A Case Study of Group Processes and Student Evaluation of Teaching

Kristian G. Mortenson

University of St. Thomas, Minnesota, kgm@stthomas.edu

Rich S. Sathe

rssathe, rssathe@stthomas.edu
A CASE STUDY OF GROUP PROCESSES AND
STUDENT EVALUATION OF TEACHING

Kristian G. Mortenson and Richard S. Sathe
Department of Accounting, University of St. Thomas, St. Paul, MN, USA

This is an Author’s Original Manuscript of an article published by Taylor & Francis in Accounting Education on 09 Jan 2017, available online: http://tandfonline.com/doi/full/10.1080/09639284.2016.1274908.

ABSTRACT

This paper documents a case study undertaken to understand the effect of group processes on student evaluation of teaching (SET). The study used interviews to investigate the experiences of students in a cohort model Master of Science in Accountancy degree program and how those experiences influenced SET. The cohort served as an extreme example in which group processes played out intensely, allowing the researchers better to examine them. The results showed that the participants’ common experience led to frequent discussion about various aspects of the classes and the program. A common topic of discussion was instructor performance. This discussion about instructors, in turn, appeared to affect SET. It also appeared that a mismatching of expectations and realized student-instructor relationships also affected SET.

Key Words: Student evaluation of teaching (SET), cohort model, group processes
INTRODUCTION

Universities’ use of student evaluation of teaching (SET)\(^1\) measures engenders different reactions from instructors. On one hand, instructors understand SET’s value as a formative tool for feedback on their teaching. On the other hand, they are concerned over its use for summative purposes, especially related to career decisions like tenure and promotion (Spooren, Brockx, & Mortelmans, 2013).

SET plays an important role in universities worldwide. For example, Spooren, Brockx, and Mortelmans’ (2013) meta-analysis of 160 peer-reviewed journal articles on SET includes work by authors in 12 countries, including Hong Kong, Japan, UK, Australia, and Belgium. While originally developed to provide instructors feedback in order to improve their teaching (Surgenor, 2013), universities subsequently began to use SET as a tool for faculty performance, promotion, and tenure evaluation. A study of US accounting instructors reported that universities used SET for evaluation of teaching effectiveness (95%), tenure (87%), promotion (85%), and merit pay (70%) (Crumbley & Reichelt, 2009). More recently, universities have begun to use SET as an instrument for institutional quality-assurance and accountability, a practice that is gaining popularity in places like Australia (Smithson et al., 2015) and Europe, particularly Ireland (Surgenor, 2013). This latter use may bear on institutional decisions on the viability of programs and individual classes. These three dimensions (teaching improvement, faculty evaluation, and institutional accountability) point to the importance of understanding SET.

---

\(^1\) This is known by different names, including student feedback questionnaires, student rating of instructions (SRI), and many others. SET was the most commonly used in the literature we reviewed.
At one institution, a number of parties involved with a Master of Science in Accountancy (MSA) degree program, including instructors, the accounting department chair, and the program director, perceived a tendency for MSA class SET results to be at the low extreme. The MSA was unique in that it operated on the closed cohort basis. A closed cohort is a set of students entering a program and pursuing a degree together, and in which all classes are on a fixed schedule and limited to cohort members. This study sought to understand how students in the MSA engaged with SET. Specifically, it used the MSA as a qualitative case study to examine their interactions and understandings of the SET process. Interviews of MSA alumni provided a rich, descriptive account of the group processes. While the cohort setting provided an extreme example in which to study these processes, which may occur across other educational models.

**Contribution**

There exists an extensive body of literature on SET, mostly quantitative studies on the validity and reliability of SET instruments and possible biases. This study examined students’ processes with respect to SET, and by documenting these processes through interviews ‘allows the data to speak.’ One specific contribution is to provide a deeper understanding of how students’ thoughts and emotions, and group dynamics affect SET. This will help instructors and administrators better to interpret SET results and manage their individual classes and overall programs. Concerning certain findings, we suggest responses based on sound teaching theory. The study also provides a basis for more nuanced analysis of SET results by those making promotion and tenure decisions, in which positions and even careers may be at stake.
The first section that follows summarizes relevant literature and states the research questions. The second section describes the MSA case setting. The third section explains the method of analysis of SET and interview data. The fourth section presents the results and discussion of the interviews. The final section provides concluding comments.

LITERATURE AND BACKGROUND

This section includes literature on relevant social theory, the cohort model, SET, and group processes. It concludes with the research questions.

Theoretical Foundations

Framing the cohort social setting as an extreme example begins with Goffman’s (1961) ‘total institution.’ In a total institution, individuals lead formally administered lives, characterized by a physical barrier to social intercourse with the outside and symbolized by prisons and boarding schools. Some have contextualized US law schools (Granfield, 1992) and both graduate schools and business organizations (Schein, 1968) as being akin to Goffman’s total institution. While professional education does not occur with the physical separation implied by Goffman, it does require substantial commitment from its students. Coser (1974) modified Goffman’s concept into the ‘greedy institution,’ still demanding commitment, but relying on voluntary instead of physical means to separate individuals from others. Scott (2010) proposed the ‘reinventive institution’ as a subset of greedy institutions, one to which people voluntarily retreat for intense self-
reflection and enrichment. One type of Scott’s reinventive institution is the high commitment educational setting. While any educational setting demands commitment in terms of meetings and interactions, a closed cohort, segregated by exclusive classes and a fixed schedule, requires a higher level of commitment and fits the definition of greedy institution or reinventive institution.

Individuals in high commitment settings make primary adjustments, that is, they are socialized (Goffman, 1961). In educational settings, such primary adjustments relate to normal activities (e.g., attending class meetings, completing SET instruments). Individuals may make secondary adjustments, outside of the normal activities (Goffman, 1961; Ng & Höpfl, 2011). Secondary adjustments may be ‘disruptive’—purposely set out to alter organizational structure, or ‘contained’—deployed to create spaces in which individuals resist institutional forces (Goffman, 1961; McCorkel, 1998; Renshaw, 2006).

High commitment educational settings yield intense interactions (Coffey, 1993) and in cohorts, these may include intragroup conflict and politics (Sathe, 2009).

The Cohort Model

There are many ways to classify curricular models. One is by the method of interaction, for example, a classroom group experience, entirely on-line, or blended. The use of the instructional time is another, for example, tutorial, traditional lecture, or ‘flipped’ classrooms. Still another is whether students choose their own classes (‘self-paced’) or are subject to a fixed schedule.

A cohort is a group of students who begin and complete a set of classes together on a fixed schedule, engaging in a common learning experience leading to a degree. Cohorts usually consist of 10 to 25 students, typically pursuing a master’s degree over 12
to 18 months. They may be closed or open to students from outside the cohort and may be full-time or part-time (Barnett & Muse, 1993). While cohorts can also be on-line or blended, this study focused on cohorts with traditional classroom experiences.

Saltiel and Russo (2001) attributed the initial development of cohort programs to professions such as law, medicine, the military, and business. While they noted that graduate business programs (especially ‘executive’ MBA programs) use the cohort model, the business and accounting literature includes very little cohort model research: Reynolds and Sitharaman (2000) found greater affective learning in cohorts and Sathe (2009) found cohorts provided knowledge about social dimensions of the accounting profession.

The primary rationale for establishing cohorts is to form supportive learning communities that emphasize collaboration instead of competition (Barnett & Muse, 1993; Hesse & Mason, 2005). Social and cultural processes, shared experiences and interactions, collective efforts, and mutual commitment to an educational goal characterize cohorts (Pemberton & Akkary, 2010). Cohorts develop close bonds because of both intentional programmatic efforts and simply from the amount of time students spend together (Beck & Kosnik, 2001; Norris, Barnett, & Basom, 1996). Students interact in face-to-face gatherings between classes and by phone, text messaging, e-mail, or other social media. Compared to the self-paced model, the community created by a cohort is less confined to the time and space of the classroom and students cannot dissolve into anonymity after a class or term (Hesse & Mason, 2005; Mandzuk, Hasinoff, & Seifert, 2003). According to Saltiel and Russo (2001) cohorts: expect to collaborate, interact, and fully share resources, information, and emotional support with classmates;
value trust and mutual respect; are known for being ‘all for one and one for all’ (p. 108); and develop a shared memory because of the close personal interactions.

Cohort students and their instructors develop a greater sense of interpersonal connection than those in self-paced programs (Barnett et al., 2000; Pemberton & Akkary, 2010; Saltiel & Russo, 2001). Students are expected to interact more freely with instructors (Wesson et al., 1996) and the cohort model tends to flatten the hierarchy between instructors and students (Maher, 2001). This may lead to student-instructor interactions that have a positive effect on student performance (Jaasma & Koper, 2015; Pascarella & Terenzini, 2005). In addition, cohorts want instructors to build knowledge upon their shared coursework experiences (Saltiel & Russo, 2001). Ultimately, however, an expectations gap may develop. Teitel quoted a student, ‘Students have high expectations of faculty members,’ (1997, p. 76), while Barnett et al. quoted an instructor, ‘Cohort groups may form expectations which cannot be met’ (2000, p. 267).

Since cohort members experience each instructor’s teaching as a group, these experiences become part of the cohort’s shared memory. This shared memory coupled with a cohort’s cohesiveness results in a group of learners who develop a power not seen in other settings (Saltiel & Russo, 2001). When a number of students complain at one time, whether as a group or by appointing a spokesperson, administrators and instructors experience ‘cohort agency’ (Maher, 2001) and take it more seriously than if only one student had complained. Cohort agency results in the expression of concerns that might never be brought up in other models (Teitel, 1997).

Student Evaluation of Teaching (SET)
SET is widely studied: ‘Several thousands of research studies have appeared since the publication of the first report on SET by Remmers and Brandenburg in 1927’ (Spooren, Brockx, & Mortelmans, 2013, p. 600). Universities worldwide use an extensive array of SET instruments. Spooren, Brockx, and Mortelmans cite a number of well-designed and validated SET instruments, but note that others appear to have no clear theory of effective teaching. A typical SET instrument consists of a series of items about class content and teaching answered on a Likert scale (with at least one item pertaining to overall teaching effectiveness) and space for written comments. It is normally anonymous and administered near the end of the class without the instructor present (Kogan, Schoenfeld-Tacher, & Hellyer, 2010).

In addition to the literature on SET validity and design, other literature links SET to effective teaching (Hativa, 2013) and addresses how SET can provide guidance for university instructors (Downs-Lombardi, 1996; McKeachie, 1997; Nilson, 2010; Svinicki & McKeachie, 2014). Standifird, Pons, and Moshavi (2008) examined classroom influence tactics and found that ‘soft’ tactics (empathy) and ‘rational’ tactics (referent and expert power) had a significant positive effect on SET, while ‘strong’ tactics (reward, coercive, or legitimate power) had no effect.

In high-commitment settings, the intensity of interactions and the extension of temporal and spatial community may affect SET. After an episode with an instructor, students may leave class and continue to process the episode in a social setting or over social media, possibly amplifying the sentiments (Lawrence, 2002; Seifert & Mandzuk, 2006). Those communications may generally shape, or directly find their way into SET. Cohorts tend to be vocal about the quality of teaching and the relevance of class material
(Barnett & Muse, 1993). They give their feedback earnestly because they believe that it will be used to improve the class or the program (Maher, 2005; Saltiel & Russo, 2001). Finally, cohorts believe that they engage in the SET process with a high level of thoughtfulness and fairness (Schulte, 2002).

Cohort agency can result in a united front in favor of or against an instructor and may lead to stronger opinions of an instructor (Barnett et al., 2000). If against the instructor, students may blame, or ‘scapegoat,’ the instructor for group problems (Mandzuk, Hasinoff, & Seifert, 2003; Mather & Hanley, 1999; Radencich et al., 1998). Radencich et al. (1998) found that scapegoating seemed particularly endemic to classes with instructors new to a program, even if the instructors had otherwise proven their mettle, quoting one of their subjects ‘how much we know about the professor (determined our behavior)’ (p. 120). Mather and Hanley (1999) characterized this behavior as mob psychology.

**Group Processes**

Shaw (1981) defined a group as two or more persons interacting with one another in a manner that each person influences and is influenced by each other person. Since research on groups and group processes is widely varied, we have focused on specific dimensions. Nonetheless, the group processes that may bear on a high commitment setting, like a cohort, present an admixture that proves difficult to separate into identifiable elements. We present a set of group processes—conformity, peer pressure, deindividuation, group polarization—and their interrelationships in this subsection. One process, group formation, is somewhat a given in this study since the group (i.e., cohort)
is formed through enrollment in the program. Also, while ‘groupthink’ is commonly used to describe collective decision-making, SET is ultimately an individual decision. Thus, we have focused on processes that are more related to the individual’s interaction with the group.

Conformity. Conformity refers to the degree to which an individual’s behavior corresponds to the norms or attitudes of the group. Deutsch and Gerard (1955) distinguished two types of conformity motivations: informational and normative. Informational conformity is the desire to form an accurate interpretation of reality and behave correctly, while normative conformity has the goal of obtaining social approval from others. Conformity is related to group cohesiveness, the degree to which members are motivated to stay in the group. The shared motivation to complete a degree program enhances group cohesiveness. Groups that are more cohesive tend to exhibit greater conformity (Hogg et al., 2004). Greater conformity also occurs in groups with decentralized communications, that is, where all group members can exert direct pressure on each other group member (Shaw, 1981). The likelihood of perceiving unanimity is greater in groups with decentralized communication patterns. When the other members are in unanimous agreement, group members are more likely to conform (Asch, 1952).

Peer Pressure. The effects of peers on university students have been widely studied (Pascarella & Terenzini, 2005). Kuh (1995) identified peer groups as formal or informal groups that influence members’ interpretive frameworks. Astin wrote that peers were university students’ ‘single most potent source of influence’ (1993, p. 398). While peer influence research in the university setting has focused on topics ranging from alcohol use to academic persistence, Weidman’s (1989) framework shows that normative
pressures may be brought to bear on formal academic contexts, which could include SET. The cohort literature (Barnett et al., 2000; Maher, 2001, 2005; Siefert & Mandzuk, 2006) suggests that peer influence is particularly strong in cohorts. Mandzuk et al. (2003) concluded that peer pressure in cohorts challenges students’ sense of individuality.

**Deindividuation.** Classical deindividuation theory refers to the loss of one’s sense of individuality resulting in behavior with little or no reference to personal internal values or standards of conduct. In a state of deindividuation, anonymity reduces the possibility that individuals will face negative consequences for their behavior (Shaw, 1981). Recent interpretations (e.g., Reicher, Spears, & Postmes, 1995) have suggested that responsiveness to a group norm is not necessarily a mindless process reflecting a reduced sense of self, but may be a conscious process relating to a meaningful sense of identity. Rather than engaging in antinormative activities, deindividuation may lead individuals to conform to group norms (Cialdini & Goldstein, 2004).

**Group Polarization.** A series of studies, beginning with Stoner (1961), examined individual versus group decisions involving risk. These studies found that, when individuals who had made decisions involving risk (e.g., low-stakes bets, securities purchases) were formed into groups to make the same decision, the resulting group decision was riskier than the average of the individual decisions. This research was named the ‘risky shift’ literature. Further research showed that while initial relatively risky tendencies elicited shifts toward riskier positions, initial cautious tendencies elicited shifts toward more cautious positions. Thus, risky shift was re-hypothesized as group polarization: the average post-discussion group response will tend to move toward the extreme in the same direction as the average of pre-discussion individual responses.
Two primary explanatory theories developed. Social comparison theory suggests that, as discussion reveals the group’s norm, individual members change their opinions or decisions to be more in accord with other persons (Festinger, 1954; Brown, 1965). Persuasive arguments theory assumes that group members become exposed to novel and persuasive arguments during group discussion and modify their opinions accordingly (Vinokur & Burnstein, 1974). Group decision-making is not essential to group polarization; a discussion followed by individual decisions also produces a shift in the group average toward the extremes (Myers & Lamm, 1976).

In summary, the intensity of high commitment educational settings, the earnestness with which cohort students undertake SET, and the group processes literature suggest that the SET process in a cohort would be a rich area to study. ‘Cohorts seem to intensify and crystallize programmatic experiences’ (Tom, 1997, p. 153).

Research Questions

Our first research question arose from observations of individuals associated with the MSA. The program director, the accounting department chair, and others observed that SET scores in the MSA program tended toward extremes:

- To what extent do SET scores in the MSA tend toward extremes?

To the extent that a cohort conforms to the idea of a “group” in the literature above, it can be characterized by strong feelings of community, conformity, peer pressure, etc. The SET process may be influenced by these characteristics. We designed
our second and third research questions to gain a general understanding of the SET process and to identify group processes that may be present in the cohort:

- How did MSA students describe their engagement with the SET process?
- What group processes or other characteristics of the cohort might be associated with the SET process?

Pursuing these questions in a cohort setting provided an environment in which issues might become known clearly and could be better isolated and analyzed. While this study focused on a cohort as an extreme example, the issues it sought to understand may be present in many educational settings.

**RESEARCH SETTING**

**MSA Program and SET Scenario**

The MSA was offered by the college of business in a private US university. Admission required a baccalaureate degree in accountancy. Most students entered the same year as their baccalaureate graduation. Typically, about one-half of the students in each cohort had earned their baccalaureate degrees from this university.

A new cohort began on a full-time basis each July in which students undertook one class, professional development seminars, and interviews for an internship the following January through March. The fall term entailed four classes. After completing the internship, students returned for four classes, finishing the MSA in June. Throughout the year, the cohort engaged in co-curricular activities, including social events and
community service. The MSA started in 2005 and eight cohorts had finished at the outset of this study.

Since 2010, the university has used the IDEA Student Ratings of Instruction system\(^2\) for SET. Prior to adopting IDEA, it had used an internally developed system. While the wording in the two systems’ response items was not identical, both included an overall instructor item. The university considered them equivalent for longitudinal purposes such as tenure and promotion.

Figure 1 is a diagram of the overall SET scenario. It represents an example of three instructors’ teaching for one academic year, showing instructors who taught across programs (Master of Business Administration, MSA, and Bachelor’s). Each rectangular box within a program represents a ‘class,’ a designated meeting (days, time, and location) when students meet with an instructor for instruction in a particular subject. SET was administered in each class. An ‘instructor-year’ is the set of classes taught by one instructor during an academic year. For example, Instructor 1 taught three classes in fall and three classes in spring. Therefore, Instructor 1 had six SET scores during the instructor-year: one in the MSA and five in other programs (three Master of Business Administration and two Bachelor’s). During the academic year, the MSA cohort took nine classes; three from instructors included in this figure and six not shown.

Figure 1. Overall SET Scenario.

---

\(^2\) For further information, see the IDEA website, [http://www.theideacenter.org/](http://www.theideacenter.org/). IDEA is one of the instruments that Spooren, Brockx, and Mortelmans (2013) cited as well-designed and validated.
Instructors administered SET near the end of the term. Each student evaluated the instructor on a 1 to 5 Likert scale (1 being the most negative and 5 being the most positive) by darkening ‘bubbles’ on a paper instrument which also provided space for written comments. The system presumed the students completed the instruments individually. University protocol required the instructor to read instructions to the assembled class and designate a student to collect the completed instruments, seal them in an envelope, and place the envelope in a campus mailbox. The instructor was required to leave the room while students completed the instruments. The completed instruments
went to the university’s institutional research department for processing. Once an
instructor submitted grades, the instructor received the completed instruments and a
summary analysis.

METHOD AND ANALYSIS

SET Data and Analysis

The first research question asked to what extent SET scores in the MSA tend
toward extremes. To answer this question, we collected SET data for all classes (across
all programs) for the 17 instructors who had taught the five most recent MSA cohorts.
Each cohort included nine classes, for 45 total MSA classes. SET data for four classes
were unavailable, resulting in 41 useable classes. Table 1 summarizes the class data.

Table 1. Class Data.

<table>
<thead>
<tr>
<th>MSA classes (5 cohorts x 9 classes)</th>
<th>MSA</th>
<th>All Programs*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data unavailable**</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Total classes taught by instructors teaching in MSA (SET score descriptive statistics)</td>
<td>41 (Mean = 4.08, SD = 0.73)</td>
<td>197 (Mean = 4.37, SD = 0.53)</td>
</tr>
<tr>
<td>Instructor taught one class during year</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Classes for Instructor-year Analysis</td>
<td>39</td>
<td>191</td>
</tr>
</tbody>
</table>

*Includes MSA and all other programs.
**See Appendix A for further details.

While the IDEA instrument includes many criteria, Benton and Cashin (2012) and
Marsh (1984) concluded that it is the instructor item in SET instruments that primarily
determines students’ opinions of their educational experience. We used the mean score of
the ‘instructor’ item as the SET data in two ways: Instructor-year Analysis and Outlier
Analysis.

**Instructor-year Analysis.** This analysis determined the number of instances the
MSA class SET score was the lowest (excluding ties) of the SET scores in an instructor-
year. This compared an instructor’s SET result in their MSA class to the remainder of
their classes during an academic year. The instructors in this study taught an average of
4.80 classes (across all programs) per instructor-year. No instructor taught more than one
MSA class in an instructor-year.

We delimited two MSA classes and corresponding instructor-years because the
instructor taught only that MSA class in the instructor-year. In the remaining 39
instructor-years, the score for the MSA class was the low extreme in 16 of the 39
instances (41.02%). During the study period, the 39 MSA classes comprised only 20.41%
of the 191 total classes taught by the instructors. This suggested that MSA classes more
often result in SET scores at the low extreme for an instructor-year.

**Outlier Analysis.** We determined the overall mean and standard deviation for all
classes (MSA and others) taught by the instructors involved in the MSA. This analysis
compared MSA classes to all classes without regard to individual instructor. SET data
were available for 197 total classes during the study period, with an overall mean
instructor score of 4.37 and standard deviation of 0.53. A simple analysis to identify
individual classes lying outside of two standard deviations from the overall mean resulted
in defining those below 3.31 as outliers. Eleven of the 197 classes were below 3.31, or
outliers. Six of these 11 outliers (54.55%) were MSA classes. Given that only 41 of the
197 (20.81%) were MSA classes, this suggested that MSA classes more frequently result in outlier instructor evaluations.\textsuperscript{3}

We compared the proportion of outliers in non-MSA classes (5 in 156 or 3.21%) to the proportion of outliers in MSA classes (6 in 41 or 14.63%) using a two-sample one-tailed Z-test. Consistent with the intuitions that led us to ask research question one, the proportion of (outliers) in MSA classes is significantly higher than the proportion in non-MSA classes ($z = 2.84$, $p = .002$).

**Student Engagement with SET**

To address the second and third research questions we used interviews and interpretational analysis techniques to find the themes and patterns that describe the students’ experiences. The primary goal was to understand how the participants’ sense-making in the SET process shaped their actions. Research aimed at understanding human actions in their social contexts in terms of the meanings that these actions hold normally follows a qualitative approach (Atkinson & Pugsley, 2005; Berg, 2004). This allows researchers to draw findings from the significant themes gleaned from the raw data (Thomas, 2003). The aim of the process is to organize and structure the findings according to the issues and topics identified by the participants as being important to understanding the phenomena of interest (Shaw, 1999). Attempting to understand how students engage with SET requires a method that can describe their practices and processes, and explore the effects of those practices and processes on outcomes. This

\textsuperscript{3} The 41 classes for which data was available and 197 total classes apply in this analysis. The delimitation of the classes in the Instructor-year Analysis in which the instructor taught only one class does not apply in the aggregated data used here. Additional details regarding the instructors and SET data are included in Appendix A.
research used a case study paradigm and unstructured interviews of members of multiple cohorts from the MSA. The case study approach is especially appropriate when time and a particular structure bound the inquiry (Stake, 1995). The specifics of the MSA (e.g., classes and policies) and those of the particular cohorts (e.g., individual students and instructors) served as boundaries for this study. Guided by the research questions, this study explored the students’ group processes regarding their engagement with SET in the specific context of one program.

**Data Collection**

Exploratory research conducted within a qualitative paradigm must be flexible enough to permit the researchers to uncover and investigate issues that emerge as interesting and potentially helpful in understanding the research questions (Shaw, 1999). We used group interviews in order to encourage open, expansive discussion. We conducted four group interviews with two to five participants from multiple cohorts in each. The interviews were conducted face-to-face by one of the researchers in an unstructured manner, a method appropriate to allow participants the freedom to explain their thoughts and highlight particular issues (Bogdan & Biklen, 2007) and to explore emerging lines of enquiry (Scapens, 2008). We used open-ended questions to prompt discussion and elicit understandings of the SET process. The questions used to guide the interviews are included in Appendix B. Interviews lasted approximately 60 minutes and were audio recorded. Within two days after the interviews, we e-mailed each participant and asked if they had afterthoughts, additions, or changes to their interview comments. The interview recordings were professionally transcribed.
Participants

The MSA program director identified alumni for the researchers to invite to participate in group interviews. This identification was purposive, with the goals of selecting alumni more likely to share their experiences freely and seeking representation across cohorts. More specifically, we used heterogeneous sampling in which participants were chosen to give a detailed picture of a particular phenomenon (Patton, 2002). We excluded students enrolled concurrent to the study due to conflict of interest, and alumni from the earliest three cohorts due to the lapse of time since their experience. Initially, we invited 25 alumni (five members from each of the five most recent cohorts at the time of the study). When an invitee declined, we invited another member of the same cohort. Interviews began once we achieved a critical mass in order to begin scheduling interview sessions. Once we reached data saturation and all cohorts were represented, invitations ceased. Ultimately, 45 alumni were invited and 15 agreed to participate. The five cohorts were not represented equally, however at least one participant represented each. While such a method is not objective, the participants represented the overall MSA makeup. Additional details regarding the participants are included in Appendix C.

Ethical Considerations

One of the researchers taught one MSA class each year. Since the participants were not current students, there was no conflict of interest specific to the dual roles of instructor and researcher. The reader must consider that this researcher navigated the roles of researcher and former instructor. We obtained permission to conduct the study
from the MSA program director, accounting department chair, dean of the college of business, and the university’s Institutional Review Board (IRB). Pursuant to IRB permission, each participant signed an informed consent form that addressed data confidentiality and risk. In addition, the university considered SET data confidential. All instructors consented to provide their SET results after reviewing the IRB-approved description of the research.

**Interview Data Analysis**

Data consisted primarily of the group interview transcripts. Data analysis followed a general inductive approach (Thomas, 2006). This approach incorporates common strategies used in several qualitative analysis traditions, including grounded theory, phenomenology, narrative analysis, and discourse analysis (Black *et al.*, 2010). In a general inductive approach, the purpose is to create a small number of summary categories that capture the key themes in the raw data judged the most important in terms of the research objectives. Theory building is limited to the presentation and description of the most important categories, typically between three and eight themes given the evaluation objectives. The outcome is the development of a framework that summarizes the raw data and conveys key themes and processes (Thomas, 2006). It seeks the core meanings evident in the data, relevant to the research questions.

Data analysis began with the researcher who conducted the interviews closely reading the interview transcripts while listening to the audio recordings to ensure that the transcripts accurately reflected the interviews, followed by coding the raw data transcripts. The second researcher was provided the categories and their descriptions,
closely read the transcripts while listening to the recordings, and separately coded without reference to the first researcher’s coding. A process of axial coding in which the categories were refined into themes followed the initial open coding (Strauss & Corbin, 1990).

During data analysis, 202 comments received a primary code from both researchers. Of these 202, both researchers gave 78 comments a secondary code and 20 comments a tertiary code. In order to quantify the degree of agreement between the two researchers, we computed Percentage of Agreement by dividing the number of comments receiving the same code by the number of comments coded by both researchers. The average Percentage of Agreement across the three code levels is 83.82%. See Table 2.

Table 2. Percentage of Agreement.

<table>
<thead>
<tr>
<th></th>
<th>Coded by both (1)</th>
<th>Codes agreed (2)</th>
<th>Percentage of Agreement (3)</th>
<th>Coded by first researcher only (4)</th>
<th>Coded by second researcher only (5)</th>
<th>Coded by neither (6)</th>
<th>Total 1,4,5,6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>202</td>
<td>178</td>
<td>88.12%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>202</td>
</tr>
<tr>
<td>Secondary</td>
<td>78</td>
<td>65</td>
<td>83.33%</td>
<td>33</td>
<td>10</td>
<td>81</td>
<td>202</td>
</tr>
<tr>
<td>Tertiary</td>
<td>20</td>
<td>16</td>
<td>80.00%</td>
<td>17</td>
<td>2</td>
<td>163</td>
<td>202</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td>83.82%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percentage of Agreement is a commonly reported measure included here to aid in comparison with other research and for its intuitive appeal (Hallgren, 2012). However, Percentage of Agreement does not correct for the possibility that two coders will agree by chance (Gwet, 2014). To correct for chance agreement, we calculated Cohen’s (1960) kappa as a statistic of inter-rater reliability for two raters (Haley & Osberg, 1989). Table
3 reports Cohen’s kappa for each of the three coding levels (primary, secondary, and tertiary) as well as for the average across codes.

Table 3. Cohen’s Kappa (k).

<table>
<thead>
<tr>
<th></th>
<th>Cohen’s Kappa (k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>86.53%</td>
</tr>
<tr>
<td>Secondary</td>
<td>71.55%</td>
</tr>
<tr>
<td>Tertiary</td>
<td>88.60%</td>
</tr>
<tr>
<td>Average</td>
<td>82.23%</td>
</tr>
</tbody>
</table>

Cohen’s Kappa values between 40% and 75% are ‘fair to good,’ and above 75% are ‘excellent,’ (Landis & Koch, 1977).

While the interviews were the primary source of data, the researcher who had taught in the MSA had observed firsthand, or had otherwise heard about, certain experiences described in the interviews. Discussions with the program director also confirmed particular incidents. Finally, the SET data used in the Instructor-year and Outlier Analyses corroborated certain interview phenomena. These secondary data sources provided a degree of triangulation (Marginson, 2004).

RESULTS AND DISCUSSION

Data coding initially resulted in many categories. Iterative reviews reduced the number of categories to 17. Certain categories (e.g., Cohort) had been anticipated while others (e.g., Sharing) arose from the interviews. Table 4 presents the categories and the number of times they appeared in the transcripts. While the number of elements coded to
a category provided one indication of a category’s importance, the qualitative nature of the elements (e.g., emphasis, humor, passion) provided another. Through analysis and discussion, we grouped the 17 categories into five themes as shown in Table 4. The themes showed a progression: The participants experienced the MSA in common. The common experience generated communication about various aspects of the MSA and specific classes; a common discussion topic was instructor performance. This communication about instructors, in turn, affected SET. After developing the SET theme, we considered the group processes arising from the data that could affect SET. We also developed a group of categories specifically related to student-instructor relationships.

Our second research question asked how students would describe their engagement with the SET process. Our third research question relates to possible group processes and other characteristics that might explain the observed extreme SET scores. We believe the themes arising from the inductive interview analysis in general, and the SET and Group Process themes specifically, allow us to address these questions.

Table 4. Categories and Themes.

<table>
<thead>
<tr>
<th>Categories</th>
<th>n</th>
<th>Themes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cohort</em>—cohort issues: community, relationships</td>
<td>108</td>
<td>Common Experience</td>
</tr>
<tr>
<td><em>Graduate</em>—general graduate school issues</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td><em>Team</em>—team assignments</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td><em>Common</em>—common, shared experiences</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td><em>Discussion</em>—discussion, all topics</td>
<td>110</td>
<td>Communication</td>
</tr>
<tr>
<td><em>Communication</em>—multimedia communication</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td><em>Evaluations</em>—discussion of SET</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td><em>Sharing</em>—sharing instructor opinion across cohorts</td>
<td>35</td>
<td>SET</td>
</tr>
<tr>
<td><em>Grading</em>—comments about grading</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td><em>Benchmarking</em>—comparing instructors</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td><em>Undergrad</em>—undergraduate SET experience</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>
Common Experience

Participants discussed their common experience in the MSA, especially related to the team-oriented class projects. ‘Teams’ were groups of 4-5 students, formally assigned at the start of each cohort by the MSA program director to mirror an accounting workplace team. The MSA used teams extensively across the curriculum and they proved an early step in community building. The team experiences, while common in the sense that they were a locus of interaction, also led to shared memory (Saltiel & Russo, 2001).

With the project-based teams in classes, I mean, you’re constantly talking with your team and kind of have the same views as your teammates because you’re getting the same experience as they are. (Student 2/Cohort C/Interview I)^4

While students initially experienced the MSA within their teams, once they came to know other members of the cohort better over time, their shared experience expanded to an identification with all members.

---

^4 We identified participants by ‘Student #,’ their respective cohorts by letter, and interview sessions by Roman numeral. Readers can trace comments by participant, cohort, or interview, and can see, for example, that members of the same cohort generated similar comments independently.
The students’ developing relationships led to feelings of closeness, even family (Saltiel & Russo, 2001). As their shared experiences expanded into broader relationships, with instructors for example, these feelings shaped the relationships.

In my cohort, everybody had each other’s back. … If the professor did something to one person, we’re going to support this broken cohort member. (Student 11/Cohort C/Interview III)

This comment exemplified the potential closeness of relationships and reminds instructors to be aware of their own behaviors that may alienate a student (Nilson, 2010). In certain settings, like cohorts, alienating one will likely alienate many. Instructors should consider an intervention (e.g., an individual conference) with students who may have become alienated or are having academic trouble (Downs-Lombardi, 1996).

**Communication**

The team projects, coupled with co-curricular activities and other common experiences stimulated communication. The MSA had a ‘commons,’ a student-only lounge, which was fertile ground for discussion.

The commons is a pretty safe environment because everybody knows they can speak their mind and hopefully it won’t get out. (Student 15/Cohort E/Interview IV)
The literature (Lawrence, 2002; Saltiel & Russo, 2001) had noted the temporal and spatial extension of community. ‘You talk to people in your [team] and then someone talks to someone else,’ commented Student 2, while mimicking keyboarding. While the shared experience and the existence of the commons facilitated communication, modern technology expanded its bounds, a condition that applies to many educational settings.

Communication included many topics, but all 15 participants agreed that discussion of instructors was significant, for example:

When we were frustrated, we’d talk about it and talk about what the teachers were doing that made us so frustrated. (Student 7/Cohort A/Interview II)

Their shared experience and easily facilitated communication led to on-going discussion of instructors, a fact of which instructors must be mindful (Saltiel & Russo, 2001).

**Student Evaluation of Teaching (SET)**

The following discussion describes how students engage with the SET process and addresses our second research question. The students expressed the belief that cohorts take the SET process seriously and that administrators will use their feedback to improve the overall experience (Maher, 2005; Saltiel & Russo, 2001; Schulte, 2002), for example:
We always tried to improve the class or the professor’s teaching style so that the cohorts in the future might catch a break. (Student 15/Cohort E/Interview IV)

Students’ expectation of influence may play a part in SET. The literature (Maher, 2005; Saltiel & Russo, 2001) suggested that cohort students provide feedback with a sense of earnestness, which may add to the overall intensity suggested by Tom (1997). In addition, Standifird, Pons, and Moshavi (2008) found a link between perception of influence and satisfaction with the instructor. Instructors may wish to foster the sense that SET feedback will be used to improve the class rather than sending the message that SET is an administrative exercise.

The students’ common experience and communication appeared to result in a general influence on SET, for example:

So, in the end it’s like you kind of come to a consensus that this is … our opinion of the professor, we’ve all talked about it and we’ve all shared experiences. So, it’s like a group effort. (Student 9/Cohort C/Interview II)

Not only did this comment point to the influence of communication on individual members, but also it explicitly pointed to the effect on SET of their shared experience (Saltiel & Russo, 2001).

How the students’ discussions about instructors ultimately affected the way they engaged with SET differed by class and by cohort. Most often, the administration of the
SET instruments went as designed, individually and without discussion. However, exceptions occurred, suggesting Goffman’s (1961) contained secondary adjustments, acts of resistance during the SET administration process.

The students’ familiarity with each other’s opinions of instructors led some to try to either counter, or build momentum behind, their perception of another’s opinion. A number of participants, from separate cohorts, mentioned this phenomenon, described by one as follows:

If someone was … writing a lot and you knew if that person liked the professor or not, because we had talked … I’d feel more inclined to write some notes either countering what they said or supporting what they said.

(Student 2/Cohort C/Interview I)

While discussions about opinions of instructors prior to SET administration were prevalent, in two instances the students discussed their sentiments during completion of the SET instruments. As a member of one cohort described it:

People would shout things out, like, ‘Remember this on this day and what the professor did here?’ So that might sway [people]. But people would shout out a couple of things, for like 30 seconds and then everyone would settle down and fill in the bubbles. … They’re probably hoping maybe the person will bring it down a point or two. (Student 15/Cohort E/Interview IV)
Given an overall mean of 4.37 and standard deviation of 0.53 (Table 1) if a number of students were to ‘bring it down a point or two’ it may have a significant effect on an instructor’s overall results. The second cohort was more serious in its discussions.

I know when our cohort did our evaluations we would talk while we were doing it, yeah, it would be like, ‘I’m giving this teacher a 1 right now.’ And, you’d have that open communication while you’re filling it out, so you know exactly what everyone else is giving them. (Student 7/Cohort A/Interview II)

We definitely had discussions about what we were writing, or what someone was going to write up beforehand to print and put in the evaluations. (Student 4/Cohort A/Interview I)

Yeah, the [Instructor W] debacle ended up that when it came to be about time for evaluations I decided I can’t even write as much as I want to write on the back of this little form, so I typed it up beforehand. [I asked] ‘Well, can I read it?’ Everybody wanted me to read it, and I read it. I can’t tell you whether that influenced what [another] person thought or said. It probably had the potential to. (Student 6/Cohort A/Interview II)
Cohort A and Instructor W had clashed over a range of issues. The MSA director, the accounting department chair, and the researcher who was an instructor in the MSA had heard about these clashes, and how Student 6 had read comments during SET administration. The class was Instructor W’s first experience in the MSA despite many years teaching at the university. Instructor W’s evaluation for this class was the low extreme in the instructor-year. It was also an outlier. While the discussion of instructors during the term and up to the completion of the SET instruments likely had general influence on the evaluations, the discussion at the very moment of completing the instrument may have had a direct effect.

An instructor experiencing open conflict with an entire class should consider an intervention to attempt to work through the issue (Downs-Lombardi, 1996). In this case, it may have been a conference with influential individuals or the entire class, and would help foster the perception of influence (Standifird, Pons, & Moshavi, 2008).

To discourage discussion during SET administration an instructor may have it proctored. This avoids intimidation through the instructor’s presence and ensures completion on an individual basis, a typical assumption of SET systems. As more institutions move to on-line administration, there is less likelihood of whole-class discussion. However, SET designers and users must consider that the opportunity for small-group discussion increases as students may opt to complete their SET together.

**Group Processes**

The discussion in this section as well as the following section on Student-Instructor Relationships provides guidance on answering our third research question. The
literature posited that processes endemic to groups might affect activities like SET. Each of the 15 participants commented that the discussions during the term affected their opinions of instructors. Certain interview comments focused on routine, term-long discussion.

Some people were very critical of the professors, overly critical I think, but those are also the people that were most outspoken so it kind of snowballed in a way after a test or something. Most of us would be talking and then other people would join the conversation and then eventually I think most were kind of like on the same page in a sense. (Student 12/Cohort A/Interview III)

The outspokenness of certain group members may have exerted peer pressure that led to conformity.

Some comments pointed to the group specifically influencing a participant’s rating on the SET instrument, for example:

I can think of an occasion where maybe the group consensus was that this professor was good, but for whatever reason I didn’t like them. Come evaluation time, I might say, ‘Well I’m just being mean and they’re probably not as bad as I’m making them out to be so I think I’ll be a little bit nicer because I know everyone else likes them.’ So, it might not reflect
my true feelings because of the group’s opinion of this person, but I might
try to be more in line with them. (Student 8/Cohort C/Interview II)

‘This might not reflect my true feelings’ suggests deindividuation. This change in a
student’s evaluation, resulting from peer pressure, would move the SET score further in
the direction of the group’s more positive overall assessment as compared to this
student’s more negative assessment. It provides some evidence of polarization. This
student further articulated the idea that the knowledge of others’ opinions provided
confirmation of, or influenced, a position.

You talk about it beforehand [and someone says] ‘I’m going to rate them
really bad.’ And I think, if they’re going to rate them really bad, I can rate
them really bad, versus if I didn’t know what anybody else is doing, I
might [have been] a little nicer. (Student 8/Cohort C/Interview II)

During the group interview session, the quote immediately above generated comments
about strength in numbers influencing the evaluations. Regarding Student 6’s reading of
prepared comments described in the SET subsection, another participant from that cohort
noted:

Well, everyone knew that at least one person was going to say these
things, so everyone else can feel comfortable saying these things too,
because it’s not just going to be you alone saying it. (Student 7/Cohort A/Interview II)

This sense of understanding the group consensus appears to lead to a tendency to conform.

While the literature (e.g., Radencich et al., 1998; Saltiel & Russo, 2001) cited the use of collective power in overt situations, for example, complaining to instructors and administrators, this study showed that collective power also could influence private processes, like SET.

**Student-Instructor Relationships**

Similar to findings in the literature (Barnett et al., 2000; Saltiel & Russo, 2001) participants frequently commented that they expected to develop, or did actually develop, closer relationships with instructors as compared to other settings.

In the cohort, you do get to know your professors, sometimes their family, and things that are going on in their life—a little more personal. (Student 10/Cohort C/Interview II)

When discussing SET, students pointed out these closer relationships, for example:

I would care more if the professor cared about me and I would write down specific examples. (Student 2/Cohort C/Interview I)
The expectation of close student-instructor relationships, and a mismatching of the actual relationships to the expected, played out vividly in one episode, known to the MSA director and the researcher who was an instructor in the MSA. Two instructors, married but not sharing the same surname, were new to the university. The first quote in the series below begins one participant’s unprompted narrative of the cohort’s discovery that the instructors were married. The remainder of the series is the subsequent discussion during the group interview. It shows Cohort A’s group process and the reactions of participants from another cohort. The latter show, first, surprise at the ends to which Cohort A went, and finally, sympathetic understanding.

[Instructor X] and [Instructor Y], you know, that’s two professors that came to [the university] and they did not tell anybody in the cohort that they were married. Somehow, you know, wheels started turning, people are talking, ‘Oh, they both came from [their prior university], that’s strange, and they both say they live in [nearby town], well that’s strange.’ Somebody all of a sudden says, ‘They must be married,’ and then somebody ‘Googled’ it and found, like, their on-line Christmas card.

(Student 6/Cohort A/Interview II)

This is intense! (Student 10/Cohort C/Interview II)
Right! The cohort was like, well, ‘Why didn’t they tell us they were married, what are they trying to hide?’ And you know all of a sudden you have your angry mob. (Student 6/Cohort A/Interview II)

It was just more weird that they wouldn’t tell us … they weren’t going to share anything personal. (Student 7/Cohort A/Interview II)

Yeah. I think a sense of betrayal or something. (Student 8/Cohort C/Interview II)

Instructors must make their own decisions on how much personal information to share with students. Some will be open about spouses, children, etc., subscribing to Downs-Lombardi’s guidance that ‘the secret to great teaching is to treat people as if they are in your own living room’ (1996, p. 63). While that may be too much sharing for some, Hativa describes one of the most important factors students consider when rating instructors in SET as ‘rapport with students: care, respect, patience, cooperation, approachability’ (2013, p. 37).

While the instructors’ marital status was not germane to the classes, the students’ comments reveal that they took exception to the omission. Both instructors’ evaluations were their low extreme in the instructor-year and both were outliers. This episode recalls Radencich et al.’s (1998) subject’s comment that student knowledge of an instructor determines their behavior. One interpretation of this data is that the students perceived the instructors as deviating from the group’s expected level of rapport, and they responded to
this perceived deviation through SET. In addition, Student 6’s ‘angry mob’ comment evokes Mather and Hanley’s (1999) characterization of this behavior as mob psychology and suggests deindividuation.

A different interpretation, offered by the accounting department chair during Cohort A’s program year, and possibly supported by the data, is that the students appeared to take out general program experience frustrations on their instructors. Student 7’s comment in the Communication subsection had specifically mentioned that Cohort A had frustrations, and discussed them. The prepared notes episode involving Instructor W (the result of on-going friction) occurred in Cohort A. One class’s schedule was chaotic as the result of an instructor’s hip replacement surgery. Another significant source of frustration was that Instructor V, a beloved, highly respected instructor, known to the students who completed their baccalaureate degrees at the university, was diagnosed with cancer prior to their start date. He was unable to teach their cohort and passed away during their fall term. In addition to being saddened by his death, the students who knew him from their undergraduate experience were disappointed that they would not have him for the two classes he was scheduled to teach. Those who did not know him had been told of his high regard and were likewise disappointed. One of the replacement instructors was Instructor W.

We were supposed to have [Instructor V]. He passed away, and yeah, we ended up with [Instructor W]. (Student 6/Cohort A/Interview II)
The other replacement instructor for Instructor V was Instructor Z. As with Instructors X and Y, the students’ had expected a rapport that did not develop.

That connection was just never established and at least we felt like she wasn’t trying to establish that with us. (Student 12/Cohort A/Interview IV)

An additional aspect of the interpretation that Cohort A took out its frustrations on instructors W, X, Y, and Z is the fact that they were new. Instructor W was new to the MSA, but not to the university. Instructor’s X and Y were new to the university (and the MSA) but had substantial prior experience in the topical areas of their Cohort A classes and in teaching graduate students. Instructor Z was an accounting professional with teaching experience, but new to the university (and the MSA). No other Cohort A instructors were new to the MSA.

Their newness may have been reason for Cohort A to target them. As noted, the Cohort A SET scores were the low extremes for Instructors W, X, and Y and were outliers. Of the 197 classes in the Outlier Analysis, SET scores from Cohort A for Instructors W, X, Y, and Z ranked 197, 191, 196, and 195, respectively—four of the seven lowest. The mean of these four instructors’ evaluations was 2.48, compared to an overall mean of 4.37, and 4.58 for the other five Cohort A instructors.

Table 5 presents Instructor W, X, Y, and Z’s SET scores for their Cohort A classes and their mean SET scores (and standard deviation) for all other classes taught during the years of the study. Their mean SET score in other classes (4.22) is near the

---

5 Instructor Z’s evaluation was delimited from the Instructor-year Analysis since it was Instructor Z’s only class.
overall mean SET score for all instructors (4.37 from Table 1). In each instance, the Cohort A SET Score is below the Other Classes Mean SET Score by more than four times the standard deviation for the other classes. We also compared their MSA SET scores to their mean and low SET scores for their non-MSA classes during their first year. Table 5 shows that Instructor W, X, and Y’s mean, and even their low SET scores in the non-MSA classes in the Cohort A year were much closer to their means for All Study Years, Other Classes.
Table 5. Instructors W, X, Y, and Z: Cohort A SET Compared to Other Classes

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Cohort A SET Score</th>
<th>All Study Years Other Classes SET</th>
<th>Cohort A Year, Other Classes SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>1.50</td>
<td>Mean 4.16 SD 0.55 Mean 4.16 Low 3.30</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>3.10</td>
<td>Mean 4.47 SD 0.25 Mean 4.38 Low 4.10</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>2.70</td>
<td>Mean 4.05 SD 0.33 Mean 4.20 Low 3.80</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>2.60</td>
<td>See footnote 5</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.48</td>
<td>4.22</td>
<td></td>
</tr>
</tbody>
</table>

The Table 5 SET scores for Other Classes provided no *ex ante* reason to expect such low SET scores in Cohort A. We interpret this difference between the MSA and non-MSA SET scores as further evidence that Cohort A may have taken out its frustrations on these new instructors. Radencich et al. (1998) had found students to be particularly hard on new professors in a program even if they were otherwise proven. This may also suggest a contained secondary adjustment (Goffman, 1961).

It appears that student expectations of instructors and group processes bear on the student-instructor relationship. These may be present in any educational setting. Instructors should understand that students might expect a higher degree of relationship even if instructors find it unnecessary and that processes such as deindividuation and peer pressure may find their way into SET results if those expectations are not met.

**CONCLUDING COMMENTS**

This study used a cohort model program as an extreme example in order to study how group processes affect SET. As with any case study, this study’s focus on one setting limits
its generalizability. As a high commitment setting, the cohort model provided a view of group processes played out intensely. These processes may be present to a greater or lesser extent in any educational setting. Readers must assess the study’s degree of correspondence to other educational settings.

This study provides an understanding of the manner in which group processes affect students’ approach to SET. The interview results show how common experiences may build a communication environment that leads to widespread discussion of instructor performance. This discussion, in turn, effects SET both through the influence of term-long discussions of instructors and sometimes through specific behaviors during administration of the SET instrument. This study contributes to the literature by examining group processes in SET.

The results show how group processes may affect SET outcomes. The results also suggest that students may expect a level of rapport with instructors and react through SET if the expectations are not met.

SET has evolved from its original intended use of providing instructors formative feedback to providing summative data for the evaluation of teaching and for critical decisions on tenure and promotion. We have included recommendations specific to particular study findings (individual or group conferences, developing the sense that SET has influence, proctoring SET administration) in the Results and Discussion section.

Other recommendations are conceptual. The study makes instructors more aware that students perceive them, and instructors can shape that perception. It shows that student interaction and communication can amplify the mundane (e.g., non-disclosure of marital status). While improvement in SET is associated with better student-instructor interactions (Benton & Cashin, 2012), better informal student-instructor interaction is important to learning, persistence, and other student success dimensions. Instructor
classroom behavior signals openness to such interaction (Jaasma & Koper, 2015; Pascarella & Terenzini, 2005). A better understanding of how students undertake the SET process is important concerning the three uses that have evolved for SET. First, instructors will find that a better understanding of SET will lead to better use of information from its feedback dimension. Resources like McKeachie (Svinicki & McKeachie, 2014) and Nilson (2010) provide specifics instructors that can tailor to their needs. Second, those who use it to evaluate instructors can better analyze SET data to make more responsible decisions regarding tenure and promotion. McKeachie’s suggestion to ‘sensibly use broad categories rather than attempting to interpret decimal-point differences’ (1997, p. 1219) is particularly germane when institutions offer a variety of class formats, including cohorts. Finally, institutional decision makers can understand SET information better prior to making decisions on classes or programs.

REFERENCES


Maher, M. A. (2005) The evolving meaning and influence of cohort membership, 


**APPENDIX A**

This Appendix provides selected details about the instructors. We note in Table 1 (Class Data) in the body of the paper that data was unavailable for four classes. This included one class for Instructor T who was new to the MSA and taught only one class, one class
for Instructor U, and two classes for Instructor V, the latter two frequently taught in the MSA. Instructors T and U were women and Instructor V was a man. Thus, the missing data was gender-balanced.

Table A1 shows whether the instructor was teaching in the MSA for the first time during the years covered in the study (New) and gender. It also includes columns regarding the Outlier Analysis. The columns show the instructor’s number of outliers in the MSA program and other programs. If an outlier occurred in the year in which the instructor was new to the program in which the outlier occurred, we designated it ‘New.’
Table A1. Instructor Demographic Data.

<table>
<thead>
<tr>
<th>Instructor</th>
<th>New</th>
<th>Gender</th>
<th>Outlier</th>
<th>MSA</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>No</td>
<td>M</td>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>J</td>
<td>Yes</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>Yes</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>Yes</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Yes</td>
<td>M</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>O</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>No</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>Yes</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>Yes</td>
<td>F</td>
<td></td>
<td>Unavailable</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>Yes</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>No</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>No</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>Yes</td>
<td>M</td>
<td></td>
<td>1, New</td>
<td>1</td>
</tr>
<tr>
<td>X</td>
<td>Yes</td>
<td>M</td>
<td></td>
<td>1, New</td>
<td>-</td>
</tr>
<tr>
<td>Y</td>
<td>Yes</td>
<td>F</td>
<td></td>
<td>2, New</td>
<td>-</td>
</tr>
<tr>
<td>Z</td>
<td>Yes</td>
<td>F</td>
<td></td>
<td>1, New</td>
<td>-</td>
</tr>
</tbody>
</table>

Table A2 shows the association of outliers with the instructor’s status as new to the MSA.

The table shows a higher tendency for new instructors to have outlier classes.

Table A2. Outliers and New Instructors

<table>
<thead>
<tr>
<th>New</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Total classes</td>
<td>10</td>
</tr>
<tr>
<td>Outliers</td>
<td>4</td>
</tr>
</tbody>
</table>

Table A3 shows the association of outliers with instructor gender. The table shows little, if any, association between gender and outliers.
Table A3. Outliers and Instructor Gender

<table>
<thead>
<tr>
<th>Gender</th>
<th>F</th>
<th>M</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total classes</td>
<td>13</td>
<td>28</td>
<td>41</td>
</tr>
<tr>
<td>Outliers</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>23.08%</td>
<td>10.71%</td>
<td>14.63%</td>
<td></td>
</tr>
</tbody>
</table>

Table A4 shows the following SET data for each instructor for (1) the MSA, and (2) classes taught in all other programs during the years of the study: the number of classes; minimum, mean, and maximum SET score; and the difference between the mean values for the MSA and all other programs.

Table A4. Instructor SET Data

<table>
<thead>
<tr>
<th>Instructor</th>
<th>MSA n</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
<th>Other n</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
<th>Difference Mean</th>
<th>Difference %</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>5</td>
<td>3.00</td>
<td>3.67</td>
<td>4.20</td>
<td>23</td>
<td>3.00</td>
<td>3.86</td>
<td>4.60</td>
<td>-0.19</td>
<td>-4.95%</td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td>3.78</td>
<td>3.78</td>
<td>3.78</td>
<td>2</td>
<td>4.57</td>
<td>4.58</td>
<td>4.58</td>
<td>-0.80</td>
<td>-17.38%</td>
</tr>
<tr>
<td>K</td>
<td>1</td>
<td>3.90</td>
<td>3.90</td>
<td>3.90</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>2</td>
<td>4.80</td>
<td>4.86</td>
<td>4.91</td>
<td>17</td>
<td>4.70</td>
<td>4.84</td>
<td>5.00</td>
<td>0.02</td>
<td>0.41%</td>
</tr>
<tr>
<td>M</td>
<td>2</td>
<td>4.25</td>
<td>4.33</td>
<td>4.40</td>
<td>8</td>
<td>4.40</td>
<td>4.62</td>
<td>4.90</td>
<td>-0.29</td>
<td>-6.33%</td>
</tr>
<tr>
<td>N</td>
<td>2</td>
<td>4.10</td>
<td>4.30</td>
<td>4.50</td>
<td>7</td>
<td>3.30</td>
<td>3.97</td>
<td>4.60</td>
<td>0.33</td>
<td>8.27%</td>
</tr>
<tr>
<td>O</td>
<td>5</td>
<td>4.40</td>
<td>4.70</td>
<td>4.90</td>
<td>24</td>
<td>4.10</td>
<td>4.61</td>
<td>5.00</td>
<td>0.09</td>
<td>1.94%</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>3.50</td>
<td>3.90</td>
<td>4.50</td>
<td>7</td>
<td>4.00</td>
<td>4.33</td>
<td>4.50</td>
<td>-0.43</td>
<td>-9.90%</td>
</tr>
<tr>
<td>Q</td>
<td>5</td>
<td>4.10</td>
<td>4.56</td>
<td>5.00</td>
<td>23</td>
<td>4.30</td>
<td>4.69</td>
<td>5.00</td>
<td>-0.13</td>
<td>-2.80%</td>
</tr>
<tr>
<td>R</td>
<td>2</td>
<td>4.40</td>
<td>4.45</td>
<td>4.50</td>
<td>3</td>
<td>4.30</td>
<td>4.40</td>
<td>4.50</td>
<td>0.05</td>
<td>1.14%</td>
</tr>
<tr>
<td>S</td>
<td>2</td>
<td>4.20</td>
<td>4.30</td>
<td>4.40</td>
<td>6</td>
<td>4.30</td>
<td>4.65</td>
<td>5.00</td>
<td>-0.35</td>
<td>-7.53%</td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>1</td>
<td>4.60</td>
<td>4.60</td>
<td>4.60</td>
<td>5</td>
<td>4.50</td>
<td>4.72</td>
<td>4.90</td>
<td>-0.12</td>
<td>-2.54%</td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>2</td>
<td>4.40</td>
<td>4.65</td>
<td>4.90</td>
<td>0.35</td>
<td>7.53%</td>
</tr>
<tr>
<td>W</td>
<td>1</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>5</td>
<td>3.30</td>
<td>4.16</td>
<td>4.70</td>
<td>-2.66</td>
<td>-63.94%</td>
</tr>
<tr>
<td>X</td>
<td>3</td>
<td>3.10</td>
<td>3.80</td>
<td>4.30</td>
<td>13</td>
<td>4.10</td>
<td>4.52</td>
<td>4.90</td>
<td>-0.72</td>
<td>-15.84%</td>
</tr>
<tr>
<td>Y</td>
<td>3</td>
<td>2.70</td>
<td>3.23</td>
<td>3.80</td>
<td>11</td>
<td>3.80</td>
<td>4.15</td>
<td>4.50</td>
<td>-0.92</td>
<td>-22.17%</td>
</tr>
<tr>
<td>Z</td>
<td>1</td>
<td>2.60</td>
<td>2.60</td>
<td>2.60</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td>156</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.76</td>
<td>3.97</td>
<td>4.16</td>
<td>4.07</td>
<td>4.45</td>
<td>4.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Overall average SET scores (minimum, mean, and maximum) for the MSA were lower when compared to all other programs. On an individual instructor basis, we compared the mean value for the MSA to the mean value for all other programs. The average mean difference for instructors whose MSA mean was higher than the mean for all other programs was 0.17 (3.86%). The average mean difference for instructors whose MSA mean was lower than the mean for all other programs was -0.66 (-15.34%). That is, when an instructor’s average MSA SET score is lower relative to other programs, the decrease is nearly four-times as large as when the average MSA SET score is higher.

APPENDIX B

The following questions guided the interviews:

- What discussions occur within the cohort regarding instructors?
- What discussions occur with specific reference to the instructor evaluation process?
- How do these discussions differ from other settings you experienced?

APPENDIX C

Table C1 presents the MSA gender breakdown during the years covered by this study. The Total Cohort columns show the breakdown for each cohort and in total. The Invited columns show the breakdown for those invited to participate in the study. The
Participants columns show the gender for study participants. The Invited and Participants gender breakdowns reflected that of the Total Cohort.

Table C1. Gender Breakdown by Cohort.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Total Cohort</th>
<th>Invited</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>F</td>
<td>Total</td>
</tr>
<tr>
<td>A</td>
<td>10</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>C</td>
<td>13</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>D</td>
<td>15</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>E</td>
<td>14</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>Totals</td>
<td>67</td>
<td>49</td>
<td>116</td>
</tr>
<tr>
<td>%</td>
<td>57.76%</td>
<td>42.24%</td>
<td>55.55%</td>
</tr>
</tbody>
</table>

Of the 116 total cohort members, all but three had begun the MSA in the same year as their baccalaureate degree, and those three had delayed only one year. As a result, we made no analysis of age effect.