Identifying and analyzing actions of effective group work

Jennifer Keil, Rebecca Stober, Emily Quinty, Bridget Molloy and Nicholas Hooker

University of Colorado, School of Education
249 UCB, Boulder, Colorado, 80309

Abstract: This study investigates the characteristics valued by teacher-researchers in collaborative group learning. Video data of groups engaging in the Physics and Everyday Thinking High School activities was collected from one teacher implementing the curriculum, and salient characteristics of group interactions were identified. These characteristics were organized into codes. The frequency of these codes was analyzed for trends. Based upon the code frequency analysis, the group that teacher-researchers identified as most effective engaged in more sense-making, displayed greater equity of voice, and demonstrated more equitable contributions. These findings may have implications for what actions teacher-researchers value in student groups. These valued behaviors may provide potential for interventions to improve group effectiveness and student achievement.

PACS: 01.40.-d, 01.40.Fk, 01.40.Ha, 0.14.Ek

I. INTRODUCTION

The Physics and Everyday Thinking High School (PET-HS) curriculum activities require that students regularly collaborate in small groups to complete tasks, such as collecting and interpreting evidence and model building [1,2,3]. Discourse is a central practice with the PET-HS curriculum and high school students learning in this context have demonstrated growth in their use of argumentation strategies and in their willingness to ask questions and challenge other students’ ideas [4]. Incorporating student discourse may improve student understanding by providing opportunities for students to share and discuss their ideas [5]. Additionally, during group discussions in-depth scientific thinking and sense-making may occur [6]. Sense-making involves students delving into the class activity to understand the experiment itself and to reveal the underlying scientific principles which explain their observational evidence [7].

Three high school science teacher-researchers implementing the PET-HS curriculum (first three authors) anecdotally concluded that small groups in their physics classes had varying levels of “effectiveness.” Without clear criteria to define what it means to be “effective,” it is difficult to assess group work and determine possible interventions to improve group interactions so that students achieve desired goals. To better understand group work, this research group1 engaged in inductive analysis of small group discussions [8] by characterizing group verbal interactions and behaviors to their snap judgments of whether or not the group interactions were considered “effective.” Specifically in this study researchers sought to answer the question: How do groups engage and what do teachers value as effective group behaviors?

II. RESEARCH CONTEXT

This study used data from one section of one teacher-researcher’s class taught in a small, urban public high school where 67% of students qualified for free or reduced lunch programs. The ethnic diversity of the sample was 70% Hispanic, 20% Asian, and 10% White.

A. Classroom setting

Students in this study were taught by the same teacher for two years, first in 11th grade chemistry and then in 12th grade physics. In both the physics and chemistry courses, the class structure involved small-group work throughout all components of the learning cycle. Each learning cycle began with students identifying and discussing their initial ideas and background knowledge with group members. They then performed experiments together and analyzed their results in small-group discussions. Finally, each group shared and modified their thinking in a whole class discussion. Throughout this process, the instructor’s role was to help students clarify their ideas through questioning rather than presenting them with the answers.

The physics course was taught with the PET-HS curriculum. The physics activities and classroom structures were designed for students to drive their own learning with support, as necessary, from the instructor. Throughout the year, the instructor “regrouped” members of the small groups approximately seven times, using a mixture of homogeneous, heterogeneous, and random grouping

---

1 Members of a Teacher Research Team: three practicing teachers (NSF Master Teaching Fellows), one undergraduate pre-service teacher, and one former high school teacher.
strategies. Given these parameters, students knew all members in the course and were acclimated to working with a variety of individuals. At the time of the study, the students were randomly grouped.

III. METHODOLOGY

A. Data collection

Three 10-minute videos, from a single PET-HS class were selected for analysis within the same instructional unit on magnetism. Videos were taken of three different groups, which consisted of 3, 4, and 4 students, respectively. The number of participants totals 11, however only 10 distinct individuals participated as a result of video filming on different days, during which one student in Group 2 was randomly added to Group 3. This student was treated as an independent member of both groups. The videos were chosen randomly from a set of videos recorded on different days during the unit. All three videos were from the experimental portions of the learning cycle in which students engaged in analyzing experimental evidence and inducing scientific principles through carefully sequenced questions. Prior to coding and analysis, each teacher-researcher independently watched each video once and rated each group as either effective or ineffective. Researchers named these ratings snap judgments, which were later compared to analyzed data.

B. Data analysis

Each video was transcribed and coded in three ways: (1) each student’s verbal utterances (each new turn of speech was considered an utterance, utterances with teacher were not counted) were counted, (2) characteristics of the utterance (e.g. off-task, experimentation) were coded using a generative coding scheme, and (3) physical behaviors (e.g. leaning toward others, looking away) were coded using a generative coding methodology [8]. After each teacher-researcher coded the videos independently, the group met to go over each line of transcript and reconcile codes. Researchers then grouped the codes into five categories to comprise the common coding scheme: sense-making, experimentation, documentation, procedural, and off-task behavior (see Table I for examples). Each teacher-researcher then independently re-coded each of the transcripts using the common coding scheme, and then met again to negotiate and resolve differences until all researchers agreed 100% on all the coded data.

The amount and percentage of time each group spent on each code was determined. Since both verbal and physical interactions were coded, it is possible students were engaging in multiple coded behaviors simultaneously. For example, a student conducting the experiment could be verbally sense-making at the same time, or one student could be off-task while the other group members participate in sense-making. The percentage of total time at least one group member spoke (compared to time in silence) was calculated, as well as the percentage of total time that each individual group member spoke. From the total time spent speaking by each group, researchers constructed an “equity of voice” profile for each individual. They also calculated the amount of time each group worked together on a single task.

Finally, the coded data were compared to the initial snap judgments made by teacher-researchers, and conclusions were made regarding what the teacher-researchers valued in group work.

<table>
<thead>
<tr>
<th>Code</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimentation</td>
<td>Stands up, grabs the rubbed nail, and holds the head of the rubbed nail to the head of the floating nail.</td>
</tr>
<tr>
<td></td>
<td>Brings head of the nail close to the North end of the compass.</td>
</tr>
<tr>
<td>Sense-making</td>
<td>“Remember where we did the other experiment? When we did the unrubbed nail it still, like, did something because one of them was magnetized.”</td>
</tr>
<tr>
<td></td>
<td>Student 3: “Okay, so is that more like a magnetized nail or the unmagnetized nail”</td>
</tr>
<tr>
<td></td>
<td>Student 2: “Wouldn’t it be unmagnetized?”</td>
</tr>
<tr>
<td></td>
<td>Student 1: “Yeah, when we put it, it still did something, and this didn’t do anything.”</td>
</tr>
<tr>
<td>Procedural</td>
<td>“So, we already did step six?” (reading instructions) “Set the rubbed nail aside for a moment. Imagine that you...”</td>
</tr>
<tr>
<td></td>
<td>(reading instructions) “In the next steps, you will use a test tube filled with iron filings to help you build your model...”</td>
</tr>
<tr>
<td>Documentation</td>
<td>Sits down and writes in notebook.</td>
</tr>
<tr>
<td></td>
<td>Observed experiment, and then continues to write in notebook.</td>
</tr>
<tr>
<td>Off-Task</td>
<td>Looks over shoulder at a different group</td>
</tr>
<tr>
<td></td>
<td>Continues flipping through pages in notebook</td>
</tr>
</tbody>
</table>
IV. FINDINGS

In this section we present each finding individually, then synthesize them in the following section.

Finding 1: Snap judgments of group effectiveness by teacher-researchers were varied. All teacher-researchers classified Group 1 as effective and four of the five teacher-researchers classified Group 2 as ineffective. Classification of Group 3 was split, with three teacher-researchers classifying the group as effective and two as ineffective (Fig. 1).

Finding 2: Group 1 spent more time speaking than Groups 2 and 3. Analysis of percentage speaking time indicated that Group 1 spent 83% of total time speaking, which was higher than Group 2 (51%) and Group 3 (65%), as shown in Fig. 2.

Finding 3: Students in Group 1 showed more equity of voice than Groups 2 and 3. Group 1 demonstrated greater equity of voice; during the times the group members spoke with one another, each student spoke 34%, 35%, and 32% of the time (Fig. 3). Groups 2 and 3 showed an imbalance of the amount of time each individual spoke to the group, which we defined as low equity of voice. In Group 2 and 3 only two students spoke a majority of the time.

Finding 4: Group 1 spent more time engaged in sense-making than Groups 2 and 3. As shown in Fig. 4, Group 1 spent 4.5 minutes out of the total 10 minutes engaged in sense-making, while Group 2 and 3 spent 3.7 and 2.8 minutes, respectively. Additionally, the amount of time that contributed to the code of experimentation, but one member spent more time experimenting (1.5 minutes) than the other two students, 0.7 & 0.9 minutes, respectively. Likewise, all three students participated in procedural tasks and documentation though at different levels (see Fig. 5). In Group 1, no off-task behaviors were noted (Fig. 4).

Finding 5: Group 2 spent more time engaged in experimentation, documentation, and off-task behaviors than Groups 1 and 3. As shown in Fig. 4, the amount of time Group 2 spent in behaviors other than sense-making was larger than for Group 1 and 3. Notably, Group 2 spent about 2.1 more minutes engaged in documentation and about 2.5 more minutes in off-task behavior than either Group 1 or 3.
Group 1 and 3. Although Group 2 spent nearly as much time sense-making as Group 1, this was not distributed equitably between group members. As shown in Fig. 5, one of the students did most of the sense-making. This student also talked 65% of the time (Fig. 3). Similarly, one student spent more time documenting than the others, one of whom spent 0 minutes documenting (Fig 5).

**Finding 6:** Group 2 spent less time working together on the same task than Groups 1 and 3. Collectively, Group 2 spent only 5.9 out of 10 minutes working together to complete a single task. When one or two students were experimenting or sense-making, for example, the other group members were engaged in documentation or off-task behaviors. This group worked together much less time than either Group 1 or 3, which worked collectively on the same task for 10 and 9.1 minutes, respectively (Fig.6).

**FIG 6:** Time groups spent working on the same task.

V. CONCLUSIONS

In this study, all of the teacher-researchers rated Group 1 as effective (Fig. 1). Group 1 was the group that demonstrated the most equity of voice (Fig. 3), spent a higher percentage of time speaking (Fig. 2), spent the most amount of time sense-making (Fig. 4), and all their time working on the same task (Fig. 6). By selecting Group 1 as effective in their snap judgments, teacher-researchers may place high value on the qualities Group 1 exhibited: the time spent sense-making, high equity of voice of each group member and time spent on the same task. Each member of Group 1 engaged in comparable amounts of time sense-making (Fig. 5), indicating that teachers may also value equitable types of contributions. Equitable contributions may also indicate that each member had a vested interest not only in their own understanding, but of that of the group as well.

Four of the five teacher-researchers classified Group 2 as ineffective in their snap judgments (Fig. 1). This group only spent 51% of the time speaking to one another (Fig. 2), and of that time, member contributions were not equitable (Fig. 3). Group 2 showed the highest values in experimentation, documentation, and off-task behavior (Fig. 4). Further analysis reveals that an imbalance in the amount of time each member engaged in the different coded behaviors existed. The inequitable contributions may be due to the apparent distribution of self-selected group roles. The selection of Group 2 as ineffective by teacher-researchers may have been due to the high level of off-task behavior, which may have led to the oversight of other potentially valued behaviors. Group 3 was split amongst the teacher-researchers’ snap judgments, presenting a complex set of behaviors and qualities. It may be worth noting that the group rated as effective (group 1) had 3 members while the group rated as ineffective (group 2) had 4 members. Additionally, it might be of value to investigate the use of distributed group roles compared to shared group responsibilities structure as seen in Group 1.

Further analysis could be used to develop a methodology to assess group dynamics, as well as develop interventions to promote and cultivate effective group work. Such interventions may be leveraged to promote the skills and behaviors valued by teachers and observed in effective groups. Additionally promoting and improving these valued group attributes may lead to higher mastery levels and increased retention of physics concepts. Our findings suggest the snap judgments about group effectiveness made by teachers may align with group characterizations that tend to demonstrate higher sense-making, greater equity of voice, and more equitable contributions.

ACKNOWLEDGEMENTS

The authors thank the Streamline to Mastery program, support by NSF Grant # 937941, as well as Valerie Otero.