Zara was initiated when Alice approached Zara’s desk and asked for help.

When working together, Zara showed Alice a previously completed quote as a model and explained the rationale behind using specific quote forms: “Because this is so different from anything that we usually have, it’s not standard… I’m going to show you an example of one that Warren [the other expert engineer] did.” After opening Warren’s quote to use as an example, Zara explained to Alice the similarities between the current quote and Warren’s quote.

Alice nodded through the explanation, and then confirmed her next steps by asking another question: “Okay, so I fill this out, send it to custom products and then, I dig up all these prints for the cylinder?” Zara nodded her head, and Alice asked, “You said the plunger and the base and all the mounting were…?”.

Zara answered Alice’s question (*the move of asking questions*) and explained that Alice would have to use a custom form for those two parts: “the plunger and base are all custom… otherwise they are pretty much a standard cylinder,” (*move of making design choices*) and then provided a rationale based on the performance and design of the product: “because of the way they mount to the steel structure itself. Otherwise, they are pretty much a standard cylinder” (*rationale of performance parameters*).

When the researcher interviewed Zara about her interactions with Alice, Zara discussed the importance of the move of asking questions: “The thing is being confident in your abilities and being comfortable with asking questions (*move of asking questions*) because engineering is not necessarily about knowing the answers… [it’s about] being able to figure out or verify that they’re going to provide what the customer is actually looking for” (*rational of customer requests*).

These excerpts above illustrate the context in which Zara made a variety of connections among moves and rationale. In the examples above, she made connections between the move of asking questions and the rationale of communication, and also justified design moves based on customer requests and the performance of the product.

Virtual internship study

Methods

Setting

Nephrotex is an 8-week long engineering virtual internship program in which students role-played as interns at a fictional biomedical engineering design company (Chesler et al., 2013, 2015). Students worked on teams to design dialyzers for hemodialysis machines. Research investigations, design activities, and team interactions all took place through the web platform that supported the internship (Fig. 2). Students began by logging into the company website, which included a fictional email and chat interface. Acting as interns, they sent and received emails to and from their supervisor (a pre-scripted character) and used the chat window for instant messaging with other team members and their design advisor (an instructor who live chatted with students).

Interns received emails from their supervisor in the simulation who asked them to complete tasks. Some of the tasks were prescriptive in nature. For example, in the first few days of the internship interns were asked to read provided documents on the physical principles of hemodialysis and membrane diffusion and summarize their research in their notebooks. Students were required to write a notebook entry at the completion of every activity. Supervisors specified criteria for each notebook entry via email. For the notebook analyzed in this study, the criteria were that students (1) listed five prototypes, (2) provided a justification for each prototype as to why it was chosen, and (3) identified which stakeholders

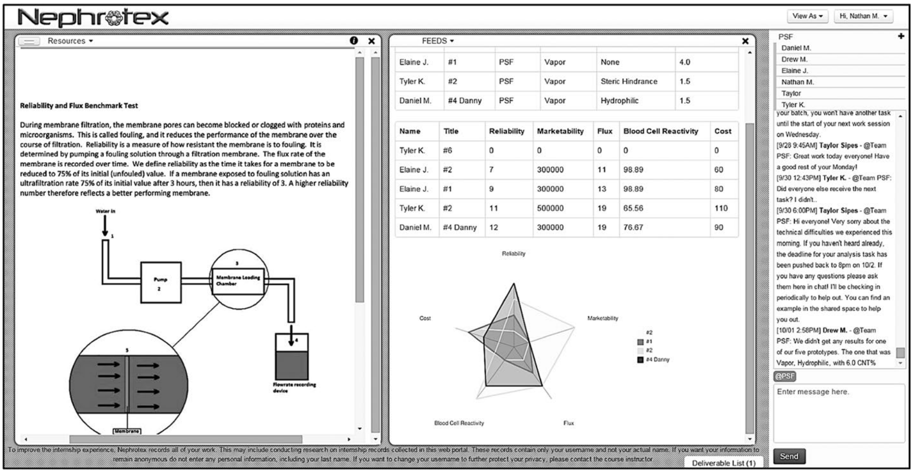


Fig. 2 Nephrotex simulated company portal interface, including a research report, data analysis tool, and chat window

would accept the prototype. This was a standardized grading procedure that all design advisors followed. After students submitted their notebook entry, design advisors assessed the notebook entry and provided feedback. If a student met the criteria, they were allowed to continue to the next activity. They received an email from the boss acknowledging their accomplishments and identifying the next task. If a student did not meet the criteria, they were asked to revise the notebook entry and were not allowed to move on to the next task.

As interns gained knowledge about hemodialysis membranes, they engaged with more open-ended, collaborative tasks centered on choosing specifications from a list of predefined selections for their design membrane prototype. During this process, they developed hypotheses based on their research, tested these hypotheses in the provided design space, and analyzed the results. Interns were also asked to meet multiple internal consultants' requests as closely as possible. The challenging aspect of this design problem was that the consultants' concerns were often in conflict with one another (e.g., as biocompatibility increases, cost also increases), and so the interns addressed and justified tradeoffs associated with their proposed design solution. Based on their research, they selected prototypes by combining different selections and parameter levels. Then, interns submitted their prototypes to a virtual testing lab and received performance results for each device. During the final days of the internship, students presented their final device design and justified their design decisions to the class and instructor.

Participants and data collection

Participants were first-year undergraduate engineering students ranging from ages 18–20. Students were enrolled in an introductory engineering course in which they participated in *Nephrotex*. Data were collected in two forms: (1) chat logs from teams of students from one activity in the program and (2) each student's engineering notebook entries from the end of that same activity. This activity was chosen because it was the main design activity in which students collaborated on a design prototype in their groups. The chat logs were analyzed for connected design rationale and the notebooks provided evidence of their individual design performance, which was the basis for separating students into two groups—low and high-performing.

The data were collected from nine instances of *Nephrotex*. All nine instances ranged from seven to eight weeks long and contained five teams of three to five students each, for a total of 65 teams and 314 students overall. All participants were first-year biomedical engineering majors and were selected to be in this study because their instructors indicated an interest in implementing *Nephrotex* into their course curriculum. The self-identified gender and racial demographics of the students were 70% male and 30% female; 73% White, 12% Asian, 7% Hispanic, 3% Other or Prefer not to respond, 1% Black, and 1% American Indian. These numbers are fairly consistent with the demographics of engineering bachelor degrees awarded in the U.S.: 79% male, 21% female; 62% White, 14% Asian, 11% Hispanic, 4% unknown, 4% Black, and 4% other (Yoder, 2017).

Segmentation and coding

Chat logs from the virtual internship were segmented by utterance—every time someone sent a response in a chat conversation. The coding scheme that was developed from the real-world internship was adapted for use in the internship analysis. Although the activities and participant structures in the real-world internship were not identical to the virtual internship, the adaptation was done to align as best as possible with an existing, real-world professional culture. For example, in the real-world setting, interns often worked individually on projects, but in the virtual internships, students often worked collaboratively. However, the coding scheme was developed to identify common engineering design moves and rationale at the intersection of the two research settings. The goal of aligning a real-world culture to an educational simulation for design or research purposes is not to duplicate what exists in the professional culture, but rather to create developmentally appropriate tasks that can fit within the constraints of an educational setting (Hod & Sagy, 2019).

Because a high volume of data was obtained from students' chat logs (26,867 utterances), an automated coding algorithm was used to code the chats (Cai et al., 2019; Eagan et al., 2017) and the statistic rho was used to determine whether the inter-rater reliability statistic of Cohen's kappa conducted on the sample could be generalized to the remainder of the dataset (Shaffer, 2017). The inter-rater reliability results for the virtual internship chat logs show that all pairwise agreements among rater one, rater two, and the automated algorithm had rho values of less than .05, which means the kappa statistic from the coded sample can be generalized to the entire dataset (Table 2). Cohen's kappa values ranged from .71 to 1.0.

Epistemic network analysis

ENA (Shaffer, 2017; Shaffer & Ruis, 2017) was used to measure the development of connections that engineers and interns made between design moves and justifications as defined by the coding scheme. ENA measures the connections between discourse elements, or codes, by quantifying the co-occurrence of those elements within a defined *stanza*. Stanzas are collections of utterances such that the utterances within a stanza are assumed to be closely related topically. Once the size of a stanza is identified, for any two codes, their strength of association is computed based on the frequency of their co-occurrence within each stanza in the discourse. Because the virtual internship discourse data were in the form of chat conversations, a *moving stanza window* model was used (Siebert-Evenstone et al., 2017). In this approach, co-occurrences are identified within one person's utterance and between that person's utterance and others in the conversation within a specified window segment. This window slides along the chat data and accumulates co-occurrences of codes for each person within their own utterance and between their utterance and other utterances that occurred before their own within the given window segment.

Table 2 Cohen's Kappa among rater 1, rater 2, and the computer for virtual internships chat codes

Code	Rater 1 & computer	Rater 2 & computer	Rater 1 & rater 2
J.Customer & Consultant Requests	.91**	.91**	1.0**
J.Performance Parameters & Requirements	.80**	.75*	.71*
J.Communication	.82**	.79**	.80**
M.Experimental Testing	.86*	.73*	.85**
M.Making Design Choices	.84**	.73*	.88**
M.Asking Questions	.89**	.87**	.98**

J indicates design justification code; M indicates design move code. * indicates rho < .05; ** indicates rho < .01.

In this study, ENA was used to sum each student's sliding windows and visualize each student's discourse as a weighted node-link network representation. To analyze several networks at one time, this study also used an alternative ENA representation in which the centroid (center of mass) of each network is calculated and plotted in a fixed two-dimensional space that is mathematically created by conducting a multi-dimensional scaling routine. Before the multi-dimensional scaling was performed, the data were sphere-normalized so that students with more discussion are not weighted more heavily than people who have less discussion but still used the same configuration of connections in their discourse. A t-test was used to determine if there were statistically significant differences between the centroids of high and low performing groups. Shaffer and Ruis (2017) discuss the mathematics ENA in more detail.

Notebook entries

Notebook entries that students completed at the end of the activity were analyzed to assess individual design performance in the virtual internship. These notebook entries served as formative assessments of the previous activity. However, the assessment criteria used by design advisors during the implementation lacked depth and validity and thus, in this study, a new rubric was developed and used to assess students' design thinking. The notebooks contained pre-determined sections, "List of five prototypes" and "Justifications for the selection of five prototypes," which were used to segment the notebook data for coding. The assignment prompt, which was an email from the supervisor, is shown in the Appendix.

A variation of the Delphi method (Dalkey, 1969; Skulmoski et al., 2007) was used to develop a rubric for the quality of student entries to support the validity of the rubric. For this study, two domain experts—an engineering educator and a professional engineer—examined the notebooks and developed a rubric that identified high quality design work (Table 3). The two experts then met with me to discuss and clarify the rubric. The use of scoring rubrics in engineering design education have been shown to be effective in terms of improving inter-reliability and providing well-defined criteria for assessing individual engineering design skills (Bailey & Szabo, 2006). Moreover, for validity and reliability purposes it is important to have multiple judges with expertise in engineering design and education score student artifacts and compare (Bartholomew, 2017; Bartholomew et al., 2018).

After the rubric was developed, the two experts and one researcher coded a sample of 48 notebook entries using this rubric. Then, the researcher developed an automated coding algorithm to code the remainder of the notebooks. To assess the validity of the rubric, inter-rater reliability was measured using Intraclass correlation (ICC). The inter-rater reliability results for the notebook entries showed that the average ICC metric for a two-way multi-rater agreement was .75 with a 95% confidence interval from .64 to .83 ($F(47,142)=13$, $p<.001$).

The median score of the notebook entries was calculated and used to classify students as high and low-performing. If a student received a notebook score equal to or higher than the median score of 4, then they were identified as high-performing and if a student received a notebook score lower than the median score of 4, then they were identified as low-performing, which resulted in 120 low-performing students and 194 high-performing students.

Iacobucci et al. (2015) claim this approach of a "median split" is appropriate when focusing on group differences, such as low and high performing students in this study.

Table 3 Rubric for evaluating quality of individual notebook entries

Category	Good (2 points)	Fair (1 point)	Poor (0 points)
Team skills	Acknowledgement of team member contributions Mention team member(s) by name and provide record of accomplishments	Team is mentioned but team member contributions not noted	No mention is made of team
Technical resources	Uses (cites) 2 or more technical reports	Cites 1 technical report	No report citations
Internal consultant citations	Uses (cites) all five consultant requirements or preferences	Cites at least 1 consultant	No consultants cited
Design justification	Makes justifications based on both quantitative and qualitative analysis	Makes justification based on EITHER quantitative OR qualitative analysis	Unjustified design decision
Testing plan logic	An approach to new design testing is coherent—using a control device for comparison or testing extremes. Has a full testing plan that is clear for all five devices	Approach to design testing is partial. Plan is not applied to all five devices	No clear approach

Max = 10, Min = 0

However, to support the use of a median split, researchers should provide empirical evidence and justification. In this study, students with a score of 4 were included in the high group because (1) not including them would ignore a large number of the students and the power in this study would be greatly reduced, (2) putting them in the low group would cause the data to be more unevenly split than putting them in the high group, and (3) an empirical test was conducted to show that the ENA results do not significantly change when students are grouped into a high (scores = 5, 6, 7, 8; n = 77), medium (score = 4; n = 117), and low (scores = 0, 1, 2, 3; n = 120) group. The results of the test showed that there were no significant differences between the medium (M = .08, SD = .58) and high (M = .12, SD = .61) groups' mean centroids ($t(157) = .50$, $p > .05$; $d = .07$), but there were significant differences between the medium and low (M = − .15, SD = .46) groups' mean centroids ($t(220) = 3.32$, $p < .05$; $d = .43$) and the high and low groups' mean centroids ($t(130) = 3.33$, $p < .05$; $d = .52$). Thus, the medium and high were combined together into one group labeled as "high" (n = 194) and the low group remained "low" (n = 120).

Results: qualitative and quantitative analyses of students

This section addresses RQ 2: *Are there differences in how low-performing and high-performing students make connections among moves and rationale in a virtual internship?* To investigate the applicability of connected design rationale to student learning, the same codes were applied to the discourse available from the virtual engineering internship. Discourse was analyzed from one design activity in *Nephrotex* in which students chatted with their team members to discuss their research thus far and decided collectively on five prototypes to send to the lab for testing. First, this study presents a qualitative analysis of the chat discourse of one representative low-performing student, Grace, and one representative high-performing student, Levi, to provide an in-depth examination of connections among moves and rationale in the virtual internship. These two students were chosen because their centroid was located in close proximity to the mean centroid of their respective group. Thus, these students' networks were similar to the average network of their group and provided a representative visualization. Second, a quantitative ENA analysis was performed to compare high and low-performing students to experts' patterns of discourse.

Low-performing student: Grace

After individually reading *Nephrotex* research reports on the various design parameters, students held a meeting in the online chat tool with their team members to discuss what they had learned so far and to decide on a batch of five devices to submit to the lab for testing. Grace, a low-performing student, was in a group with four other individuals: David, Jared, Matthew, and Austin.

At the start of discussion Jared asked, "OK so which prototypes should we use for our batches?" Austin advocated for one of Grace's prototypes which used a hydrophilic surfactant, which he claimed was the most reliable and the cheapest surfactant choice. David continued the conversation:

David: I'm hearing hydrophilic so that sounds like our best bet.
Austin: but the biological one has a low percentage of blood cell reactivity which is good
Grace: Are you talking about making new prototypes [or] are you still looking at the already made ones?

Jared: I think we should stick with the ones that are already made since that would make it easier.

David and Austin offered suggestions for which surfactant to choose. Grace did not offer suggestions for a surfactant, but instead asked a clarifying question about whether the team was making new prototypes. In this excerpt, Grace connected the move of asking questions: “Are you talking about making new prototypes [or] are you still looking at the already made ones?” to David’s move of making a design decision and to Austin’s justification based on performance parameters.

Later in the discussion, the team decided to use Grace’s previously designed prototype as one of the devices to submit for testing. Austin asked Grace to explain her reasoning for choosing 4% carbon nanotube for her device:

Austin: Reason for going with 4% nanotube?
Grace: Just because it was the highest percentage that I could see data for and therefore the highest I could trust 100%

In this moment, Grace connected the rationale of performance parameter: “it was the highest percentage that I could see data for” to Austin’s move of asking questions. In contrast to the previous exchange, Grace answered a question instead of asking a question but still made connections to the move of asking questions.

Grace’s talk centered on asking clarifying questions and providing direct responses to questions asked by her team members. Visualizing her talk as a discourse network confirms this finding (Fig. 3). The connections in her discourse network focused on the move of asking questions, which she connected to making design decisions and justifications about communication and performance parameters. The strongest connections in her network are between the move of asking questions and making design choices and between the move of asking questions and the justification of performance parameters.

Fig. 3 Example of a low-performing student (Grace) Discourse network

High-performing student: Levi

In contrast to Grace's discourse network which had a central focus of asking and answering questions, Levi, a high-performing student, made connections among a variety of moves and rationale. Levi was on a team with three other individuals: Francesca, Priya, and Lee. When it was time to start the meeting, Levi initiated the discussion by typing, "Okay everybody, so we are looking at the prototypes in the FEEDS option under the tools tab." Levi continued the discussion:

Levi: So we don't have to type out all of the explanations for each prototype again, would you just like to put that notebook in the shared space file and then read each others' explanations and go from there? I think we might want to choose some of each based on explanations.

Francesca:	how do you do that?
Levi:	Go to the "individuals design 5 prototypes Notebook" under the notebook tab, on the top left of the notebook you should see a box that says "available in shared space" next to it. Just click that box.
Francesca:	okay i did
Levi:	Awesome! Okay so we'll read each other's justifications and then discuss prototypes.
Francesca:	okay sounds good!

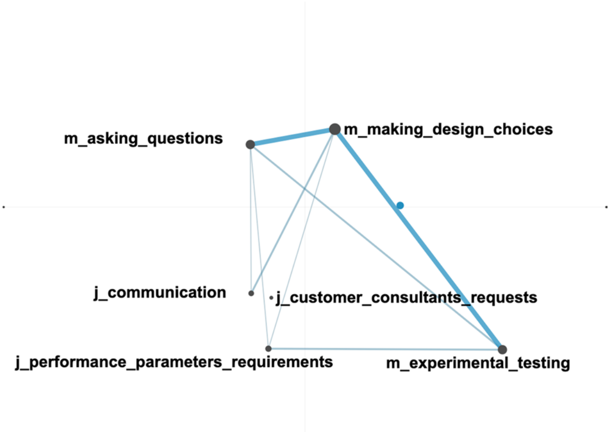
In this excerpt, Levi connected the moves of making design decision: "I think we might want to choose…" and asking questions: "Would you just like to… read each others' explanations and go from there?" to the rationale of communicating with his teammates: "So that we don't have to type out all the explanations for each prototype again."

After gaining access to their team members' notes, Francesca asked if the team should only choose two different surfactants, biological and steric hindrance:

Francesca:	Do you think we should pick either steric or biological for all of them? or choose some of each?
Francesca:	I feel like if we have a lot of different variants we wont have anything to really compare our results with
Levi:	I like the idea of using mainly steric hindrance because it was the most versatile surfactant and voted the best choice by our group previously, but i think we should try to include at least one prototype using a different surfactant to test the results of changing a surfactant.
Francesca:	okay that sounds good...do you want to do 3 and 2?
Francesca:	Want to do the three steric having 1.5% nanotube and then do one vapor, one dry-jet wet, and one phase?
Levi:	Sure, that sounds good if we can find enough similarities between at least two designs to justify comparing the results of each to each other. Like each design has a different design that varies by only one factor so we can compare results.

In this excerpt, Levi connected the moves of making design decisions: "I like the idea of using mainly steric hindrance…" to the rationale of performance: "because it was the most versatile surfactant and voted the best by our group previously." These specific connections

Fig. 4 Example of a high-performing student (Levi) Discourse network



were made within his own utterance. However, these statements also connected to Francesca's previous utterances in which she asked a question about choosing design specifications and proposed an experimental approach.

Levi made a variety of connections among moves and rationale, focusing on the move of making design decisions. Visualizing his talk as a discourse network confirms this finding (Fig. 4). His strongest connections were among making design decisions, asking questions, and suggesting an experimental approach for testing. Similar to Grace, Levi connected asking questions to justifications centered on better team communication, but these connections were not as common as the ones focused on making design choices and conducting experimental tests. Because Levi's network more strongly focuses on making design decisions based on performance parameters, it more closely resembles the discourse of the expert engineers than Grace's network.

Network analysis

This section addresses RQ 3: *How do low-performing and high-performing students' connections among moves and rationale in a virtual internship compare to expert engineers' in a real-world setting?*

The analysis was expanded to examine the discourse networks of 314 virtual internship students and whether high-performing students have patterns of connections more like those of experts than low-performing students.

First, mean (average) network representations were created for the low and high-performing students. Then, a subtracted network was created between the mean networks of the high-performing and low-performing students. To create a subtracted network, the weights of each corresponding link in the two mean networks are subtracted to obtain one network (Fig. 5). The subtracted network suggests high-performing students had more strongly weighted connections on average than low-performing students because the largest weight differences are in favor of the high-performing students, as indicated by the prominence of thick blue lines. These connections were the move of making design decisions and the move of experimental testing, the move of asking questions and the move of experimental testing, and the move of making design decision and the rationale of customer requests. In contrast, the low-performing students had smaller weight differences in their

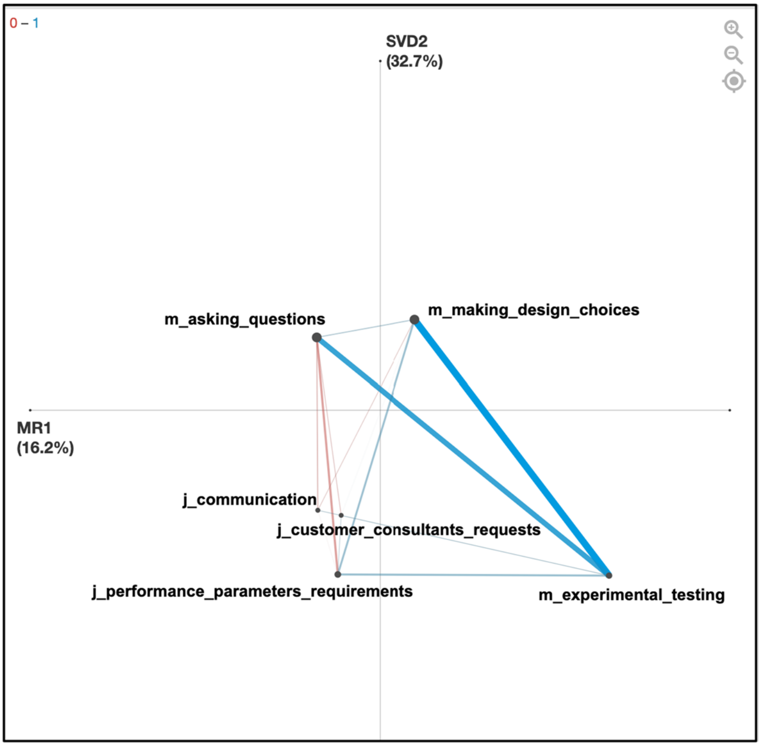


Fig. 5 A subtracted network comparison between high-performing (blue) and low-performing (red) student discourse networks. This network shows the weight difference between the mean high-performing student network and the mean low-performing student network. (Color figure online)

favor, as indicated by the thin red lines: the move of asking questions and the rationale of performance parameter, the move of asking questions and the rationale of communication, and the move of making design decisions and the rationale of communication. This indicates that low-performing students made slightly more connections among these connections than high-performing students, but the results more so suggest that low-performing students are *lacking* the connections that high-performing students were making.

The analysis of the centroids of the networks confirms and extends these results statistically (Fig. 6). The results of the multi-dimensional scaling show that the first dimension (x-axis) accounts for 16.2% of the variance in the data, and the second dimension (y-axis) accounts for 32.7% of variance in the data. A statistical t-test was conducted on the centroids of high and low performing student networks. High-performing students ($M = .10$, $SD = .50$) had significantly higher discourse network centroids in the x-direction than low-performing students ($M = -.16$, $SD = .40$; $t(290.45) = 4.98$, $p < .001$) with a moderate effect size ($d = .55$). Thus, high-performing students focused on connections to performance parameter rationale when making design decisions and running experimental tests. These patterns were similar to those made by experts.

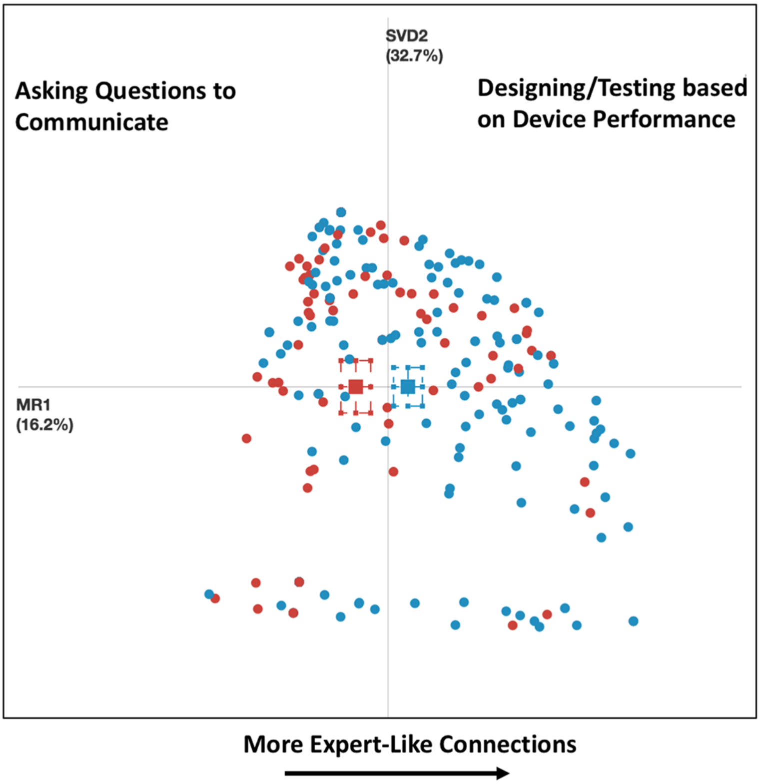


Fig. 6 Virtual internship students' centroids. Plot shows significant differences between low (red) and high (blue) outcome virtual internship students on the x-axis. (Color figure online)

Discussion

Connections to the learning sciences and engineering education

Design is a central activity in engineering and, thus, is central to engineering education. One space for effective design learning is virtual reflective practicums in which student novices interact with tools and experts to learn the ways of the practice. However, a significant issue in virtual design education is how to measure and evaluate design learning as it emerges in authentic practices. This study proposes one approach for how design learning can be measured and evaluated in authentic practice during virtual design education.

The results in this study from the real-world internship investigation revealed that experts made a variety of connections among moves and rationale. Connections made the most often were specific to the domain of design, such as making design decisions in order to meet performance parameters and making design decisions based on the customer's preferences. Using the same coding scheme of moves and rationale from the real-world study, the results from the virtual internship investigation showed a significant difference between high and low-performing students in terms of their patterns of connections. The networks suggested that high-performing students made connections that were more like those of experts.

The analysis of the professional engineer and student discourse in the real-world and virtual settings respectively showed that all participants made moves (Schön, 1987) and provided rationale (Rittel, 1988) in some form. In the domain of engineering education, this study aligns with Atman and colleagues' (Atman et al., 1999, 2000, 2008) claims that expert engineers prioritize identifying customer and performance constraints while novices primarily focus on effective communication and planning. The experts in their study were professional engineers with an average of 19 years of experience in the field, mostly white males, and from a variety of engineering disciplines. In the domain of learning sciences, the study's claims about learning are similar to diSessa (1993)—that expert knowledge systems have more reliable and productive connections among knowledge elements than novice knowledge systems. The progression from novice to expert occurs through the reorganization and refinements of connections in the knowledge system. However, the results suggest that knowledge systems that develop in immersive virtual design learning environments, such as Nephrotex, also incorporate the ways in which people in the domain talk, act, and interact that are socially meaningful. Developing expertise requires more than acquiring knowledge; it requires developing ways of knowing, being, and doing that are particular to a professional culture and enacting the relationships among Discourse elements that are characteristic of the community (Shaffer, 2006). In this study, these ideas of sociocultural learning, specifically in regard to professional cultures, have been applied to design thinking through the examination of relationships among design moves and rationales to propose a connected design rationale framework.

Connected rationale as a theoretical lens for research and practice

The results in this study provide an example of how a connected design rationale model measures student design learning in a situated, authentic environment. This is a significant contribution to learning sciences and engineering education because, as Atman et al. (2008, p. 309) argue, "Research exists on engineering students' knowing and thinking, yet how it is enabled through discourse and a community of practice is not well understood." This study is a step towards the goal of learning more about how the design learning process works when co-constructing knowledge in a social environment because it provides one example of measuring expert and novice design discourse in a connected manne.

More pragmatically, a connected design rationale could be a useful theoretical lens for the future design of learning environments and assessments and in particular, for evaluating collaborative design activity in virtual environments. Learning objectives and assessments can be designed to purposefully guide students toward creating a complex web of moves and rationale when engaging in collaborative design thinking. For example, if one requirement is for students to develop a testing plan, an instructor can provide opportunities for students to provide rationales for such a testing plan either by submitting an engineering notebook or engaging in discussions with teammates. Consequently, instructors can create rubrics that do not measure isolated instances of design skills, but rather resemble a web of connected design rationale consisting of potential moves and justifications that would constitute various levels of expertise in engineering design work. In this way, instructors can evaluate the strength of students' cohesive arguments and how their reasoning mirrors professional engineering thinking.

Epistemic network analysis as a tool for measuring design learning

This study also adds to the collection of work which demonstrates the usability of ENA as a tool to measure connections in discourse and to measure complex thinking (Arastoopour et al., 2014; Nash & Shaffer, 2011; Svarovsky, 2011). As the results suggest, the key affordances of ENA are its ability to visualize the co-occurrences of qualitative codes through network representations, analyze a large sample size through centroid representations, and perform statistical tests to draw quantitative conclusions about connections made in qualitative data. Using these features in this study, ENA reproduced and highlighted patterns of connections among moves and rationale in discourse and modeled a connected design rationale in a virtual simulation environment.

Although ENA was used specifically as a research tool in this study, Atman et al. (2008) claim that engineering education research tools may be used as assessment instruments to help guide instruction if adopted in a useful manner. In the context of virtual learning environments, ENA may be useful for assessment of complex thinking. Virtual learning systems tracks large quantities of chat logs, notebooks, and other forms of digital artifacts, which would be difficult and time-consuming for educators to assess manually. As a solution, instructors can use ENA to visualize connections among moves and rationale in student discourse and interpret models of student learning instead of creating the models themselves. Such connected design rationale network models can help instructors to better understand individual and collaborative design practices in their classrooms, which may result in an instructor's decision to reconfigure teams, change the pace of the course, or plan new lessons to facilitate certain patterns of connection-making in students' discourse networks.

Some initial work has been conducted adapting ENA for assessment use in the classroom. Herder et al. (2018) designed a teacher interface embedded within the virtual internship that displayed real-time connections made by students. Through observations and interviews with teachers, the researchers discovered that teachers did not use the interface to assess students very often, although they expressed a desire to use it. Mainly, the information presented was too abstract for teachers to interpret while managing their classroom environments. However, teachers conveyed that the advantage of the ENA models was that they maintained the complexity of student learning through dynamic network representations instead of only numeric values. Thus, there are significant challenges that remain with adapting ENA as a real-time assessment tool for authentic virtual learning, but it is a promising avenue for further research.

Limitations

There are several limitations in this study. First, this study only examined the discourse of two professional engineers in the real-world internship environment. The benefit was in an in-depth qualitative analysis of the two experts' practices and discourse. However, this focus on only two engineers greatly limited the generalizability of the findings. Expert practices vary across the sub-cultures of professional engineering and among different individuals. Thus, this study offered one worked example of how a connected design rationale approach could be used to measure design learning in one specific context but does not necessarily generalize to other engineering learning scenarios.

Relatedly, the generalizability of this study is also limited because of the demographics of the student population: mostly white and male. Although the demographics of the undergraduate population in the virtual internship was fairly representative of all engineering undergraduates in the U.S., this breakdown suggests that historically dominant Discourses were present and other forms were not accounted for. This limitation is especially important to identify in the domain of engineering, which has been historically dominated by white males and has largely excluded men and women of color and white women (Pawley, 2017). In a more recent study, there has been some investigation with this study's dataset on women of color's experiences and discourses within Nephrotex (Arastoopour Irgens, 2019).

In short, the intention of this study is not to propose an absolute measurement of design learning but rather, to propose one approach for measuring design learning through a connected design rationale lens and using ENA. In future studies, this proposed approach may be applied across different populations with more nuanced Discourses to explore its limitations and boundaries.

Conclusion

This work proposes innovative and aligned theoretical and methodological approaches for investigating design thinking and learning. This study merged learning sciences and design education research by proposing a connected model of design learning that examined the interactions among moves and rationale during the design process. Beginning with a qualitative investigation of real-world engineering design practices, this study identified professional engineers' patterns of moves and rationales when mentoring interns. Using these identified moves and rationales as a grounding for investigating design practices in a virtual internship, the results revealed differences between students and these differences were associated with an independent outcome measure. The results suggest that a connected design rationale model distinguishes between levels of expertise in virtual simulation environments and can be a useful approach for both learning sciences and engineering education research for a better understanding of the development of design thinking—a cognitive process fundamental to twenty-first century education.

Appendix

See Fig. 7.

USERNAME,

During today's work session meet with your project team via chat to choose which designs to test. Because actually building and testing a filtration membrane is very expensive, we test early-stage designs with a sophisticated virtual simulator.

As a team, you should select five designs you would like to test. You can choose from among the prototypes that team members have already created, or you can design new ones based on your conversation. This should be a meeting with just the interns, but your design advisor will be present to answer any questions you may have. Remember, the goal of these tests is to figure out how different design parameters will affect the membrane attributes that Nephrotex's internal consultants prioritize. Once you agree on 5 designs to test, collect your designs into a batch in FEEDS.

After you meet and select your designs, each team member should create a notebook entry including:

1. A list of each prototype and its specifications
2. A justification for each prototype explaining why it was selected for testing
3. An attachment of your team's batch

Your team needs to submit your notebooks by DUEDATE. Once all team members have submitted their notebook entries, I will upload your team's 5 designs to the server for testing. Because many Nephrotex researchers (and a number of our affiliates) use our virtual simulators to run tests, it may take a day or two to get results back.

Have a productive meeting! Alex

Fig. 7 Prompt that was given to *Nephrotex* students in the design activity for writing in their engineering notebooks

Acknowledgements This work was funded in part by the National Science Foundation (DRL-1661036, DRL-1713110), 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.

Declarations

Conflict of interest The author declares no conflict of interest.

Ethical Approval The data collection and analysis procedures employed in this study were granted approval by the Institutional Review Board at the University of Wisconsin-Madison.

References

Adams, R. S., Daly, S. R., Mann, L. M., & Dall'Alba, G. (2011). Being a professional: Three lenses into design thinking, acting, and being. *Design Studies, 32*(6), 588–607.

Arastoopour, G., Chesler, N. C., & Shaffer, D. W. (2014). Epistemic persistence: A simulation-based approach to increasing participation of women in engineering. *Journal of Women and Minorities in Science and Engineering, 20*(3), 211–234.

Arastoopour, G., & Shaffer, D. W. (2015). Epistemography and Professional CSCL Environment Design. In Lindwall, O., Häkkinen, P., Koschman, T. Tchounikine, P. & Ludvigsen, S. (Eds.) Exploring the

Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1 (pp. 204–211) Gothenburg, Sweden.

Arastoopour, G., Shaffer, D. W., Swiecki, Z., Ruis, A. R., & Chesler, N. C. (2016). Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. *International Journal of Engineering Education, 32*(2), 1492–501.

Arastoopour Irgens, G. (2017). Connected design rationale: Modeling and measuring engineering design learning [Ph.D., The University of Wisconsin—Madison]. In *ProQuest Dissertations and Theses* (1901451547). ProQuest Dissertations & Theses Global. https://www.proquest.com/docview/1901451547/

Arastoopour Irgens, G. (2019). Facilitating a sense of belonging for women of color in engineering. In R. D. Roscoe, E. K. Chiou, & A. R. Wooldrige (Eds.), *Advancing diversity, inclusion, and social justice through human systems engineering* (pp. 221–239). CRC Press. https://doi.org/10.1201/9780429425905

Atman, C. J., Adams, R. S., & Turns, J. (2000). Using multiple methods to evaluate a freshmen design course. In *30th Annual Frontiers in Education Conference. Building on a Century of Progress in Engineering Education. Conference Proceedings (IEEE Cat. No. 00CH37135)* 2000, pp. S1A/6-S1A13 vol.2, https://doi.org/10.1109/FIE.2000.896605.

Atman, C. J., Chimka, J. R., Bursic, K. M., & Nachtmann, H. L. (1999). A comparison of freshman and senior engineering design processes. *Design Studies, 20*(2), 131–152.

Atman, C. J., Eris, O., McDonnell, J., Cardella, M. E., & Borgford-Parnell, J. L. (2014). Engineering design education. In A. Johri, & B. M. E. Olds (Eds.), *Cambridge handbook of engineering education research* (pp. 201–226). Cambridge University Press. https://doi.org/10.1017/CBO9781139013451.015

Atman, C. J., Kilgore, D., & McKenna, A. (2008). Characterizing design learning: A mixed-methods study of engineering designers' use of language. *Journal of Engineering Education, 97*(3), 309–326.

Bagley, E., & Shaffer, D. W. (2009). When people get in the way: Promoting civic thinking through epistemic gameplay. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS), 1*(1), 36–52.

Bagley, E., & Shaffer, D. W. (2010). The epistemography of urban and regional planning 912: Appropriation in the face of resistance. In K. Gomez, L. Lyons, & J. Radinksy (Eds.), *Learning in the disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010)—volume 1 full papers (ICLS 2010): Vol. volume 1* (pp. 81–88).

Bailey, R., & Szabo, Z. (2006). Assessing engineering design process knowledge. *International Journal of Engineering Education, 22*(3), 508–518.

Bartholomew, S. R. (2017). Assessing open-ended design problems. *Technology and Engineering Teacher, 76*(6), 13.

Bartholomew, S. R., Strimel, G. J., & Jackson, A. (2018). A comparison of traditional and adaptive comparative judgment assessment techniques for freshmen engineering design projects. *International Journal of Engineering Education, 34*(1), 20–33.

Bergsman, K. (2018). The "Structured" engineering design notebook: A new tool for design thinking within a studio design course. In *2018 ASEE Annual Conference & Exposition Proceedings*, 31076. https://doi.org/10.18260/1-2--31076

Buchal, R. O. (2018). Collaborative digital design notebooks. In *Proceedings of the Canadian Engineering Education Association (CEEA)*. https://doi.org/10.24908/pceea.v0i0.10362

Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D. W., Hu, X., & Graesser, A. C. (2019). nCoder+: A semantic tool for improving recall of nCoder coding. In B. Eagan, M. Misfeldt, & A. Siebert-Evenstone (Eds.), *Advances in quantitative ethnography* (pp. 41–54). Springer International Publishing. https://doi.org/10.1007/978-3-030-33232-7_4

Chesler, N. C., Arastoopour, G., D'Angelo, C. M., Bagley, E. A., & Shaffer, D. W. (2013). Design of a professional practice simulator for educating and motivating first-year engineering students. *Advances in Engineering Education, 3*(3), n3.

Chesler, N. C., Ruis, A. R., Collier, W., Swiecki, Z., Arastoopour, G., & Williamson Shaffer, D. (2015). A novel paradigm for engineering education: Virtual internships with individualized mentoring and assessment of engineering thinking. *Journal of Biomechanical Engineering*. https://doi.org/10.1115/1.4029235

Dalkey, N. C. (1969). *The Delphi Method: An experimental study of group opinion*. RAND Corporation. https://www.rand.org/pubs/research_memoranda/RM5888.html

DiSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10*(2–3), 105–225.

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education, 94*(1), 103–120.
Eagan, B., Misfeldt, M., & Siebert-Evenstone, A. (Eds.). (2019). *Advances in quantitative ethnography: First international conference, ICQE 2019, Madison, WI, October 20–22, 2019, Proceedings* (1st ed. 2019 edition). Springer.
Eagan, B. R., Rogers, B., Serlin, R., Ruis, A. R., Arastoopour Irgens, G., Shaffer, D. W. (2017). Can We Rely on IRR? Testing the assumptions of inter-rater reliability. In Smith, B., Borge, M., Mercier, E., & Lim, K.Y. (Eds.), Making a Difference: Prioritizing Equity and Access in CSCL: Proceedings of the 12th International Conference on Computer Supported Collaborative Learning Volume 1 (pp. 529–532). Philadelphia, PA.
Febrian, A., Lawanto, O., & Cromwell, M. (2015). Advancing research on engineering design using e-Journal. In *2015 IEEE Frontiers in Education Conference (FIE)* (pp. 1–5). https://doi.org/10.1109/FIE.2015.7344191
Gee, J. P. (2011). *An introduction to discourse analysis: Theory and method* (3rd ed.). New York: Routledge.
Gero, J. S., & Kannengiesser, U. (2008). An ontological account of Donald Schön's reflection in designing. *International Journal of Design Sciences and Technologies, 15*(2), 77–90.
Glaser, B. G., & Strauss, A. L. (1967). *Discovery of grounded theory: Strategies for qualitative research*. Chicago: Adline Publishing Co.
Hatfield, D. (2015). The right kind of telling: An analysis of feedback and learning in a journalism epistemic game. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS), 7*(2), 1–23.
Hatfield, D., & Shaffer, D. W. (2010). The epistemography of journalism 335: Complexity in developing journalistic expertise. In K. Gomez, L. Lyons, & J. Radinksy (Eds.), *Learning in the disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010)—volume 1 full papers (ICLS 2010): Vol. volume 1* (pp. 628–635).
Herder, T., Swiecki, Z., Fougt, S. S., Tamborg, A. L., Allsopp, B. B., Shaffer, D. W., & Misfeldt, M. (2018). Supporting teachers' intervention in students' virtual collaboration using a network based model. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 21–25). https://doi.org/10.1145/3170358.3170394
Hod, Y., & Sagy, O. (2019). Conceptualizing the designs of authentic computer-supported collaborative learning environments in schools. *International Journal of Computer-Supported Collaborative Learning, 14*(2), 143–164. https://doi.org/10.1007/s11412-019-09300-7
Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015). Toward a more nuanced understanding of the statistical properties of a median split. *Journal of Consumer Psychology, 25*(4), 652–665. https://doi.org/10.1016/j.jcps.2014.12.002
Lee, J. (1997). Design rationale systems: Understanding the issues. *IEEE Expert, 12*(3), 78–85.
Lee, J., & Lai, K.-Y. (1991). What's in design rationale? *Human-Computer Interaction, 6*(3–4), 251–280.
Linn, M. C., Eylon, B.-S., & Davis, E. A. (2004). The knowledge integration perspective on learning. In M. C. Linn, E. A. Davis & P. Bell (Eds.), Internet environments for science education (pp. 29–46). Mahwah, NJ: Lawrence Erlbaum.
Lloyd, P. (2019). You make it and you try it out: Seeds of design discipline futures. *Design Studies, 65*, 167–181. https://doi.org/10.1016/j.destud.2019.10.008
McCall, R., & Burge, J. (2016). Untangling wicked problems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AI EDAM, 30*(2), 200–210. https://doi.org/10.1017/S089006041600007X
Moyne, M. M., Herman, M., Gajos, K. Z., Walsh, C. J., & Holland, D. P. (2018b). The development and evaluation of DEFT, a web-based tool for engineering design education. *IEEE Transactions on Learning Technologies, 11*(4), 545–550. https://doi.org/10.1109/TLT.2018.2810197
Moyne, M., Herman, M., Walsh, C., & Holland, D. (2018). An evaluation of an engineering design class using mixed methods techniques. In *2018 ASEE Annual Conference & Exposition Proceedings*, 29782. https://doi.org/10.18260/1-2--29782
Nash, P., & Shaffer, D. W. (2011). Mentor modeling: The internalization of modeled professional thinking in an epistemic game. *Journal of Computer Assisted Learning, 27*(2), 173–189. https://doi.org/10.1111/j.1365-2729.2010.00385.x
Pawley, A. L. (2017). Shifting the "default": The case for making diversity the expected condition for engineering education and making whiteness and maleness visible. *Journal of Engineering Education, 106*(4), 531–533.
Rittel, H. (1988). *The reasoning of designers*. Stuttgart: IGP.

Rockwell, J. A., Grosse, I. R., Krishnamurty, S., & Wileden, J. C. (2010). A semantic information model for capturing and communicating design decisions. *Journal of Computing and Information Science in Engineering*. https://doi.org/10.1115/1.3462926

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Routledge.

Schön, D. A. (1984). Problems, frames and perspectives on designing. *Design Studies, 5*(3), 132–136.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* JosseyBass.

Schön, D. A. (1988). Designing: Rules, types and worlds. *Design Studies, 9*(3), 181–190.

Shaffer, D. W. (2006). *How computer games help children learn*. Macmillan.

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

Shaffer, D. W. & Ruis, A. R. (2017). Epistemic Network Analysis: A Worked Example of Theory-Based Learning Analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic, Eds., Handbook of Learning Analytics (First Edition, pp. 175–187). Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17.015

Siebert-Evenstone, A. L., Arastoopour Irgens, G., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017). In search of conversational grain size: Modeling semantic structure using moving stanza windows. *Journal of Learning Analytics, 4*(3), 123–139.

Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge: MIT Press.

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education: Research, 6*, 001–021. https://doi.org/10.28945/199

Strauss, A., & Corbin, J. (1994). Grounded theory methodology. *Handbook of Qualitative Research, 17*(1), 273–285.

Strauss, A., & Corbin, J. M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks: Sage Publications

Svarovsky, G. N. (2011). Exploring complex engineering learning over time with epistemic network analysis. *Journal of Pre-College Engineering Education Research (J-PEER), 1*(2), 4.

Tener, R. K., Winstead, M. T., & Smaglik, E. J. (2001). Experiential learning from internships in construction engineering. *ASEE Annual Conference Proceedings*. 4889–4920

Yoder, B. L. (2017). *Engineering by the numbers*. American Society for Engineering Education.