Examining faculty choices while implementing the Next Gen PET curriculum through Revealed Causal Mapping

Julia Willison (she/her)
Department of Physics, University of Central Florida, 4111 Libra Drive, Orlando, FL, 32816

Erin M. Scanlon (she/her)
Department of Physics, University of Connecticut - Avery Point, 1084 Shennecossett Rd., Groton, CT 06340

Jacquelyn J. Chini (she/her)
Department of Physics, University of Central Florida, 4111 Libra Drive, Orlando, FL, 32816

Next Generation Physical Science and Everyday Thinking (Next Gen PET) is a research-based, active-learning physical science curriculum for general education physics courses, with a focus on pre-service and in-service elementary school teachers. During pre-COVID introduction of this curricula at higher-education institutions across the country, instructors implementing the curriculum were recruited to participate in faculty online learning communities (FOLCs). In this project, we conduct a secondary analysis on transcriptions of recorded FOLC meetings through Revealed Causal Mapping (RCM), a qualitative technique to create causal maps composed of interconnected causal statements. This method allows us to examine instructors’ decision-making while implementing the Next Gen PET curriculum. In this paper, we share insights about how instructors decide to form groups, specifically group composition and frequency of group changes. We found that instructors may have their own preconceived ideas about the best ways to form groups, sometimes contrary to what current research suggests.
I. INTRODUCTION

Implementing new curricula takes time and effort from faculty. New curricula may try to be one-size fits all, but different faculty at different institutions have different wants and needs to best serve their students. Previous research has shown that when incorporating new curricula, instructors tend to make modifications to the curricular materials that are being implemented [1]. The Next Generation Physical Science and Everyday Thinking (Next Gen PET) curriculum is a university general education science curriculum with a focus on pre-service teachers [2, 3]. It was designed to be flexible to help accommodate variations in faculty and student needs [2]. Part of the flexibility of the Next Gen PET curriculum is the instructor’s ability to make explicit choices in their implementation of the curriculum [3]. This includes choosing a class style—lecture or studio—and which of the modules and activities to use [3].

To help faculty implement the Next Gen PET curriculum, faculty online learning communities (FOLCs) were created. Two to three faculty with experience working with the curriculum were assigned as facilitators to each FOLC. Each FOLC consisted of approximately 8-13 instructors.

The Next Gen PET curriculum developers and FOLCs emphasized the importance of several active learning strategies, including whole class discussion and group work. In this paper, we focus our analysis on discussions of group work in one FOLC, and specifically on the mechanisms behind creating groups of students and shuffling groups. All the participants in the FOLC we analyzed were using the studio-version of the Next Gen PET curriculum. Our research question was: What reasoning supports instructors’ decisions to implement a variety of groupwork strategies in the Nex Gen PET curriculum?

To answer this question, we utilized Revealed Causal Mapping (RCM), a mixed-methods analysis used to create causal maps composed of interconnected causal statements which can show paths of reasoning [4]. This method is particularly useful when looking at the logic of expert decisionmakers [4]. We consider instructors to be expert decisionmakers within the domain of their instructional experiences, and we look to elevate their experiences to inform curriculum and professional development.

A feature of RCM, like other mixed methods techniques is to “facilitate the transformation from qualitative inquiry to quantitative inquiry,” [5]. Teddlie and Tashakkori’s typology of mixed-methods would likely classify RCM as a monostand conversion design [6].

II. METHODS AND DATA

A. Data

The analysis presented here is a secondary analysis conducted on a pre-existing dataset consisting of verbatim transcripts of one Next Gen PET FOLC’s meetings during the 2017 – 2018 academic year, the first year these FOLCs ran. This FOLC met about once per week for one hour per meeting, with faculty participating every other week and facilitators swapping about every other week, so that at least one facilitator was present at each meeting. Author J.J.C. was a facilitator in a separate FOLC for this project. She also has used the Next Gen PET curriculum in her general education physics course for 8 years.

The focal FOLC was made up of thirteen faculty, including 3 facilitators. The group varied in gender (7 men, 6 women), race/ethnicity (9 white, 4 non-white), and institution type (8 Master’s degree-granting, 5 Ph.D. granting). Eleven instructors are represented in the data presented in this paper. The 2 instructors not represented here did not have any relevant causal statements within the dataset. The FOLC met 33 times in total. We analyzed transcripts for 22 of the 33 meetings, with analyzed transcripts sampled by convenience. The analyzed meeting transcripts span the entire school year. Previous research suggests that attitudes and content of FOLC meetings change across the course of the FOLC, so it was imperative that we encompassed the entire year with our analyzed data [7].

B. Methods

Our analysis was primarily conducted using the NVivo qualitative analysis software. The transcripts were divided among the three of us so that at least two coders analyzed each transcript.

Nine transcripts were utilized to develop and test our coding method in a two-stage process. First, four of the nine transcripts were used to develop the keywords to search for sections related to the focal practice of group work. This was done by reading through the transcripts for relevant sections and noting common words describing the instructional practice. The keyword list we developed is as follows, (* denotes any form of the word was used as a keyword):

- Group, dysfunc*, norm, expect*, eval*, account*, partner, disc*, turn, present*, exam, project, focus, incentive, slow, fast, check, goal, collab*, dynam*, split, facil*, activit*, work, concentrat*, task, behind, share, commun*, behav*, listen, learn, together, public, everybody, engag*

After the keyword list was developed and relevant areas of the four transcripts were identified, all three coders searched for causal linking words and identified the causes and effects together. This process allowed us to develop norms for identifying causal statements and their causes and effects. The next five transcripts were coded individually by two coders each to refine the norms and the coding process.

For the final thirteen transcripts, two coders individually used the keywords to distill the data down to the sections that discussed group work. Relevant sections were found in 20 out of 22 transcripts analyzed.
1. Cause and effect identification

The next step was to find causal statements within the relevant sections of data. Causal statements consist of a cause, a linking word, and an effect [5]. First, we searched for the causal linking words outlined by previous literature: if, then, because/’cause, so, since, think/thought, know/knew, use, believe, feel/felt, and which [5,8]. We only searched for the linking words within relevant sections of each transcript. Then, once all the causal linking words in the relevant sections were identified, they were manually reviewed to determine the cause and effect linked by each word, if any.

There are some uses of the causal linking words that are not causal in nature (e.g., “I give the Lawson Reasoning Test and then I kind of try to put a group of four...” uses a temporal, not a causal, “then”). These statements were not included in our data set. Causal statements that were not relevant to group work were also not included in our dataset.

To build trustworthiness of our analysis, two individual coders coded each transcript. To ensure inter-rater agreement (IRA) after individual coding, the two coders met and compared each relevant causal statement found within the text. In the event of a discrepancy, the differing statement was discussed between the coders until an agreement was reached to keep or not keep the statement. If an agreement was not reached, the statement was brought to the third researcher, who had not coded that section, to finalize a decision on the statement. In total, 37 statements were discarded through IRA, with 7 of those being from the Creating Groups set. The Creating Groups set contained 71 statements out of the 216 statements kept after the IRA.

The statements were then sorted by the classroom practice being discussed, and the statements in each section were numbered. From there, the causes and effects of each statement were aggregated into larger grain size categories to be used as nodes in the maps. Two coders each individually aggregated the larger grain categories, one via Grounded Theory, by line-by-line coding, then aggregating those codes into larger grain size categories and the other via content analysis [9,10]. Then, the categories were discussed between the two coders and the third researcher. The aggregated categories had many similarities between the two coders, so the smaller set of categories was chosen, with a few changes being made to reflect ideas from the larger set. For example, “When I first mix up the groups, I look for all the weak students and I put them together in a group,” and “I fall into the group of people that leans towards homogenous groups I think,” both fall into the node Group same level students. The aggregated statements formed 43 nodes in total, some of which are just causes, some just effects, and some that are both.

2. Causal mapping

After the causal statements were fully identified and classified into wider causes and effects, each coder proceeded to make an initial causal map. The maps consist of a cause, effect, or combination cause and effect as a node; and lines connecting individual nodes with arrows pointing in the direction from cause to effect. Some lines may have arrows in both directions, denoting that the causal relationship goes both ways in the data set. The maps were compared for accuracy and agreement. From there, two researchers each created a second draft of the map, with statement numbers and cause and effect numbers at each node. The maps were compared once again for accuracy and agreement.

III. FINDINGS AND DISCUSSION

The set of statements coded under the “Group Work – Creating Groups” instructional practice ultimately formed two connected but thematically distinct maps, Creating Groups and Shuffling Groups. Eleven of the thirteen FOLC members are represented in these two maps (i.e., the FOLC members’ causal statements, derived from the transcripts, are included in the nodes in the causal map). The Creating Groups map focuses on the strategies discussed for making groups, while the Shuffling Groups map focuses on the frequency of forming new groups throughout the semester.

Some nodes were not connected to the larger maps. Most are single cause-and-effect statements; most are presented in Fig. 3. They will be referred to as “isolated RCMs.”

A. Creating groups

Creating Groups is centered around three large nodes, the instructional practices C. Evenly mixed group level and O. Group same level students, and the faculty reasoning H. Group works well. The next largest nodes are L. Randomized groups, F. Student dissatisfaction, and P. Improve student effort. Five of the isolated RCMs were explicitly related to the Creating Groups RCM (see Fig. 3).

It is unsurprising that the Creating Groups RCM is dominated by C. Evenly mixed group level, O. Group same level students, L. Randomized groups, and H. Group works well. Ultimately, the goal for instructors is to have groups in their class work well. As one instructor stated, “If you group them up ... really seems to make all of them better, which makes me happy.”

This data shows that the path to groups working well is varied. Both C. Evenly mixed group level and O. Group same level students led to H. Group works well. Not only that, but they both lead to P. Improve student effort. Other curricula, such as Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP), call explicitly for evenly mixed groups, with the reasoning that the top students in a group can help assist the lower students, leading to an overall improvement [11]. Additionally, research in physics education suggests that varying student ability in groups allows students to support each others’ learning [12,13].
FIG 1 Creating Groups causal map. Nodes in this figure are lettered A – S., left to right, top to bottom. The numbers in the parentheses denote (# of statements, # of faculty). Yellow square nodes are instructional practices, and indigo oval nodes are faculty reasoning. The color saturation of the nodes denotes its frequency in the dataset, with lighter colors having fewer statements.

It is interesting to note that L. Randomized groups is not connected to H. Group works well in this analysis but was connected to F. Student dissatisfaction by a two-sided arrow. Looking back at the causal statements used to make up these nodes, we find that the instructor understood that the students were dissatisfied by randomized groups, and they also used that as their reasoning behind having randomized groups. The instructor stated: “But that way, they know it’s random, and if they hate their group, they know it’s not going to be a group forever.” The lack of connection to H. Group works well is even more interesting considering that literature suggests that randomized grouping’s main drawback is potential lack of group cohesion, along with the inability to equitably distribute underrepresented individuals [14,15].

Of the four instructors to discuss mixed-level groups and the two instructors to discuss same-level groups, only one instructor discussed both. The two instructors who discussed randomized groups discussed neither mixed-level nor same-level groups. This could imply that instructors have a preconceived notion about how to group their students. It could also explain instructors’ satisfaction in using same-level and randomized groups even though they are not supported by research as much as evenly mixed groups. The lack of grouping standards for the Next Gen PET curriculum allows for instructors to use those preconceived notions to make choices in their instructional practices. There are pros and cons to this, as instructors may be more comfortable with specific grouping methods in their classrooms, but that does not necessarily mean that those grouping methods are the most effective.

B. Shuffling groups

Shuffling Groups has one central node, AM. Shuffled groups; all but two nodes are direct causes or effects of instructors shuffling groups. The nodes with the next highest number of statements are AK. Concerns about student pacing and AJ. Good point in curriculum. Other than AJ. Good point in curriculum, all the causes of AM. Shuffled groups are related to students. Three are related to the instructors’ perception of the students, AK. Concerns about student pacing, AI. Difficult student, and AA. Students struggle. The last is AH. Student request, which is not related to the instructors’ perception of the students but could be considered as an instructor not perceiving an issue, causing the student to bring it to the instructors’ attention.

The effects of AM. Shuffled groups are also majority student oriented. Like the causes, we have nodes related to the instructors’ perception of the students, AC. Minimize student dissatisfaction, AF. Students build community and AD. Students on topic. Two of these, AC. Minimize student dissatisfaction, match the instructors’ perception of the students. Other than AC. Minimize student dissatisfaction, all the effects of AM. Shuffled groups are also majority student oriented.
Creating Groups Isolated RCMs

FIG 3. Creating Groups isolated RCMs. Nodes in this figure are lettered top to bottom, left to right within each isolated RCM. A new first letter denotes a new map. The numbers in the parentheses denote (# of statements, # of faculty). Yellow square nodes are instructional practices, and indigo oval nodes are faculty reasoning. Color saturation of the nodes denotes its frequency in the dataset, with lighter colors having fewer statements.

dissatisfaction and AF. Students build community are related closely to the affect of the students, as with the causes AA. Students struggle and AI. Difficult student. This may suggest that instructors shuffle groups because they note a negative affect of their students, and they hope that by shuffling, they can see an uptick in positive affect of their students.

One isolated RCM was explicitly tied to the Shuffling Groups map. This RCM reads “If an instructor changes the curriculum, then they didn’t change groups, so the group works well.” Group works well and Changing the curriculum each had one statement from one instructor, while Didn’t change groups had two statements from two instructors.

Interestingly, the last of the isolated RCMs, one not explicitly tied to either map, reads “Instructors heard about a practice in their FOLC, so they tried that strategy.” Each node was found in two statements from two instructors. This causal statement is interesting because it shows that instructors are using the FOLC to field new instructional practices, suggesting that the FOLC was valuable as a learning community for these instructors.

IV. TAKEAWAYS

For instructors implementing the Next Gen PET curriculum, creating groups is a challenge. As can be seen from the RCMs, there are many variations in how faculty create groups, even though faculty have similar reasoning for using different grouping techniques. Ultimately, what faculty want are groups that work well where the students are not struggling.

While education research may have ideas about what grouping practices are the best, faculty may be staunch in their preconceived notions of grouping. This can make it difficult when curricula call for a particular grouping method, but faculty have their own modifications that they apply to the grouping methods.

A. Limitations and future work

This project has several limitations. The RCM method is frequently used on formal dialogue, such as company memos or political speeches [4]. Previous research in the physics education sphere has used RCM on instructor interviews [8]. Interview dialogue tends to be carefully thought out by the speaker, and interviewers are careful to not cut off or interject ideas into their interviewees’ speech, and can ask for the interviewee’s reasoning. The transcribed FOLC meetings used here are informal faculty discussions. This made identifying causal statements and their causes and effects difficult, particularly when the causal statements were spread across several turns of speech between different FOLC members.

Another limitation is the size of the dataset. This dataset encompasses a single cluster with thirteen FOLC members, and only eleven of those members are represented in the data presented in this paper. While this is a sufficient sized dataset for a case study on this particular FOLC, a larger dataset would be needed for any quantitative analysis or generalizations beyond this FOLC.

Future work on this project includes creating maps for the other facets of group work that were coded (group management, group assessment, and group activities,) and for whole class discussion. These maps will help Next Gen PET curriculum developers understand the various ways that faculty are implementing the curriculum. Additionally, RCM is intended to identify potential topics of future study. For example, future research could investigate why faculty seem to be staunch in their group formation methods or expand upon the lack of connection between Group works well and Randomized groups in the creating groups map, particularly considering the connection to student dissatisfaction. Does choosing randomized grouping indicate just a lack of time or more a lack of motivation to have cohesive groups? Future work could also include a comparison with other curricula that have had similar analysis done on faculty reasoning behind instructional practices.

ACKNOWLEDGMENTS

This work was funded by NSF DUE-1626496. Thanks to the Next Gen PET team for supporting this work.


