12-19-2018

The creation, validation, and implementation of the Engineering Process Safety Research Instrument

Brittany Lynn Butler
Rowan University

THE CREATION, VALIDATION, AND IMPLEMENTATION OF THE ENGINEERING PROCESS SAFETY RESEARCH INSTRUMENT

By

Brittany L. Butler

A Thesis

Submitted to the
Department of Experimental Engineering Education
College of Engineering
In partial fulfillment of the requirement
For the degree of
Master of Science in Engineering
at
Rowan University
December 5, 2018

Thesis Chair: Cheryl Bodnar, Ph.D.

Acknowledgements

I would like to express my gratitude toward Dr. Cheryl Bodnar for her guidance, encouragement, and assistance during the course of this research project. Her consistent feedback and dedication to this project is very greatly appreciated. This project would not have been possible without her.

I would also like to acknowledge the NSF improving undergraduate STEM education (IUSE DUE#1711376, 1711644, 1711672, 1711866) for supporting this work.

Abstract

Brittany Butler
THE CREATION, VALIDATION, AND IMPLEMENTATION OF THE ENGINEERING PROCESS SAFETY RESEARCH INSTRUMENT
2018-2019
Cheryl Bodnar, Ph.D.
Master of Science in Engineering

The purpose of this study was to develop and validate an instrument that can measure how senior chemical engineering students make process safety decisions. The Engineering Process Safety Research Instrument (EPSRI) contains dilemmas that represent process safety scenarios, followed by three decision options, and 12-15 considerations that fall into pre-conventional, conventional, or post-conventional forms of reasoning. Three studies were completed as a part of this research. The content validation study ensured the dilemmas represented process safety scenarios, the considerations matched their perceived theoretical definitions, and that no content areas were omitted. This study resulted in validation of the content, following the elimination of one dilemma and eleven considerations. The large scale validation study determined the number of underlying latent variables present on the instrument including the correlation between considerations on a factor. The instrument was not able to be fully validated, but resulted in the elimination of one dilemma and six considerations with 22 considerations being revised for further study. The think aloud protocol with the EPSRI determined how students were classified based on their EPSRI scores, and their moral reasoning approaching these dilemmas. From this study, it was found that senior chemical engineering students mainly applied post-conventional reasoning, despite all the students not being classified as post-conventional based on their EPSRI responses.

Table of Contents

Table of Contents (Continued)

Table of Contents (Continued)

Table of Contents (Continued)

List of Figures

List of Tables

Table Page

List of Tables (Continued)

Table Page

Chapter 1

Contribution to Research

Background

The importance of process safety is becoming increasingly apparent in chemical industry, as well as in the classroom. Chemical companies are addressing the need for better procedures and training through a variety of techniques and practices. For instance, Dow Chemical Company implemented discipline systems which significantly reduced their number of tier one process safety events (Champion, Van Geffen, and Borrousch 2017). BP designed process safety modules, applied through an eLearning platform and workshops in an attempt to improve their process safety (Bruyere, Fox, and Watson, 2009). Companies are also able to send their employees to a four day process safety boot camp administered by the American Institute of Chemical Engineering (AIChE), which focuses on the fundamental concepts of process safety (American Institute, n.d.).

Despite these efforts, process safety incidents continue to occur. While some process safety incidents are due to maintenance or management errors, they could also be due to whether individuals are able to identify the ethical implications of the situation during the time of a decision. A study completed by AIChE found that people working in chemical industry recognize the importance of strong ethics in the workplace, which serves to ask the question, why do poor process safety decisions continue to occur (Grubbe, 2018)? This question may be able to be answered by comparing ethics to behavioral ethics.

Behavioral ethics occurs when an individual does not recognize the ethical implication of their decision, compared to ethical decisions which occur when an individual recognizes they are in an ethical dilemma. The way ethics is currently taught is based on the assumption that an individual will recognize an ethical dilemma when it is presented to them, and does not encompass the predictable behaviors that result in unethical actions (Bazerman and Tenbrunsel, 2011). According to Bazerman and Tenbrunsel (2011), individuals experience ethical dilemmas in three phases: the "before" phase, the "during" phase, and the "after" phase. During the "before" phase, an individual makes predictions about how they believe they will behave during an ethical dilemma. These predictions are typically inaccurate, and are referred to as behavioral forecasting errors (Osberg and Shrauger, 1986). The "during" phase occurs when the individual is in the ethical dilemma. At this stage, individuals often behave how they want to, regardless of how they believed they should behave in the "before" phase. Decisions lose their ethical dimensions because the ethical principles do not seem relevant at the time, and unethical decisions can be made (Bazerman and Tenbrunsel, 2011). In the "after" phase, the full implications of a decision begin to settle in, which causes reformulation of decisions to make the individual believe they are still ethical (Bazerman and Tenbrunsel, 2011).

Process safety incidents could also be a result of how process safety is taught in the classroom. While the "consideration of hazards associated with the engineering application of basic sciences" was added to the Criteria of Accrediting Engineering Programs (ABET) student learning outcomes, it wasn't required to be taught in a curriculum until 2012 (Criteria for Accrediting, n.d.; Dee, Cox and Ogle, 2015; Shallcross, 2014). Additionally, it can be difficult to add process safety to the chemical engineering curriculum due to the amount of classes students are already required to take. Some schools have implemented process safety through the addition of new classes, which can result in students having to drop a non-engineering elective (Dee et al., 2015). Other schools have integrated process safety modules into classes that already exist, or offer programs which take place outside of the classroom.

Additionally, there has been little research that shows the evaluation of students' process safety knowledge based on these interventions. Shallcross (2013; 2014) implemented safety shares and case studies into a second year chemical engineering course to promote the importance of process safety. At the conclusion of the study, Shallcross implemented surveys that discovered the effectiveness of the methods, and found the students enjoyed the safety shares and case studies, and recognized the importance of process safety (Shallcross, 2013; Shallcross, 2014). However, the students' knowledge of process safety and their thought process while making process safety decisions was not measured. This could be due to the lack of a validated instrument that can measure how students make process safety decisions.

Purpose of Study

The purpose of this study was to design and validate the Engineering Process Safety Research Instrument (EPSRI), which assists in measuring how students approach process safety decisions based on Kohlberg's Moral Development Theory (Kohlberg and Hersh, 1977). Kohlberg's theory describes moral development in three steps: pre-conventional, conventional, and post-conventional. Pre-conventional reasoning occurs when decisions are made based on avoiding personal consequences, and satisfying one's needs. Conventional reasoning occurs when decisions are made based on the needs of friends and family, and the maintenance of social order. Post-conventional decisions are made based on general individual rights, and respect for human beings as individuals (Kohlberg and Hersh, 1977).

The EPSRI was modeled on the Defining Issues Test 2 (DIT2) and Engineering Ethical Reasoning Instrument (EERI), which both use Kohlberg's Moral Development Theory as their basis (Rest, Thoma, and Narvaez 1999a; Zhu et al., 2014). Both instruments contain five scenarios, and 12 corresponding considerations. In the initial stage, the EPSRI contained eight dilemmas which were accompanied by 15-17 considerations. Additional dilemmas and considerations were created with the anticipation that some dilemmas and considerations would be eliminated during the validation process. In its final form, the ESPRI is projected to contain five dilemmas that represent process safety incidents, and 12 considerations that fall into either pre-conventional, conventional, or post-conventional forms of reasoning.

Multiple studies were conducted as part of the instrument development and validation process. Additionally, an examination of how students were making process safety decisions was completed. The content validation study ensured that the dilemmas were representative of actual process safety incidents that might occur in chemical industry. It also confirmed that the considerations were representative of their perceived definitions, and that no content areas were omitted. Content experts from chemical industry, chemical engineering education, and learning science fields reviewed the instrument as part of the content validation. The large scale validation was completed to analyze and strengthen the correlations between the items on the dilemmas and within schemas, and determined the number of underlying latent variables. This was completed through a factor analysis that was conducted on data obtained from the large scale study. The process safety moral reasoning think aloud study was conducted to analyze how senior chemical engineering students reason through process safety decisions. This was a mixed methods study which sought to answer four research questions: 1) How can the p-score and N2 score be applied to understand students' moral reasoning? 2) What schemas of moral reasoning do senior chemical engineering students demonstrate when performing process safety decisions? 3) How do senior chemical engineering students reason through process safety decisions? 4) Do the schemas of moral reasoning students represent truly reflect their moral reasoning process when approaching process safety decisions?

Study Outcomes and Significance

Three outcomes were obtained from these studies. 1) The content of the EPSRI was reviewed and validated by a pool of content experts. As a result of this study, one dilemma and eleven considerations were eliminated. 2) The EPSRI was able to be reviewed for reliability and construct validity, and proposed changes were made to address issues that were identified during the factor analysis. Ideally, each dilemma should have four underlying latent variables which represent pre-conventional, conventional, post-conventional, and meaningless items. As a result of this study, one dilemma and seven considerations were eliminated, and 22 considerations were revised. Moving forward, the large scale validation will be redone in the Fall of 2018 to allow for another step to be made on validating the instrument. 3) Through the mixed methods think aloud study, it was found that all students mainly applied post-conventional reasoning across the instrument. It was also observed that the moral schema students represented did not necessarily reflect the full spectrum of students' moral reasoning when faced with a process safety dilemma.

Contribution to Educational Research

The validated version of the EPSRI will allow educators to examine how students are making process safety decisions. Additionally, modified versions of scoring mechanisms adapted from Rest et al. (1997a; 1997b) have been provided, which can be used to determine students' quantitative measures for the EPSRI. These scores are able to determine the extent of a students' post-conventional nature based on their responses to pre-conventional and post-conventional considerations. They will also be able to determine if a student is consolidated in their reasoning, or transitioning between two forms of reasoning. Combined with other quantitative measures, educators will be able to classify their students, which is described at length in this paper.

The codebook developed from the think aloud study has also been presented in this work. The codebook can be used to determine pre-conventional, conventional, or post-conventional themes in students' responses to the process safety scenarios included within the EPSRI. The codebook contains a list of codes that fall into pre-conventional,

conventional, or post-conventional forms of reasoning as well as descriptions and examples of each code.

Contribution to Educational Practice

The validated form of the EPSRI will be able to determine how a student morally reasons through process safety decisions. This information can assist educators in moving students toward post-conventional reasoning and evaluating the efforts of their process safety instruction. At the post-conventional stage, individuals are showing a "clear effort to define moral values," and consider much more than the people and surroundings directly involved with the dilemma (Kohlberg and Hersh, 1977, p. 55). The impact of process safety decisions affects more than the individual, and their surrounding work environment. When students make process safety decisions, they should be considering the environment, surrounding communities, and the benefit to society. It is important that students understand and eventually move towards post-conventional reasoning.

This study found that students were operating at a post-conventional level when they participated in the think aloud protocol. However, it is unlikely for senior chemical engineering students to be operating at a level this high, as they should be operating at a conventional level according to prior work by Rest et al. (1999a). These results may be due to behavioral forecasting errors, which occur when a student is working in the predictive phase of reasoning (Osberg and Shrauger, 1986). The students may not fully understand the implications of their decisions because the dilemmas are not real to them. For this reason, educators should address behavioral ethics along with ethics when teaching students about process safety. Students should be encouraged towards the "during" phase as much as possible in order for them to understand the full complexity of different perspectives that surround an ethical dilemma.

Summary

Despite the importance of implementing process safety into the engineering classroom, research has found that there is a lack of a validated instrument that can measure how students make process safety decisions. The purpose of this study is to create and validate the EPSRI, which can be used to measure where students fall in terms of moral schema when making process safety decisions based on Kohlberg's Moral Development Theory (Kohlberg and Hersh, 1977). Three studies were conducted in attempt to validate the instrument, and analyze how senior chemical engineering students make process safety decisions.

The content validation study ensured the dilemmas were relevant and reflected real process safety situations, the considerations matched their perceived definitions, and no content areas were omitted. One dilemma and eleven considerations were eliminated as a result of the content validation study. The large scale validation study sought to determine the number of underlying latent variables, and find the correlations between the items within the dilemma, and the factors. The instrument was not validated as a result of this initial study, however, one dilemma and seven considerations were eliminated, and 22 considerations were revised. The think aloud study sought to answer four research questions, which were previously stated. From this study, it was determined that the moral schemas students represent do not necessarily reflect the full spectrum of their reasoning.

Chapter 2

Literature Review

The Importance of Process Safety

Process safety in industry. The importance of process safety is becoming increasingly prominent in the chemical industry. However, the chemical safety board has still documented over 800 process safety incidents since its foundation 20 years ago. One of the well-known process safety incidents was the "ExxonMobil Baton Rouge Refinery Isobutane Release and Fire" (Key Lessons, 2016). During this incident, two operators performing maintenance on an isobutane pipeline were attempting to open a plug valve that was attached to a gearbox. When the valve failed to open, the operators removed the gearbox as directed by the company's practices. However, the gearbox they were attempting to remove was attached to a pressure retaining piece known as a "top cap with a bracket." About 3% of the company's valves contained this old design in which the bracket was attached to the side of the gearbox and the top cap. Operators incorrectly removed the gearbox and altered the top cap in a way that allowed 2,000 pounds of isobutane to leak into the atmosphere. Within seconds, the cloud of isobutane ignited and severely burned one ExxonMobil employee and three contractors who were working on site (Key Lessons, 2016).

When investigating this process safety incident, the chemical safety board discovered that ExxonMobil management accepted the practice of removing the gearbox, however, adequate written procedures and training were not available to the operators (Key Lessons, 2016). The training that the operators received discussed appropriate removal of a gearbox, but did not include information on how to remove a gearbox from the older design. The chemical safety board concluded more detailed and accurate written procedures should be provided to the operators conducting dangerous work, and that all workers participate in accurate training to ensure they can adequately perform their job tasks (Key Lessons, 2016).

Process safety incidents like this have pushed companies to become more dedicated to process safety. Some companies have decided to implement discipline systems, such as Dow Chemical Company (Champion et al., 2017). By prioritizing safety culture and leadership, process safety systems and operational discipline, Dow was able to largely decrease the amount of tier one process safety incidents from 69 in 2008 to an average of 10 between 2013 and 2015. A tier one process safety event is described as an incident with great consequence that was a result of loss of containment (Recommended Practice, 2010; Champion et al., 2017). Following the BP Texas explosion in 2005, the Baker Panel Report proposed that BP become a leader in process safety (Baker et al., 2007; Bruyere et al., 2009). In response, Bruyere et al. (2009) developed an interactive, electronic learning platform as well as workshops that encompassed process safety. The process safety modules were designed and implemented through eLearning and workshops that were meant to improve process safety performance at BP. Implementation of the modules in 2008 provided successful results, with positive feedback and comments on the easy accessibility of the platform (Bruyere et al., 2009).

A different method proposed by Carvalho Neto and Correa (2017) allows for a better understanding of the competencies necessary for employees to perform functions and roles through a three-dimensional matrix. According to Risk Based Process Safety guidelines, the development and maintenance of process safety competency includes the continuous improvement of knowledge and competency, ensuring that individuals have access to the appropriate information, and consistent application of what has been learned (CCPS, 2014). The three-dimensional matrix contains trainings and courses, positions and roles, and proficiency levels (Carvalho Neto and Correa, 2017). The application of the three-dimensional matrix allows for a more organized practice of training that encompasses the understanding and maintenance of competency within process safety (Carvalho Neto and Correa, 2017). The American Institute of Chemical Engineering offers a four day process safety boot camp to train individuals on the fundamental concepts of process safety. Companies who are interested have the ability to send their staff to the boot camp in order to have a team better trained on process safety (American Institute, n.d.)

While the implementation and creation of these platforms and projects are promising, the problem of reoccurring process safety incidents has not been solved. This may be due in part to the lack of understanding of the ethical behaviors that lie within the companies. Process safety training may help with understanding and maintenance of adequate process safety knowledge, but most times, process safety incidents occur as a results of a decision that was made in a situation that had ethical implications. The following section will describe the differences between ethics and behavioral ethics, and what may lead to some of these decisions that have resulted in process safety incidents.

Behavioral ethics vs. ethics. Process safety incidents are due in part to poor decisions that were made in a situation that has ethical implications. However, most people who work in the chemical engineering field recognize that ethics are important, and also believe that they are ethical individuals. In 2016, AIChE conducted an "ethics survey" to determine the importance of ethics and the ethical beliefs of individuals in chemical engineering (Grubbe, 2018). Surveys were sent directly to AIChE members, and received a total of 1,346 completed surveys within 17 days. The results from the survey found that 96% of respondents believed that ethical behavior was important within their job role, and 99% of respondents had rated the importance of ethical behavior from professional members as very or extremely important (Grubbe, 2018). The survey also found that respondents within the United States had rated the importance of ethical behavior and acting in an ethical manner as "extremely" important, rated their work environments as "mostly" ethical, and had faced ethical dilemmas within their careers (Grubbe, 2018). If the individuals working within these industries recognize the importance of ethics, then why are poor ethical decisions taking place that lead to process safety incidents? The answer may lie within the difference between ethics and behavioral ethics.

Behavioral ethics describes how an individual will act when faced with an ethical dilemma (Bazerman and Tenbrunsel, 2011). The difference between behavioral ethics and ethics is the level of awareness that an individual has about the situation being an ethical dilemma. Ethics focuses mainly on the behaviors and decisions that are made when an individual is aware of the ethical dilemma they are facing. Ethical training is not adequate, and will continue to fail since it is based on the assumption that an individual will recognize an ethical dilemma when it is presented to them (Bazerman and Tenbrunsel, 2011). Ethical training should encompass the predictable cognitive behaviors that result in unethical behavior, and the way ethics are bounded (Bazerman and Tenbrunsel, 2011). This can be described through behavioral ethics, which occurs when an individual is unaware that they are in an ethical dilemma. Within behavioral ethics is the recognition and understanding of bounded awareness and ethical fading. Bounded awareness describes the tendency to exclude important information from a decision due to the arbitrary and dysfunctions associated with bounds that are placed around the definition of a problem (Bazerman and Chugh, 2006). Ethical fading occurs when the ethical dimensions of a decision are eliminated in a way that an ethical decision may appear as a business decision (Tenbrunsel and Messick, 2004). For example, the explosion of the Challenger shuttle occurred as a result of ethical fading. The decision to launch the shuttle was framed as a business decision: launch the shuttle or lose the contract which was associated with a lot of money. The engineers and managers who made the decision to launch the shuttle were aware that it may fail due to the O-rings, but due to ethical fading, this ethical decision was framed as a business decision instead, which resulted in the deaths of several people (Bazerman and Tenbrunsel, 2011).

Decision making in behavioral ethics can be described by system one and system two thinking (Stanovich and West, 2000). System one thinking, or thinking employing the intuitive system, involves quick snap judgements that are made based on emotions, and are largely due to mental overloading (Kahneman, 2003; Bazerman and Tenbrunsel, 2011). System one thinking occurs more often when the mind becomes overloaded with information, such as at the end of a workday (Bazerman and Tenbrunsel, 2011). It is efficient and quick and is appropriate for most decisions made on a daily basis (Kahneman, 2003). However, system one thinking may result in decisions that are different from decisions that would have been made with more deliberation (Bazerman and Tenbrunsel, 2011). System two thinking, or the logical decision making system, requires more slow and deep conscious thought (Kahneman, 2003). The cost and benefits of both decisions are weighed and deliberated in an organized manner within system two thinking, which typically results in more ethical decisions (Bazerman and Tenbrunsel, 2011). While system one thinking is appropriate for most decisions, a problem occurs when system two thinking is never utilized. System two thinking should be used when making important decisions with ethical implications. Using only system one thinking creates a gap between the desired and executed ethical behaviors (Bazerman and Tenbrunsel, 2011).

Poor ethical judgement and decisions could also be due to the psychological process that individuals function under when considering an ethical dilemma. This process can be broken into three phases: the "before" phase, the "during" phase, and the "after" phase. The "before" phase describes the way an individual believes they will behave in an ethical dilemma (Bazerman and Tenbrunsel, 2011). The decision an individual believes they will make often encompass behavioral forecasting errors, which are inaccurate predications an individual will have about their behavior. Research has shown that individuals tend to incorrectly predict their behavior (Osberg and Shrauger, 1986). For example, a group of female college students were asked how they would behave if a male asked them inappropriate questions during a job interview. The study found that 62 percent of the women would confront the interviewer about the inappropriate questions, and 68 percent of the women said they would refuse to answer the questions. However, when the women were placed in the situation, none of the women refused to answer the questions, and 36 percent of the women asked the interviewer why it was necessary to answer those questions. This was typically done at the end of the interview and in a polite manner (Woodzicka and LaFrance, 2001).

The way the women acted when placed in the interview describes the “during” phase. Within the “during” phase, an individual’s thoughts are dominated by how they want to behave, regardless of how they believed they should behave in the “before” phase (Bazerman and Tenbrunsel, 2011). Decision making in this phase largely reflects system one thinking, where decisions are made quickly and emotionally. Ethical fading plays a role in the decision phase as well (Bazerman and Tenbrunsel, 2011). In the “before” phase, an individual is able to see the ethical aspect and impacts of their decision. However, in the “during” phase, a decision may lose its ethical dimension because a decision is viewed as a business or legal decision instead of an ethical decision. This occurs because the ethical principles of the decision don’t seem relevant at the time, so unethical decisions may be made (Bazerman and Tenbrunsel, 2011).

Following the “during” phase is the “after” phase. The “after” phase occurs once distance is gained from the decisions that were made in an ethical dilemma in the “during” phase (Bazerman and Tenbrunsel, 2011). The full implications of the decision begin to settle in, which triggers psychological cleaning (Bazerman and Tenbrunsel, 2011). Psychological cleaning is a result of the recognition of the discrepancies between how one wants to behave ethically, and how one actually behaves. Psychological cleaning is a smaller part of moral disengagement, where individuals behave unethically or contrary to their ethical beliefs while still maintaining that they are ethical people. People tend to look at their actions and reformulate their decisions to reflect why their decisions was the right one to make (Bazerman and Tenbrunsel, 2011).

The three phases of decision making fall into either the “want” or “should” self (Bazerman, Tenbrunsel, and Wade-Benzoni, 1998). The “should” self describes decisions that are made rationally, cognitively, and thoughtfully. Within the “should” self are the ethical beliefs one should have according to recognized ethical values. The “should” self often dominates during the “before” and “after” phases of the decision making process. In the “before” phase, the full ethical implications of a decision are acknowledged, and an individual believes they will behave ethically. During the “after” phase, an individual recognizes that they should have behaved ethically, and will reformulate their decisions to make them believe that they had behaved ethically (Bazerman and Tenbrunsel, 2011). The “want” self focuses on decisions that are made impulsively and emotionally, similar to system one thinking. The “want” self reflects how an individual actually behaves when placed in an ethical dilemma, and their true understanding of the ethical implications their decision carries. The “want” self dominates in the “during” phase because an individual no longer sees the decision as an ethical dilemma (Bazerman and Tenbrunsel, 2011). An understanding of behavioral ethics within the workplace may help reduce the amount of unethical decisions and process safety incidents.

Process safety in the classroom. The addition of "consideration of hazards associated with the engineering application of basic sciences" to the program criteria for chemical engineering within the Criteria for Accrediting Engineering Program (ABET) in 2012 was as a result of a process safety incident (Criteria for Accrediting, n.d.; Dee et al., 2015; Shallcross, 2013). In 2007, a fifth-year and first-year graduate student at Texas Tech decided to scale up the synthesis of nickel hydrazine perchlorate (NHP) to 10 grams without consulting their professor (T2 Laboratories, 2009). At the time, Texas Tech had no guidelines that would have informed them to seek approval from their professor prior to scaling up the synthesis. At smaller quantities, the students observed that the compound would not explode when in contact with hexane or water. Understanding this, the senior graduate student used hexane in order to break down clumps of the NHP, which resulted in the compound detonating. The senior student, who was not wearing eye protection at the time, lost three fingers, burned his hands and face, and injured one of his eyes (T2 Laboratories, 2009). As a result of this incident, process safety became a requirement for chemical engineering students who attended ABET accredited schools (Dee et al., 2015; Shallcross, 2013).

As of 2018, the ABET general criteria for learning outcomes for engineering students includes the "ability to design a system, component or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability" (Criteria for Accrediting, n.d.). The addition of a process safety module or class to an already full curriculum for chemical engineering undergraduate students is difficult for several reasons (Dee et al., 2015). A student who is taking four to six classes per semester may be detracted from other electives in order to take a class on process safety. Another option would be combining two courses, which would be difficult since it is already a challenge to cover the material of one course in a semester (Dee et al., 2015). In response to the ABET addition, institutions are teaching process safety through the addition of a course, a module, or a combination of the two (Dee et al., 2015). Through brief research, Dee et al. (2015) found a variety of ways that institutions approach the implementation of process safety into the curricula. Universities such as Georgia Institute of Technology and Texas Tech offer courses that encompass chemical process safety. The courses range from one to three credit hours, and cover a multitude of topics. Other institutions, such as Syracuse University and University of Pittsburgh, offer process safety courses outside of the chemical engineering curricula (Dee et al., 2015). Further yet, institutions have integrated process safety modules into existing courses, such as what was done at the Massachusetts Institute of Technology, or provided opportunities for students to learn about this material through co-curricular or extra-curricular activities (Johnston, n.d.; McRae, n.d.). Chemical engineering students at West Virginia University participated in a two day Process Safety Boot Camp in 2013 (Dillon, 2013). Northeastern University has implemented a process safety program that is offered during spring break which allows students to travel to a facility and complete a HAZOP analysis (Dee et al., 2015).

While the addition of process safety courses and modules to the chemical engineering curricula is promising, there is little research that shows evaluation of students' process safety knowledge. Dee et al. (2015) suggested evaluating process safety culture based on how the students value process safety, or evaluating process safety knowledge through a metric such as the Fundamentals of Engineering exam, or the ABET criterion. In 2014, Shallcross implemented safety shares in second year chemical engineering course to promote the importance of process safety. A safety share was a two to four minute presentation that took place at the beginning of every lecture and discussed different aspects of safety that were mainly drawn from the process industry. Over 50 safety shares were presented throughout the course and covered topics such as situational awareness, pressurized pipes, and enforcing safety rules at all measures. At the conclusion of the class, Shallcross (2014) implemented a survey that sought to determine the effectiveness of the safety shares in terms of instruction of process safety. Shallcross (2014) found that the students enjoyed the safety shares and recognized the importance of process safety. However, the students' knowledge of process safety was not found or measured. Shallcross (2013) had also previously implemented oral presentations in a second year chemical engineering class as a method of teaching process safety. At the conclusion of the study, Shallcross (2013) studied the effectiveness of the oral presentations, but the students' knowledge of process safety still was not measured.

While the implementation of various courses and modules to the chemical engineering curricula to teach process safety is promising, there is no validated way to measure if students are gaining process safety knowledge. Additionally, little research has been completed on how students make process safety decisions. The goal of the Engineering Process Safety Research Instrument (EPSRI) is to assist in measuring students' process safety decision making. The following sections describe the development and validation of the EPSRI.

Designing and Validating an Instrument

Creation of the EPSRI. The Engineering Process Safety Research Instrument (EPSRI) was based on two ethical dilemma instruments: the Defining Issues Test 2 (DIT2) and the Engineering Ethical Reasoning Instrument (EERI) (Rest et al., 1999a; Zhu et al., 2014). These instruments were chosen as the base for the EPSRI since they are both validated instruments that have been used to measure ethical reasoning in participants. Both instruments contain five dilemmas which each contain three decision options and twelve considerations. The structure of these instruments begins with an ethical dilemma, that is followed by three options. Two of the options take action, and one allows the participant to opt out of making a decision. The EPSRI follows this format, but the dilemmas are meant to represent possible process safety scenarios.

Considerations are meant to represent the possible fallout that could occur when making a decision. Considerations fall into either pre-conventional, conventional, or post-conventional thinking as described by Kohlberg's moral development theory (Kohlberg and Hersh, 1977). Kohlberg stated in his theory that moral development "represents the transformations that occur in a person's form or structure of thought (Kohlberg and Hersh 1977, p. 54)." Kohlberg found that moral development occurred through six stages, which represented three characteristics. Kohlberg discovered that stages are "structured wholes," which means that an individual working within that stage is constant in that form of thinking (Kohlberg and Hersh, 1977, p. 54). Kohlberg also discovered that the stages form a sequence in which individuals can only progress through, and never move backwards. Lastly, Kohlberg found that stages are "hierarchical integrations," which indicates that someone thinking at a higher level will also comprehend lower level thinking (Kohlberg and Hersh, 1977, p. 54). For example, someone reasoning at a post-conventional level will also use pre-conventional and conventional reasoning in their decision making.

There are three levels of thinking which each include two stages. The first two stages of thought are encompassed by pre-conventional thinking. At the pre-conventional level, decisions are made based on personal consequence, wants or needs. This first stage is the "punishment-and-obedience orientation," where individuals make decisions based on physical consequence, and avoiding punishment (Kohlberg and Hersh, 1977, p. 54). The second stage is the "instrumental-relativist orientation," where decisions are made based on satisfying ones needs or wants (Kohlberg and Hersh, 1977, p. 54). The next two levels fall under conventional reasoning, which represents decisions that are made based on conformity to personal expectation and maintaining social order. The third stage is the "interpersonal concordance" stage, where decisions are made to please or benefit others (Kohlberg and Hersh, 1977, p. 55). The fourth stage is the "law and order orientation," where decisions are made based on the law. At this stage, one will prioritize authority, and maintaining social order (Kohlberg and Hersh, 1977, p. 55). The final two stages fall under post-conventional thinking, where a "clear effort to define moral values" is shown

(Kohlberg and Hersh, 1977, p. 55). The fifth stage is the "social-contract, legalistic orientation," where decisions are made based on general individual rights that are generally agreed upon throughout society (Kohlberg and Hersh, 1977, p. 55). This differs from the fourth stage, because while decisions are still made based on the legality of them, there is an emphasis on changing the law based on the situation at hand. The sixth stage is the "universal-ethical-principle orientation," where decisions are made based on justice, the equality of human rights, and the respect for humans as individual persons (Kohlberg and Hersh, 1977, p. 55).

Validation of the EPSRI. Validation is essential to an instrument because it ensures that the instrument is measuring what is intended to be measured (Bannigan and Watson, 2009). Validating an instrument studies the interpretation and meaning of the items (Bannigan and Watson, 2009). There are different ways to validate an instrument, and a variety of methods is typically recommended when validating (McDowell and Newell, 1996). DeVellis (2012) mentions three forms of validation: content validation, criterion-related validity, and construct validity. This study will focus mainly on the content validity of the EPSRI through professionals and content experts in the chemical engineering field, and construct validity through exploratory factor analysis.

Content validation. According to DeVellis (2012), content validity is concerned with the appropriateness of the items in the content domain. The content domain of the instrument should be clearly defined and widely understood by the instrument's intended audience (DeVellis, 2012). Content validity ensures that the domain includes only relevant items and that all irrelevant items are excluded (Bannigan and Watson, 2009).

The objective of content validation is to ensure all items are relevant to the content domain, the items represent the definitions applicable to the construct, and that no areas have been omitted. This type of review should be completed by someone knowledgeable to the content area, or content experts (DeVellis, 2012). By completing content validation, the instrument becomes authentic, direct, and entirely relevant to the content domain that is being measured (Bannigan and Watson, 2009).

Exploratory factor analysis**.** Schonrock-Adema *et al.* (2009) stated in their research that factor analysis is a highly useful method that is used to validate the internal structure of an instrument. However, in their study they found that many of the quantitative studies completed in a classroom were being completed in an invalidated manner. They discovered that quantitative studies that utilized factor analysis to validate the results were difficult to find (Schonrock-Adema, Heijne-Penninga, Van Hell, and Cohen-Schotanus, 2009). Factor analysis is important to use while validating an instrument because it is able to identify the commonalities in a set of variables in order to create a smaller set of derived variables, or factors, that give results meaning (Briggs and Cheek, 1986).

Factor analysis can also be used to validate items that have been grouped together by a researcher. For example, the items in the EPSRI are classified as pre-conventional, conventional, post-conventional, or M-items. M-items are meaningless items that are meant to assist in detecting unreliable data (Rest, Narvaez, and Thoma, 1999a). Completing a factor analysis ensures that the items have been grouped or defined correctly. According to DeVellis (2012), factor analysis serves four purposes. The first is determining the number of latent variables in a set of items. The second purpose is condensing information, which allows for variation in the items to be summarized into a few variables. The third purpose is defining the meaning of the factors by looking at the items within the factor. For example, if one of the factors on the EPSRI contained all of the post-conventional items, the factor would be identified as post-conventional. The final purpose of factor analysis is identifying weak items. If an item weakly correlates with other items within its factor, or it does not fall into a factor at all, it is identified as a weak item (DeVellis, 2012).

Factor analysis has been used previously in engineering education for a variety of reasons. Ha *et al.* (2017) used factor analysis to identify unobserved, latent traits of civil engineering understanding of the Statistics Concept Inventory. A university in South Korea used factor analysis to discover and understand the low retention rate of women in engineering (Youn and Choi, 2016). Chu *et al.* (2014) used factor analysis to determine elements that influenced students' participation, and their significance during an engineering project. For this study, factor analysis will be used to determine the strength of the items, and to verify their classification as pre-conventional, conventional, post-conventional, or M-items.

Summary

Research has shown that chemical companies are becoming dedicated to process safety, and are attempting to improve process safety knowledge through a variety of programs, workshops and trainings. However, process safety incidents are still occurring at an alarming rate. In order to reduce the number of process safety incidents, undergraduate students must be well trained in process safety before they reach industry. Universities and colleges have begun to implement process safety into the chemical engineering curriculum through new classes, modules, or a combination of both. However, there is not a validated instrument that can measure how students are making process safety decisions.

The purpose of this dissertation was to create and validate an instrument that can assist in measuring how students make process safety decisions. The Engineering Process Safety Research Instrument (EPSRI) was modeled on two ethical reasoning instruments that use Kohlberg's Moral Development Theory as their basis. The EPSRI underwent a validation process through content validation and factor analysis. Following the validation, the ESPRI was tested in a think aloud protocol that determined how senior chemical engineering students make process safety decisions.

Chapter 3

Content Expert EPSRI Validation Study

Overview

The objective of this study was to perform a content validation of the Engineering Process Safety Research Instrument (EPSRI). This chapter will review the development of the EPSRI, the pool of experts who reviewed the instrument, the validation procedure, as well as the results and conclusions from the study. The initial EPSRI contained eight dilemmas, which were accompanied by three options and 15-17 considerations. The dilemmas represented process safety scenarios that occurred in a chemical engineering environment. Two of the options that followed allow participants to make a dilemma on the situation, and the third suggests not taking any action. The considerations fell into pre-conventional, conventional, or post-conventional forms of thinking as described by Kohlberg and Hersh (1977). Participants are meant to rate considerations based on how much they played a role in their overall decision.

Professionals who worked in the chemical industry, chemical engineering faculty members, or individuals in the engineering education and/or learning science field served as the content experts for this study. Content experts were asked to review the instrument to ensure that the dilemmas were valid process safety scenarios that realistically would occur in a chemical engineering environment. They were also asked to determine how well the considerations related to the definitions of Kohlberg's moral development theory (pre-conventional, conventional, and post-conventional) (Kohlberg and Hersh, 1977).

Two content experts from chemical industry, five chemical engineering faculty, and five learning science and engineering education faculty participated in the study.

After analyzing results from the content experts, one dilemma and eleven considerations were eliminated. Additionally, four dilemmas and seven considerations were revised.

Introduction

Engineering Process Safety Research Instrument (EPSRI). The need for stronger process safety knowledge is becoming increasingly apparent, as discussed previously in Chapter 2. Companies are making efforts to train employees, and are expecting recent graduates to be knowledgeable in process safety. In response, new process safety courses and modules have been implemented into the chemical engineering curriculum (Dee et al., 2015). However, students are not being evaluated on their process safety decision making due to the lack of a validated instrument. The EPSRI serves as a tool that evaluates students' process decision making, and assesses the development of their process safety knowledge over time.

The structure of the EPSRI closely follows the DIT2 and EERI previously discussed in Chapter 2 as it involves a series of dilemmas, followed by a decision for the individual to make given three choices and then considerations pertinent to the dilemma for the individual to rank (Rest et al., 1999a; Zhu et al., 2014). The first version of the EPSRI contained eight dilemmas each of which included 15-17 considerations. The EERI and DIT2 both contain 5 dilemmas which include 12 considerations (Rest et al., 1999a; Zhu et al., 2014). Additional dilemmas and considerations were created with the understanding that some may be eliminated during the validation process. The dilemmas were developed based on process safety case study investigations taken from the chemical safety board website (csb.org), or industrial experience. Of the five individuals involved with the development of the EPSRI, three were responsible for creating two dilemmas each, and two were responsible for creating one dilemma each. Each dilemma was reviewed and revised by every group member after the initial version of the EPSRI was created to suggest overall improvements, or to improve grammar and wording. An example of a dilemma can be seen below.

As a design engineer at a large plastics manufacturing facility, your responsibilities include identifying specifications for replacement parts and new equipment. One of the manufacturing processes in your facility requires transporting a dangerous chemical which is fatal to humans upon exposure through metal hoses lined with polymer. You must choose the hose used to carry this dangerous chemical, and you find there are two options for these hoses: Option A is a hose lined with a polymer that slowly breaks down upon exposure to the chemical - it is fairly inexpensive, but would need to be replaced on a monthly basis to avoid leaks and/or accidental sudden discharge of the chemical in the facility. Option B has a more expensive polymer liner which offers greater resistance to attack by the chemical, and so the hose would only need to be replaced each year. You determine that it would be more expensive to specify Option B on an annual basis. Option A will require additional maintenance effort where it poses higher risk, but would offer savings with a similar level of safety under normal operation. You receive a $5,000 bonus for each year that you keep equipment costs below a certain level, and you are currently projected to spend above this threshold level - specifying Option A would go a long way toward producing the savings needed to receive your bonus.

Following each dilemma were three options which allowed students to pick from two opposite actions, or choose that they could not decide on a course of action. One of the options for the dilemma described above was choosing option A, which was the cheaper hose which is replaced more frequently. The opposite option was option B, which was the more expensive hose which is replaced on a yearly basis. The final option was not deciding on a course of action. Following the options are the considerations.

Considerations fell into either pre-conventional, conventional, or post-conventional forms of thinking. Pre-conventional considerations reflected motives or decisions that are concerned with one's self (Kohlberg and Hersh, 1977). An example of a pre-conventional consideration that accompanied the dilemma described above was "What would you do with the money you could receive as a bonus?" Conventional considerations reflected motives or decisions that dealt with family, friend, or co-workers (Kohlberg and Hersh, 1977). An example of a conventional consideration that accompanied the dilemma above is "Do you really want to put more work and risk on your employees by requiring them to replace the Option A hoses each month?" Post-conventional considerations reflected motives or decision that were concerned with outside communities, the environment, and general moral values (Kohlberg and Hersh, 1977). An example of a post-conventional consideration that accompanied the dilemma above is "Is it ever a good idea to rely on active measures (employee maintenance) rather than passive measures (material of construction)?" Meaningless items, or M-items, were also included to detect unreliable data as described by Rest et. al. (1999a). While creating the DIT2, Rest et. al. (1999a) included M-items that were mixed in with the other considerations. When analyzing student responses, Rest et al. (1999a) was able to detect unreliable data if a student was rating the M-items too highly, or ranking them too often. An example of an M-item that accompanied the dilemma above is “Do the polymer linings for each option come in different colors?” Table 1 below shows the amount of pre-conventional, conventional, and post-conventional considerations in each dilemma.

Table 1

Amount of considerations per dilemma

	Pre-conventional	**Conventional**	**Post-conventional**	**M-items**	**Total**
Dilemma 1	4	5	4	2	15
Dilemma 2	4	4	4	3	15
Dilemma 3	4	5	5	1	15
Dilemma 4	4	5	5	1	15
Dilemma 5	4	4	4	4	16
Dilemma 6	4	4	4	3	15
Dilemma 7	5	4	4	3	16
Dilemma 8	5	5	5	2	17

Students who take the EPSRI are asked to rate the considerations based on how much they played a role in their decision making process. Considerations can be rated on a scale from zero (none) to five (much). In the initial state, each dilemma contained 12 considerations. Each group member added one consideration to every dilemma that was not theirs for a total of 15-17 considerations per dilemma. Each consideration was reviewed and revised by all group members.

In order to validate the content of the instrument, professionals who worked in the chemical industry, chemical engineering faculty members, or individuals in the engineering education and/or learning science field reviewed the instrument. The content experts reviewed the dilemmas for relevance, reviewed the considerations to make sure they related to their perceived definitions, and ensured that no areas were emitted. Following the content validation, one dilemma and eleven considerations were eliminated. This study serves as the first step toward validating the EPSRI.

Methods

Content validation. The EPSRI was validated following the process outlined in DeVellis (2012). First, DeVellis (2012) stated that a group of people who are knowledgeable in the content area should review the item pool. Professionals in the chemical engineering industry, the learning science and engineering education field, and chemical engineering faculty were selected as appropriate experts for this study. DeVellis (2012) also stated that professionals should review the set of items to ensure they represent the definitions applicable to the construct. Lastly, DeVellis (2012) stated that professionals should review the list of items, and suggest any content areas that may have been omitted. Content experts had the opportunity to provide feedback on the items throughout the survey. Proper human subject's approval was obtained prior to conducting this study.

Dilemma review. Content experts in the chemical industry and chemical engineering faculty participated in the dilemma review. Individuals in the learning science or engineering education field were omitted from this part of the survey as they may not have had as strong of an understanding about chemical engineering scenarios as the other two groups of content experts. The dilemmas were reviewed for relevance, and to ensure that they were situations that could occur in a chemical engineering environment. Participants were able to rate the dilemma on a scale from one (not relevant) to three (relevant). Participants were also able to provide feedback on dilemmas that received poor ratings.

After the surveys were completed, the researchers averaged the ratings for each dilemma. Dilemmas with an average rating below a 2 were eliminated from the instrument. Dilemmas with an average rating above a 2 were kept and revised according to the feedback provided. However, if all researchers felt that the feedback neither improved nor clarified the dilemma, it was left alone. A rating of 2 was chosen as the threshold because it represented a dilemma that was somewhat relevant. Falling below a two meant it was not representative of a dilemma that had importance within the chemical field.

Consideration review. All content experts participated in the consideration review portion of the survey. Prior to reading through the considerations, participants were provided with a brief definition for pre-conventional, conventional and post-conventional thinking. The considerations were rated on how well they related back to their perceived definitions. Ratings were on a scale from one (low) to three (high) in terms of alignment. M-items were not reviewed by the content experts. Participants were also able to provide feedback on considerations they rated poorly.

After the surveys were completed, researchers found the average rating for each consideration. Considerations that were rated below a 2 were eliminated. Considerations that were rated between a 2 and 2.5 were kept if there was feedback that helped clarify or improve the consideration. If there was no feedback for the consideration, or if there was a sufficient amount of considerations for that form of thinking, it was eliminated. If a consideration was rated above a 2.5, it was kept and revised according to the feedback provided. However, if all researchers felt that the feedback did not improve the clarity of the consideration, it was left alone. A rating of 2 was chosen as the minimum threshold for elimination because a 2 on the rating scale represented a consideration that moderately related back to its perceived definition. However, the researchers wanted to ensure that the considerations that remained were as close as possible to a rating of a three. If all considerations rated below a 2.5 were eliminated, some dilemmas would have fallen below the 12 consideration minimum. For this reason, a second threshold of 2.5 was used in the review process.

Results and Discussion

A total of 12 content experts participated in the study. The surveys were voluntary, and were both sent to 2 industry members, 3 chemical engineering faculty, and 3 learning science and engineering education faculty for a total of 8 experts per survey. The content experts came from a list of network contacts of the researchers. Out of the eight content experts that were sent the first survey, seven completed the assessment, one (14%) was from the chemical industry, three (43%) were chemical engineering faculty, and three (43%) were from the learning science or engineering education field. Of the eight content experts who were sent the second survey, five completed the assessment, one (20%) was from the chemical industry, two (40%) were chemical engineering faculty, and two (40%) were from the learning science or engineering education field.

Dilemma review. Based on the methods previously described for the dilemma review, one dilemma was eliminated from the instrument, and the remaining seven were kept. The initial instrument contained eight dilemmas due to the anticipation that some would be eliminated throughout the validation process. The DIT2 and EERI instruments (Rest et al., 1999a; Zhu et al., 2014) that the EPSRI are based on only contain five dilemmas. Starting with eight dilemmas allowed for flexibility when validating the instrument with the understanding that the EPSRI will eventually only contain five dilemmas. Table 2 shows the average and individual ratings for each of the dilemmas.

Table 2

Individual and average scores for the eight dilemmas

		Not relevant (1)	**Moderately relevant (2)**	**Very relevant (3)**	**Average score**
Survey 2	Dilemma 8	2	0	1	**1.7**
	Dilemma 1	0	0	3	**3.0**
	Dilemma 2	0	0	3	**3.0**
	Dilemma 3	0	0	3	**3.0**
Survey 1	Dilemma 4	1	1	2	**2.3**
	Dilemma 5	0	1	3	**2.8**
	Dilemma 6	0	1	3	**2.8**
	Dilemma 7	0	1	3	**2.8**

The eighth dilemma was eliminated from the instrument due to its low rating. This dilemma described a situation in which the participant is the manager of a design project. The manager has the final say on a design before it is built, and is waiting on a safety review from one of their best engineers. The safety review is taking longer than expected, and is due to management by 9 PM. The manager's child's birthday party is at

6 that night, and it is becoming apparent that the manager has to either trust his engineer to complete the safety review, or miss the birthday party to read and sign off on the safety review once it is finished. Content experts found that this dilemma was not relevant since reports wouldn't have an absolute deadline, and the manager would be able to go back after their child's birthday party. For these reasons, the dilemma received a low rating, and was eliminated from the instrument.

An example of a dilemma that was revised according to feedback obtained was the first dilemma. This dilemma takes place in a plastics manufacturing facility where the participant, who is a design engineer, has to choose between two polymer lined hoses. The first hose is cheap, but must be replaced on a monthly basis. The second hose is replaced yearly, but has a higher cost associated with it. The participant, or the design engineer, also gets a bonus if they keep production costs below a certain value. The dilemma is deciding which polymer lined hose to use. Content experts mentioned that it wasn't much of a dilemma since there wasn't any difference in safety stated in the dilemma. Specifically, the dilemma had stated that "option A [would] require additional maintenance effort where it poses higher risk, but would offer savings with a similar level of safety under normal operation." This was meant to imply that safety was similar during normal operation, but a risk was associated when the hoses were being changed, or while it wasn't under normal operation. The dilemma was revised to reflect the potential hazard that associated a more frequent change of hose. This was done by adding the statement "each monthly replacement of the Option A hose brings with it elevated risk of exposure to the fatal chemical, but maintenance procedures are in place to minimize the risk under normal conditions."

An example of a dilemma where suggestions for change were not implemented was the fifth dilemma. In this dilemma, the participant is an operator at an oil company who has been asked by their supervisor to open a pumps inlet plug. However, the valve won't open the way it is meant to, so the co-worker suggests to manually open it even though it may be unsafe. The dilemma is whether to listen to the co-worker, or ask the supervisor. Feedback on this dilemma suggested that one should not worry about asking for help. The researchers decided to not revise the dilemma based on the feedback since the instrument is aimed toward students. Through informal conversations with employees at chemical companies, the researchers found that students, or new hires, often do not recognize when they need to ask for help. This student perspective may be different for someone who is familiar with or has worked in the chemical industry which could be the reason for the feedback obtained from the experts evaluating the instrument.

A summary table that shows which dilemmas were eliminated, revised, or kept is shown in Table 3 below. One dilemma was eliminated, and three were kept. Minor revisions were made to the remaining four dilemmas, such as word changes, or implementing statements that were suggested by content experts.

Table 3

Summary table for dilemmas.

Dilemma Number							
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Revised	Kept	Kept	Revised	Revised	Revised	Kept	Eliminated

Consideration review. Eleven considerations were eliminated from the instrument, and five considerations were revised following the content validation procedure. Of the eleven eliminated considerations, six (55%) were pre-conventional, three (27%) were conventional, and two (18%) were post-conventional. This total does not account for the considerations that were eliminated as part of the first dilemma. The average rating for the pre-conventional, conventional and post-conventional considerations were 2.7, 2.8 and 2.8 respectively. Table 4 shows the average ratings for each consideration.

Table 4

Average ratings for each consideration.

Dilemma 1					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	3.0	3.0	2.8	2.6	
Conventional	3.0	2.8	2.3	2.8	3.0
Post-Conventional	3.0	3.0	2.8	2.8	

Dilemma 2					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	3.0	3.0	3.0	2.8	
Conventional	3.0	3.0	2.5	2.8	2.5
Post-Conventional	3.0	3.0	3.0	3.0	

Dilemma 3					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	3.0	3.0	3.0	2.4	
Conventional	2.8	3.0	3.0	2.8	2.4
Post-Conventional	3.0	2.4	3.0	3.0	2.5

Dilemma 4					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	2.5	2.8	2.8	3.0	2.3
Conventional	3.0	3.0	3.0	3.0	2.3
Post-Conventional	2.5	3.0	2.8	3.0	3.0

Dilemma 5					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	2.6	3.0	1.8	3.0	
Conventional	3.0	3.0	2.5	3.0	
Post-Conventional	3.0	3.0	3.0	2.5	

Dilemma 6					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	2.2	2.0	2.8	3.0	
Conventional	3.0	2.8	2.8	3.0	
Post-Conventional	3.0	2.8	2.8	2.3	

Dilemma 7					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	2.0	3.0	2.8	2.8	2.8
Conventional	2.3	3.0	2.8	2.8	
Post-Conventional	3.0	3.0	3.0	2.8	

Dilemma 8					
Type of Consideration	Consideration # 1	2	3	4	5
Pre-Conventional	3.0	2.4	2.6	2.2	2.4
Conventional	3.0	3.0	2.8	2.6	2.6
Post-Conventional	3.0	2.3	2.3	2.8	2.8

The initial instrument that was sent to content experts contained a total of 107 considerations, with each dilemma containing 15 to 17 considerations. The EERI and DIT2, which the EPSRI is modeled after, have 12 considerations per dilemma (Rest et al., 1999a; Zhu et al., 2014). Similar to the dilemmas, a surplus of considerations were created with the anticipation that some considerations would be eliminated during the validation process.

An example of a consideration that was revised according to feedback was one of the conventional considerations from the fifth dilemma. This consideration asked participants how an isobutene leak would affect the company's image. Content experts mentioned that an isobutene leak would not affect the company's image, giving the example of how BP has had major oil spills, yet continues to do business in the US. This consideration was revised to reflect this feedback, instead asking participants how a hazardous leak might impact the company's image.

An example of a consideration that did not implement suggested changes was a post-conventional consideration from the third dilemma. This consideration asked the participant what their level of desire was to continue to work for a company that doesn't follow protocol correctly. Content experts felt that this prompt was too personal. Researchers felt that the personal aspect of this consideration did not alter the fact that it was still post-conventional, so the consideration was not revised.

An example of a consideration that was eliminated was a pre-conventional consideration from the seventh dilemma. This consideration asked the participant how important it was to follow the instructions of their co-worker. In regards to the dilemma, the content experts found the consideration to lack a connection to consequences or rewards. This consideration was given an average rating of two, and was accompanied by four other pre-conventional considerations. Taking into account all of these reasons, it was decided that the consideration should be eliminated from the instrument.

A total of six pre-conventional considerations were eliminated from five different dilemmas. One pre-conventional consideration was eliminated from dilemma three due to its low rating, however, it was replaced by another pre-conventional consideration that was suggested by one of the content experts. Two pre-conventional considerations were eliminated from the fourth dilemma. One of the pre-conventional considerations was replaced with a conventional consideration that content experts believed related more to the pre-conventional definition. One pre-conventional consideration was eliminated from the fifth dilemma due to its low rating. One pre-conventional consideration was eliminated from the sixth dilemma, but was replaced with a suggested consideration that one of the content experts included in their feedback. One pre-conventional consideration was eliminated from the seventh dilemma due to low ratings and negative feedback.

A total of three conventional considerations were eliminated from three different dilemmas. One conventional consideration was eliminated from the first dilemma due to its low rating, and the sufficient amount of conventional considerations for that dilemma. Put differently, there were originally five conventional considerations for this dilemma, so eliminating one of the conventional considerations was fair. One conventional consideration was eliminated from dilemma three due to its low rating and sufficient amount of conventional considerations for that dilemma. One conventional consideration was eliminated from the seventh dilemma due to its low rating and negative feedback.

A total of two post-conventional considerations were eliminated from two different dilemmas. One post-conventional consideration was eliminated from dilemma three due to its low rating and sufficient amount of post-conventional considerations for that dilemma. One post-conventional consideration was eliminated from the sixth dilemma due to its low rating and negative feedback from content experts.

The table below shows the number of pre-conventional, conventional, and post-conventional considerations that were eliminated, revised, or kept per dilemma. Revisions to the considerations were classified as either minor or major revisions. A minor revision entails re-phrasing, or word changes. A major revision means that the consideration was either replaced with a suggested consideration, or was changed to a different form of thinking.

Table 5

Summary table showing results of consideration review.

		Eliminated	Minor Revision	Major Revision	No Change
Dilemma 1	Pre-Conv	-	-	-	5
	Conv	1	-	-	4
	Post-Conv	-	-	-	4
Dilemma 2	Pre-Conv	-	-	-	4
	Conv	-	-	-	5
	Post-Conv	-	-	-	4
Dilemma 3	Pre-Conv	1	-	1	3
	Conv	1	-	-	4
	Post-Conv	1	-	-	4
Dilemma 4	Pre-Conv	2	-	1	3
	Conv	-	1	-	5
	Post-Conv	-	-	-	5
Dilemma 5	Pre-Conv	1	1	-	3
	Conv	-	1	-	4
	Post-Conv	-	1	-	4
Dilemma 6	Pre-Conv	1	1	1	3
	Conv	-	-	-	4
	Post-Conv	1	-	-	3
Dilemma 7	Pre-Conv	1	-	-	4
	Conv	1	-	-	3
	Post-Conv	-	-	-	4

Overall, eleven considerations were eliminated, five had minor revisions, and three had major revisions.

Limitations. The largest limitation of this study was the small sample size of content experts that were able to validate the EPSRI. There were twelve content experts who participated in the study, however, each of the content experts were not able to review every dilemma and consideration. Due to the length of the instrument, two surveys were created that each contained half of the instrument. The first survey contained dilemmas four through seven, with their accompanying considerations, and was reviewed by five content experts. The second survey contained dilemmas one through three and eight, with their accompanying considerations, and was reviewed by seven content experts. While the sample size of content experts who participated in the study was low, we believe that they are a sufficient representation of the chemical engineering education population.

Conclusions

The objective of this study was to validate the content within the EPSRI to ensure the dilemmas were relevant, and the considerations related back to their perceived definitions. In order to accomplish this, researchers sought out professionals in the chemical industry, engineering education or learning science fields, as well as chemical engineering faculty to determine the relevance of the dilemmas. The professionals also reviewed the considerations to ensure they represented the applicable definitions and suggest any items that might have been omitted.

Dilemmas with an average rating below a 2.5 were eliminated, and dilemmas with an average rating above a 2.5 were kept and revised if all researchers believed that the corresponding feedback helped improve or clarify the dilemma. One dilemma was eliminated, and four were revised.

Considerations with an average rating below a 2 were eliminated. Considerations with an average rating between a 2 and 2.5 were eliminated if there was a sufficient amount of considerations at that level of thinking, or kept if there was feedback that improved the consideration. Considerations above a 2.5 were kept and revised if the feedback provided helped to improve or clarify the consideration. Eleven considerations were eliminated, five considerations underwent minor revisions, and three underwent major revisions.

The initial EPSRI contained eight dilemmas with 15 to 17 considerations per dilemma. Other instruments such as the EERI and DIT2 (Rest et al., 1999a; Zhu et al., 2014), which the EPSRI is modeled upon, only contain five dilemmas, with each dilemma containing twelve considerations. A surplus of dilemmas and consideration were created in anticipation of poor relevance and alignment, and with the knowledge that the EPSRI will eventually contain five dilemmas with 12 accompanying considerations. Following the content validation, one dilemma and eleven considerations were eliminated from the instrument leaving the second version of the instrument with a total of 7 dilemmas and 13-15 considerations per dilemma. A revised version of Table 1 is shown below with the updated number of considerations per dilemma. The item numbers of the considerations are shown in the brackets.

Table 6

Considerations per dilemma after content validation

	Pre-conventional	Conventional	Post-conventional	M-items	Total
Dilemma 1	4 [1, 2, 3, 5]	4 [4, 6, 7, 13]	4 [8, 9, 10, 14]	2 [11, 12]	14
Dilemma 2	4 [2, 3, 9, 14]	4 [4, 6, 8, 10]	4 [1, 5, 12, 15]	3 [7, 11, 13]	15
Dilemma 3	4 [1, 5, 7, 13]	4 [2, 6, 10, 11]	4 [3, 8, 9, 12]	1 [4]	13
Dilemma 4	4 [4, 10, 11, 12]	4 [1, 5, 8, 9]	4 [2, 6, 7, 13]	1 [3]	13
Dilemma 5	3 [2, 3, 12]	4 [6, 8, 11, 15]	4 [1, 7, 13, 14]	4 [4, 5, 9, 10]	15
Dilemma 6	4 [5, 6, 8, 12]	4 [1, 2, 4, 9]	4 [3, 7, 13, 15]	3 [10, 11, 14]	15
Dilemma 7	4 [5, 8, 12, 14]	3 [3, 6, 10]	4 [1, 4, 9, 13]	3 [2, 7, 11]	14

Chapter 4

Large Scale Validation Study

Overview

The objective of this study was to conduct a factor analysis to validate the EPSRI based on data obtained from a large scale study. The factor analysis provides the opportunity to analyze the strength of the correlations between the items across the dilemma and within each moral schema, and to determine the number of underlying latent variables. Senior chemical engineering students in their capstone classes across three institutions participated in the study. Rose Hulman Institute of Technology had 49 responses, University of Connecticut had 79 responses, and North Carolina State University had 109 responses for a total of 237 student responses.

Before the factor analysis was conducted, the data underwent a bogus data analysis that identified unreliable responses through three tests; rate-rank score, missing data, and repeating data (Rest et al., 1999a). The bogus data analysis eliminated 14 data sets from the protocol, which resulted in 223 student responses being included in the factor analysis.

The factor analysis was completed on two data sets: one with the complete data set, and one that consisted of all data after the bogus data was eliminated. Eliminating the bogus data ensured the protocol was robust, and that only reliable data was included in the protocol. However, results from both analysis were found to be similar. From the factor analysis, one dilemma and seven considerations were eliminated, and 22 considerations were revised. The instrument was not yet able to be validated at the conclusion of this study. Moving forward, the instrument will be implemented again in the fall of 2018, and will hopefully lead to the validation of the instrument.

Introduction

Instrument validation process. So far, the Engineering Process Safety Research Instrument (EPSRI) has undergone a content validation, which ensured the relevancy of the dilemmas, and the alignment of the considerations with their perceived definitions. Content experts from chemical industry and engineering education/learning science fields, and chemical engineering faculty reviewed the instrument as part of the validation process. The initial version of the EPSRI contained eight dilemmas which each contained 15-17 considerations. Following the content validation, one dilemma and eleven considerations were eliminated, and four dilemmas and seven considerations were revised. At the end of this first step in the validation process the EPSRI had a total of 7 dilemmas and 99 considerations.

The next step of the validation process is to perform a factor and reliability analysis, which is described in this chapter. The factor analysis allowed for the relationships between the items to be determined, and the correlations between the considerations and the moral schemas (pre-conventional, conventional and post-conventional) to be analyzed.

Overview of methods. ***Bogus data analysis.*** Prior to conducting the factor analysis, unreliable or bogus data was identified to ensure the results obtained from the factor analysis were derived from reliable and consistent data (Rest et al., 1999a). The bogus data analysis was completed through three analyses; rate-rank score, missing data, and repeating data. The rate-rank score identified random responses by comparing the items that were ranked with their respective ratings (Rest et al., 1999a). The missing data identified unreliable data sets based on how many responses were omitted (Rest et al., 1999a). The repeating data identified unreliable data sets based on how many considerations were rated the same within a dilemma (Rest et al., 1999a). Following the bogus data analysis, 14 data sets were purged from the protocol.

***Factor analysis and reliability*.** Prior to the factor extraction, appropriateness of data testing was conducted on the data to determine which considerations to eliminate from a dilemma based on its poor correlation with the other considerations, and which dilemmas may need to be eliminated entirely. Following this process, the factor extraction took place to determine the number of underlying latent variables within a dilemma. Ideally, each dilemma should contain four underlying latent variables to represent the pre-conventional, conventional, post-conventional considerations and meaningless items (M-items). The factor extraction also determined which considerations were loading together. Ideally, each schema's considerations and M-items should load together on their respective factors. Following the factor extraction, reliability testing was completed on the items that loaded together to determine the strength of correlations between the items, and which items to eliminate to strengthen the correlations. If an item loaded onto more than one factor, the reliability testing could also assist in determining which factor the item would load onto best although review of the intended loading of the consideration with its desired latent variable was also a consideration.

Overview of results. The factor analysis was conducted on the data set containing the bogus data and the data set without the bogus data. Most items were loading similarly between the two trials, and the conclusions made from both analysis were the same indicating the same considerations for elimination or revision. Based on this observation, it was decided that to move forward with just the analysis using the final data set after bogus data set removal. Following the factor analysis, one dilemma and seven considerations were eliminated. Considerations that did not load with other considerations of its schema were revised to behave more similarly to the other considerations in the schema. The instrument was not able to be validated at this point during the study, due to some of the considerations not yet loading on their intended latent variables, and lower reliability scores than suggested for instrument validation. The instrument will be implemented again during the Fall 2018 semester to senior chemical engineering students to acquire more data for the instrument validation process.

Methods

Data collection. The EPSRI was implemented at three universities, Rose Hulman Institute of Technology, the University of Connecticut, and North Carolina State University. It was administered to senior chemical engineering students in their senior capstone class. A total of 237 students participated in the study; 49 students were from Rose Hulman Institute of Technology, 79 students were from University of Connecticut and 109 students were from North Carolina State University. Student responses were electronically recorded through Qualtrics.

Students were first prompted to read the dilemma, then choose one of the three options about their course of action before moving onto the considerations. Students would read considerations, then rate them on a scale from one (none) to five (great) based on how much the consideration played a role into their thinking. Once students were finished rating considerations, they ranked the top four considerations from most to least important.

Bogus data analysis. Reliability checks were administered to the data sets to ensure that unreliable, or bogus, data sets were eliminated from the protocol. Rest et al. (1999a) created a set of checks for bogus data for the Defining Issues Test, which was revised for the DIT2. The EPSRI utilized modified versions of the bogus data checks from the DIT2 due to the fact that the EPSRI follows the structure of the DIT2. Rest et al. (1999a) identified four checks to test for unreliable data: random responding, missing data, alien test-taking sets, and nondiscrimination of items.

Random responding occurs when a participant responds in a way that does not reflect their moral cognition (Rest, 1999a). Missing data occurs when a participant omits large sections of responses, or quits the protocol part way through (Rest, 1999a). Alien test-taking sets occur when a participant chooses an answer based on high level syntax or wording (Rest, 1999a). Rest et al. (1999a) created a set of meaningless items, or M-items, in order to identify alien test-taking sets. The EPSRI also contains meaningless items (M-items), however, these items were not validated during the initial content validation study. Therefore, alien test-taking sets were not included in data checks for the EPSRI. Nondiscrimination of items occurs when a participant rates or ranks most of the items the same (Rest et al., 1999a). The following sections will provide more detailed descriptions of each data check along with an example of identifying the bogus data.

Random responding. Items that are ranked should be consistent with how they were rated. In other words, no item should be rated higher than the number one ranked item. The only item that should be rated higher than the second ranked item is the first item. The first and second ranked items should be the only two items rated higher than the third ranked item. Only the first, second and third ranked item should have a higher rating than the fourth ranked item. An item is considered an inconsistency if it fails to follow this approach. In order to account for these inconsistencies, Rest et al. (1999a) derived a rate-rank score.

The rate-rank score is able to determine the consistency of a data set by utilizing multipliers and summing the amount of inconsistencies. The amount of inconsistencies rated above the first ranked item is given a multiplier of four. The amount of inconsistencies rated above the second ranked item is given a multiplier of three. Inconsistencies rated above the third ranked item are multiplied by two and inconsistencies rated above the fourth ranked item are multiplied by one. These values were assigned by Rest et al. (1999a), and adapted for this study. The vales are summed across the dilemmas to obtain the rate-rank score. An example from the data set is given below to demonstrate the calculation of the rate-rank score. Tables 7 and 8 below show the ratings and rankings of items.

Table 7

Ratings of considerations from dilemma 3

Consideration	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***13***
Rating	4	3	4	3	5	5	4	4	3	2	4	4	3

Table 8

Ranking of considerations from dilemma 3

Rank	***First***	***Second***	***Third***	***Fourth***
Consideration Number	7	6	8	11

This student ranked consideration seven as the most important item. Table 7 shows that consideration seven was given a rating of 4, however, items five and six were rated 5. These two inconsistencies are multiplied by four for a total of eight added to the rate-rank score. Consideration six was ranked as second most important and was given a rating of 5. There are no considerations rated above a 5, so there are no inconsistencies at this level. Consideration eight was ranked third most important, and was given a rating of 4. Consideration five and six were rated higher than consideration eight, however consideration six was ranked higher than the third ranked item. For this reason, the one inconsistency is multiplied by two, and added to the rate-rank score for a cumulative total of ten. Consideration eleven was ranked fourth most important and was given a rating of 4. Consideration five was given a higher rating, but was not ranked above consideration eleven. The inconsistency multiplied by one and added to the rate rank score for a total of eleven. The calculation for this score is shown below.

$$4 \times 2 + 3 \times 0 + 2 \times 1 + 1 \times 1 = 11$$

In order to determine which data sets should be omitted from the protocol, a cut off value was obtained. Rest et al. (1999a) used a cut off value of 200 for the DIT2. Based on the five dilemmas that were each accompanied by 12 considerations, the rate-rank score for the DIT2 could range from 0-600. At this point of the study, the EPSRI contained seven dilemmas and were accompanied by 13-15 considerations. In order to stay consistent with the DIT2, the cut off value for the EPSRI would be one third of the maximum rate-rank score. The maximum rate-rank score for the EPSRI was calculated to be 920. The cut off value was determined to be 306.67, or 307.

Missing data. Rest et al. (1999a) determined that a data set was inconsistent if three or more ratings were omitted from two dilemmas or more, or if six or more ratings were missing from the protocol overall. These methods were adopted for the EPSRI with no modifications. If a student omitted two or more ratings from three or more dilemmas, or at least six ratings overall, the data set was eliminated from the protocol.

Repeating data. Rest et al. (1999a) determined that a data set was inconsistent if 11 or 12 items were rated the same within one dilemma. If this occurred in more than one dilemma, the data set was purged from the protocol (Rest et al., 1999a). These methods were slightly modified for the number of considerations on the EPSRI. All dilemmas on the DIT2 contained 12 considerations compared to the EPSRI which contains 13-15 considerations per dilemma. In order to remain consistent with the DIT2, a data set was considered inconsistent if all of the considerations, or if one less than all of the considerations, were rated the same. For example, if a dilemma contained 15 considerations, it was considered inconsistent if 14 or 15 considerations were rated the same. If this occurred more than once across the seven dilemmas, the data set was eliminated from the protocol.

Results and Discussion

Rationale for methods. Factor analysis was completed on each dilemma, however, it was not completed on the instrument as a whole. This was due to the number of student responses that were received compared to the number of items present on the EPSRI. Hair et al. (1995) expressed that the minimum number of responses necessary for factor analysis is five responses for every item on the instrument, and a more acceptable range being a 10:1 ratio of student responses to items on the instrument. The EPSRI contains 99 items, which means over 1000 student responses would be needed to satisfy the 10:1 ratio. Since each dilemma contains a maximum of 15 items, a minimum of 150 student responses were needed to satisfy the 10:1 ratio. The amount of student responses that were received exceeded this minimum amount, proving a sufficient sample size.

Prior to completing the factor analysis, the data was tested for appropriateness following Pett et al. (2011). Data underwent the Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) test for sampling to determine if the results were statistically significant and if the sample size was sufficient. The Measure of Sampling Adequacy testing was also completed as a part of this process to ensure the items were well correlated. This process is explained in detail in the following sections. This process was chosen based on its implementation in social science research (Pett, Lackey, and Sullivan, 2011).

Factor analysis was completed using a principal component analysis with an oblique rotation approach. The goal of principal component analysis is to summarize the interrelationships among a set of original variables in terms of a smaller set of uncorrelated principal components that are linear combinations of the original values (Pett et al., 2011). The analysis assumes a large amount of variance can be explained by the extracted factors. It is useful when wanting to summarize a large number of variables into a small number of components. Each dilemma in the EPSRI contains 13-15 considerations that should fall into one of four categories. For that reason, principal component analysis was chosen for factor extraction.

Oblique rotation predicts that the underlying latent variables somewhat correlate with one another and that items can be classified with respect to a single category (Devellis, 2012). The theory behind the EPSRI is based on the development of moral reasoning, which insinuates that the levels of moral reasoning are believed to have relationships between one another (Kohlberg and Hersh, 1977). Within the theory, Kohlberg defined the stages of moral reasoning as hierarchical integration, which means that an individual operating at a higher level of moral reasoning will still understand and reason through lower level reasoning (Kohlberg and Hersh, 1977). While there may be overlap between pre-conventional, conventional and post-conventional reasoning, the items should still load independently, based on the schema they are meant to represent.

Oblique rotation was appropriate to use because it accounts for the underlying relationship while still separating proper items into their respective latent variables.

If successful results are obtained from this study, the EPSRI will be validated in its current form for senior chemical engineering students. If any changes are made to the EPSRI, or if the EPSRI is used by participants who are not senior chemical engineering students, the reliability of the constructs would need to be re-evaluated to ensure they are measuring the intended variables. Validation of the EPSRI would lead to validated results or measures from the students which have meaning (Briggs and Cheek, 1986) Additionally, validating the EPSRI would entail that the items created to represent pre-conventional, conventional and post-conventional reasoning were well reflective, and had been grouped appropriately (DeVellis, 2012).

Factor analysis procedure. Each dilemma underwent a test for appropriateness of data, factor extraction, and reliability testing to analyze the strength of the data for each dilemma and consideration. Testing for appropriateness of data was completed by first analyzing the correlation between items and ensuring that the absolute value did not fall below 0.001. The determinant of this matrix was then evaluated to ensure it fell between zero and one. KMO and Bartlett's test of sphericity were analyzed to ensure the results were statistically significant, and that the sample size was sufficient (Pett et al., 2011). MSA values were also analyzed to ensure the considerations were correlated well enough to proceed with the factor extraction (Pett et al., 2011). Considerations that obtained an MSA value below 0.6 were eliminated, and the analysis was run again. Factor extraction was completed by the principal component analysis with the direct oblim approach. Ideally, each dilemma should have four factors that contained pre-conventional, conventional, post-conventional and M-items. However, the dilemmas were not forced to have four factors as would have occurred when doing a confirmatory factor analysis. Factor loadings under 0.4 were suppressed from the structure matrix.

Once the factor extraction was completed, the reliability analysis was done to determine the strength of the items on each factor, the best placement for items that loaded onto more than one factor, and which items to eliminate. The reliability analysis tested for Cronbach's alpha and inter-item correlations. A Cronbach's alpha value over 0.7 indicated strong reliability for the factor (Tavakol and Dennick, 2011). Inter-item correlations determined if items were too similar, and were expected to stay below 0.8. Following the reliability analysis, suggestions were made about which considerations to revise or eliminate. It is beneficial to revise items instead of eliminating the ones that didn't load properly for two reasons. First, each dilemma should contain a minimum of 12 items. By eliminating the items that didn't load properly, some dilemmas may fall below the 12 consideration benchmark. Second, the factor extraction would have to be re-run after the elimination of the items in order to properly validate the instrument. Items are not guaranteed to load the same way they had before the items had been eliminated.

Dilemma 1. ***Appropriateness of data.*** The initial appropriateness of data test revealed a KMO of 0.691, and all items except for item five had an MSA above 0.6. Item five, which had an MSA value of 0.568, was eliminated and the process was run again. The secondary run yielded a KMO of 0.696, and all items except item eight had an MSA above 0.6. Item eight, which had an MSA value of 0.556, was eliminated, and the process was run again. The final run yielded a KMO of 0.697, and all items had an MSA above 0.6. These successful results allowed for the process to move to factor extraction.

Factor extraction. Factor extraction was completed by using the principal component analysis approach and an oblique rotation with delta set to zero. Considerations with factor loadings below 0.4 were considered weak and were suppressed from the structure matrix in order to determine where the considerations were most strongly loaded. All considerations strongly loaded onto at least one factor. Items 2, 6 and 14 loaded onto more than one factor during the factor extraction. Table 9 shows the results from the factor extraction for dilemma one.

Table 9

Dilemma one considerations and factor loadings.

	Factor		
Consideration Number and Prompt	**1**	**2**	**3**
2. Do you really want to put more work and risk on your employees by requiring them to replace the Option A hoses each month? (Conv)	0.617*	~~-0.488~~	
6. How much risk is associated with replacement of hoses for each option? (Post-conv)	0.421	~~-0.581~~	
9. Do you believe there is a level of risk that is acceptable considering the savings of Option A? (Post-conv)	0.703		
10. Is it ever a good idea to rely on active measures (in this case, employee maintenance) rather than passive measures (in this case, material of construction)? (Post-conv)	0.727		
13. Would choosing Option A benefit your company and co-workers in other ways besides saving money? (Conv)	0.769*		
7. Would product quality be affected by degradation products from the hose being present during the plastics manufacturing process? (Conv)	0.406*		

Table 9 (continued)

	Factor		
Consideration Number and Prompt	**1**	**2**	**3**
3. Would you gain personal satisfaction of "doing a good job" if you choose one option over the other? (Pre-conv)		**0.722**	
11. Would your fellow employees have a more positive opinion of you if you chose one option or the other? (Pre-conv)		0.774	
1. What would you do with the money you could receive as a bonus? (Pre-conv)			0.715
14. On top of the bonus, what kind of further recognition and opportunities for career advancement could you receive by keeping costs low? (Pre-conv)	~~0.463~~		0.577
4. Do you think that management would prefer the name of one option more than the other? (M-item)			0.689
12. Do the polymer linings for each option come in different colors? (M-item)			0.702

Loadings with a strikethrough indicate an item that was double loaded and was removed from one of the factors. Items with an asterisk were not well correlated, and were revised following this study.

Factor 1 – Post-conventional. Factor one contained item 14 which was pre-conventional, items 2, 7 and 13 which were conventional, and items 6, 9 and 10 which were post-conventional. Item 14 loaded onto two factors, which allowed for the removal of this item from the first factor. All the post-conventional items loaded onto the first factor, compared to the conventional items which loaded across the factors. The correlation of the post-conventional items had a Cronbach's alpha of 0.568. Inter-item correlations remained below 0.8, which determined that there were no redundant items. The low alpha value is indicative of a weak correlation (Tavakol and Dennick, 2011). In order to make this factor truly post-conventional, revisions will be made to the items moving forward. The conventional items will be revised to behave less like the post-conventional items. Currently, the students are not identifying the difference between the conventional and post-conventional items, which resulted in them loading onto the first factor together. Ideally, there should be four factors that contain either pre-conventional, conventional, post-conventional considerations, or M-items. Revising the conventional items should result in them unloading from the post-conventional factor and moving them onto their own new factor. Item 14 will also be revised to behave more like items 3 and 11, which will result in it unloading from the first factor.

Factor 2 – Pre-conventional. Factor two contained items 3 and 11, which were pre-conventional, item 2 which was conventional, and item 6 which was post-conventional. Items 2 and 6 loaded onto more than one factor, which allowed for the removal of these items from the second factor. The remaining items on factor were pre-conventional. The correlation of the items was determined by Cronbach's alpha, which was 0.658. Inter-item correlations remained below 0.8, which determined that there were no redundancies. The alpha value falls below 0.7, so the correlation cannot be considered as strong. In order for all of the pre-conventional items to load on the second factor, and to strengthen the correlation between the pre-conventional items, items 1 and 14 will be revised to behave more like items 3 and 11.

Factor 3 – M-items. Factor three contained items 1 and 14 which were pre-conventional, and items 4 and 12 which were M-items. Since it is anticipated that items 1 and 14 will be revised and unload from this factor, the third factor will contain the M-items. The correlation between these items was determined from Cronbach's alpha, which was 0.417. Inter-item correlations remained below 0.8, which insinuated that there were no redundancies. The alpha value is indicative of a weak relationship between the M-items. However, Cronbach's alpha value increases with the number of items being measured, so it is not anticipated that there will be a high value between two items (Tavakol and Dennick, 2011).

Summary. In order to obtain stronger correlations, and correct loading of the items according to their schema, proposed revisions will be made. Moving forward, the conventional items will be revised to unload from the post-conventional factor onto their own factor. Items 1 and 14 will be revised to behave more like item 3 and 11, which will result in factor 2 containing the pre-conventional items. Table 10 shows the factor extraction results for dilemma one, as well as the results that are expected moving forward based on the proposed revisions.

Table 10

Factor extraction results and desired results for dilemma one.

	Name	Extraction Results	Desired Results	Revised
Factor 1	Post-conventional	2, 6, 7, 9, 10, 13	6, 9, 10	None
Factor 2	Pre-conventional	3, 11	1, 3, 11, 14	1, 14
Factor 3	M-items	1, 4, 12, 14	4, 12	None
Factor 4	Conventional		2, 7, 13	2, 7, 13

Dilemma 2. ***Appropriateness of data.*** The appropriateness of data test revealed that all items had an MSA above 0.6. None of the items were removed during this step of the process. The analysis gave a KMO value of 0.742, which is above the 0.6 threshold (Pett et al., 2011). These successful results allowed for the process to move forward with factor extraction.

Factor extraction. Factor extraction was completed by using the principal component analysis approach and an oblique rotation with delta set to zero. Considerations with factor loadings below 0.4 were considered weak and were suppressed from the structure matrix in order to determine where the considerations were most strongly loaded. All of the items strongly loaded onto at least one factor. Items 3 and 5 loaded onto more than one factor. Table 11 below shows the results from the factor extraction on dilemma two.

Table 11

Dilemma two considerations and factor loadings.

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
1. What is the potential for negative impact to the environment if the tanks release their contents? (Post-conv)	0.786			
4. Is there the potential for the exploding tanks to damage the surrounding neighborhood and infrastructure adjacent to the plant? (Post-conv)	0.725			
12. What is the potential for negative health effects on residents who live in the areas surrounding the plant if the tanks release their contents? (Post-conv)	0.839			
2. What would be the impact on your own job if the plant is damaged or destroyed due to explosion caused by failure of the cooling loops? (Pre-conv)		0.595*		
10. How much would negative press impact the company if an explosion occurs? (Conv)		0.718*		

Table 11 (continued)

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
13. What is the potential for wind damage to the plant? (M-item)		**0.604**		
14. Does staying and working at the plant during the storm make it difficult for you to secure your personal belongings from damage? (Pre-conv)		0.687*		
15. What responsibility does your company have to locate its facilities in areas where negative impacts to the surrounding community are minimized? (Post-conv)		0.645*		
5. What is your level of comfort in soliciting volunteers to stay on-site during what may be a life-threatening situation? (Conv)			0.613	~~0.492~~
7. How confident are you in the accuracy of the predictions of the storm's impact on the plant? (M-item)			0.637*	

Table 11 (continued)

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
8. What is your confidence that you and/or your teams will be able to keep the generators functioning under the storm conditions? (Conv)			0.778	
11. How high the floodwaters are predicted to get? (M-item)			**0.718***	
3. What is the possibility that your colleagues could be injured or killed if they stay at the plant during the hurricane? (Conv)	~~0.502~~			0.653*
6. What is your level of concern regarding your own personal safety if you choose to stay on-site during the storm? (Pre-conv)				0.81
9. Would staying to prepare the tanks jeopardize you and your family's safety in the upcoming storm (Pre-conv)				0.638

Loadings with a strikethrough indicate an item that was double loaded and was removed from one of the factors. Items with an asterisk were not well correlated, and were revised following this study.

Factor 1 – Post-conventional. Factor one contained items 1, 4 and 12 which were post-conventional, and item 3 which was conventional. Item 3 had loaded onto multiple factors and was able to be removed from the first factor. This resulted in factor one containing only post-conventional items. The Cronbach's alpha for the post-conventional items on factor one was 0.756. Inter-item correlations remained below 0.8, which indicates there were no redundancies. Due to the strong correlation between the items, and the post-conventional items loading together, no revisions were made (Tavakol and Dennick, 2011).

Factor 2 – M-items. Factor two contained items 2 and 14, which were pre-conventional, 10, which was conventional, 15, which was post-conventional and 13 which was an M-item. Pre-conventional, conventional and post-conventional items were strongly loaded on the remaining factors, which allowed for the second factor to contain the M-items. In order for factor two to only contain M-items, multiple revisions will be made. Items 2 and 4 will be revised to behave more like items 6 and 9 which will unload them from this factor. Item 10 will be revised to behave more like items 5 and 8, and unload them from this factor. Item 15 will be revised to behave more like items 1, 4 and 12 on the first factor. Since there is only one item that remains, a reliability analysis was unable to be completed.

Factor 3 – Conventional. Factor three contained items 5 and 8 which were conventional, and items 7 and 11 which were M-items. Conventional items loaded onto all the factors, but loaded best onto factor three. The correlation between items 5 and 8 was measured using Cronbach's alpha, and obtained a value of 0.523. Inter-item correlations remained below 0.8, which indicates there were no redundancies. In order to make this factor more conventional, and strengthen the correlation of the conventional items, revisions will be made. Items 7 and 11 will be revised to behave more like item 13, which should load onto the second factor. Items 3 and 10 will be revised to behave more like items 5 and 8, which will result in all the conventional items loading together.

Factor 4 – Pre-conventional. Factor four contained items 6 and 9, which were pre-conventional, and items 3 and 5 which were conventional. Item 5 loaded onto multiple factors, which allowed for the removal of this item from the fourth factor. Moving forward, item 3 will be revised and should unload from this factor. The correlation between items 6 and 9 was measured using Cronbach's alpha, and obtained a value of 0.460. This low alpha could be a result of the low number of items being tested. Moving forward, the alpha should increase when all the pre-conventional items load together. Inter-item correlations remained below 0.8, which indicated that there were no redundant items. In order for this factor to become truly pre-conventional, and to strengthen the correlations between the pre-conventional items, items 2 and 14 will be revised to behave more like items 6 and 9.

Summary. In order to obtain stronger correlations, and correct loading of the items according to their schema, proposed revisions will be made. Moving forward, items 2 and 14 will be revised to behave more like items 6 and 9, which will result in the pre-conventional items loading together. Items 3 and 10 will be revised to behave more like items 5 and 8, in order for the conventional items to load together. Item 15 will be revised to behave more like items 1, 4, and 12 which will result in the post-conventional items loading together. Items 7 and 11 will be revised to behave more like item 13, in order for the M-items to load together. Table 12 shows the factor extraction results for dilemma one, as well as the results that are expected moving forward based on the proposed revisions.

Table 12

Factor extraction results and desired results for dilemma two

	Name	Extraction Results	Desired Results	Revisions
Factor 1	Post-conventional	1, 4, 12	1, 4, 12, 15	15
Factor 2	M-items	2, 10, 13, 14, 15	7, 11, 13	7, 11
Factor 3	Conventional	5, 7, 8, 11	3, 5, 8, 10	3, 10
Factor 4	Pre-conventional	3, 6, 9	2, 6, 9, 14	2, 14

Dilemma 3. ***Appropriateness of data.*** The appropriateness of data test revealed that item 11 had an MSA below 0.6. This item, which had an MSA value of 0.507, was eliminated and the process was run again which resulted in successful results. The secondary analysis gave a KMO value of 0.785, and all remaining items having an MSA value above 0.6. These successful results allowed for the process to move forward with factor extraction.

Factor extraction. Factor extraction was completed by using the principal component analysis approach and an oblique rotation with delta set to zero. Considerations with factor loadings below 0.4 were considered weak and were suppressed from the structure matrix in order to determine where the considerations were most strongly loaded. All of the items strongly loaded onto at least one factor. Items 5, 7, and 13 loaded onto more than one factor. Table 13 shows the results from the factor extraction on dilemma three.

Table 13

Dilemma three considerations and factor loadings.

	Factor			
Consideration number and prompt	1	2	3	4
1. What is the possibility of your continued employment at the company? (Pre-conv)	0.785*			
2. Is it possible that your co-workers might lose their jobs when you file the report? (Conv)	0.754			
5. What is your manager's opinion of you? (Pre-conv)	0.633*			~~-0.61~~
6. What would be the negative impact on your family or dependents if you lose your job? (Conv)	0.811			
13. Can you avoid being placed in the same position as the previous engineer who was put under pressure while preparing the report? (Pre-conv)	0.567	~~0.408~~		
9. Is it ever okay to purposefully misrepresent data? (Post-conv)		0.879		

Table 13 (continued)

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
12. What is your desire to continue to work for an employer who doesn't follow protocol correctly? (Post-conv)		**0.828**		
3. Who would be the most impacted by the spills? (Post-conv)		0.733		
10. Is it your duty to change the report if it is for the good of your company? (Conv)		0.588*		
7. What if the next chemical spill has an impact on you personally? (Pre-conv)			0.690	~~-0.451~~
8. What is the likelihood of another, more serious chemical spill if the data is presented inaccurately? (Post-conv)			0.778*	
4. How long has Pam worked for the company? (M-item)				-0.872

Loadings with a strikethrough indicate an item that was double loaded and was removed from one of the factors. Items with an asterisk were not well correlated, and were revised following this study.

Factor 1 – Conventional. Factor three contained items 1, 5 and 13 which were pre-conventional, and items 2 and 6 which were conventional. Although factor one contained more pre-conventional items, most of the conventional items had loaded onto factor one. Additionally, the pre-conventional items loaded across the factors, and the conventional items only loaded onto two factors. The correlation between items 2 and 6 was measured using Cronbach's alpha, and obtained a value of 0.628. Inter-item correlations remained below 0.8, which indicates there were no redundancies. In order for factor one to contain only conventional items, revisions will have to be made. Items 1 and 5 will be revised to behave more like item 7, which will result in them unloading from the first factor. Item 13 was originally slated to be removed from the instrument at the time revisions were being made. However, eliminating item 13 resulted in 11 considerations for dilemma three, which is not sufficient. Item 13 was added back into the instrument however, it was not revised.

Factor 2 – Post-conventional. Factor two contained item 13 which was pre-conventional, item 10 which was conventional, and items 3, 9, and 12 which were post-conventional. Item 13 loaded onto multiple factors, which allowed for it to be removed from this factor. Most of the post-conventional items loaded onto factor three, and had a strong correlation. The correlation between the post-conventional items was measured using Cronbach's alpha, and had a value of 0.799. Inter-item correlations remained below 0.8, which indicates there were no redundant items. In order for this factor to only contain post-conventional items, item 10 will be revised to behave more like items 2 and 6. This will result in the conventional items loading together, as well as factor two only containing post-conventional items.

Factor 3 – Pre-conventional. Factor three contained item 7 which was pre-conventional, and item 8 which was post-conventional. Since it was previously determined that the post-conventional items remain on the second factor, item 8 will be revised to behave more like items 3, 9 and 12. This results in factor three only containing item 7. The correlation cannot be measured since there is only one item left on the factor. In order for factor three to contain the pre-conventional items, revisions to item 1 and 5 will be made so they behave more like item 7. Item 13 was not revised, but should still correlate well with the other pre-conventional items.

Factor 4 – M-items. Factor four contained items 5 and 7 which were pre-conventional and item 4, which is an M-item. Items 5 and 7 loaded onto multiple factors and were able to be removed from the fourth factor. This results in factor four only containing item 4. Since only one item remains, the correlation cannot be determined. Dilemma three only contains one M-item, which should load onto a factor by itself.

Summary. In order to obtain stronger correlations, and correct loading of the items according to their schema, proposed revisions will be made. Moving forward, items 1 and 5 will be revised to behave more like item 7 which will result in the pre-conventional items loading together. Item 13 was not revised due to it originally being eliminated, but should still load with the other pre-conventional items. Item 10 will be revised to behave more like items 2 and 6 in order for the conventional items to load together. Item 8 will be revised to behave more like items 3, 9, and 12 which will result in the post-conventional items loading together. Table 14 summarizes the results from the factor extractions, as well as the anticipated results following the revisions.

Table 14

Factor extraction results and desired results for dilemma three.

	Name	Extraction Results	Desired Results	Revisions
Factor 1	Conventional	1, 2, 5, 6, 13	2, 6, 10	10
Factor 2	Post-conventional	3, 9, 10, 12	3, 8, 9, 12	8
Factor 3	Pre-conventional	7, 8	1, 5, 7, 13	1, 5
Factor 4	M-item	4	4	None

Dilemma 4. ***Appropriateness of data.*** The appropriateness of data test revealed that item 6 had an MSA below 0.6. This item, which had an MSA value of 0.478, was eliminated and the process was run again which yielded successful results. The secondary analysis gave a KMO value of 0.780, and all remaining items having an MSA value above 0.6. These successful results allowed for the process to move forward with factor extraction.

Factor extraction. Factor extraction was completed by using the principal component analysis approach and an oblique rotation with delta set to zero. Considerations with factor loadings below 0.4 were considered weak and were suppressed from the structure matrix in order to determine where the considerations were most strongly loaded. All of the items strongly loaded onto at least one factor. Items 1, 4, 8, 10, and 11 loaded onto more than one factor. Table 15 below shows the results of the factor extraction on dilemma four.

Table 15

Dilemma four considerations and factor loadings.

	Factor			
Consideration number and prompt	**1**	**2**	**3**	**4**
1. How often is maintenance performed on the equipment in the plant? (Conv)	0.549		~~0.598~~	
3. What type of valve is leaking? (M-item)	0.780*			
5. How often is the valve used? (Conv)	0.757			
10. What other people or equipment may be exposed to the steam leak? (Conv)	0.527		~~0.682~~	
4. Will the leaking valve be a common nuisance for you, or is it located in a part of the plant you will seldom visit? (Pre-conv)	~~0.653~~	0.433		
9. Will you face negative repercussions from reporting the leak? (Pre-conv)		0.818		
11. How much time or effort would it take you to have the valve inspected? (Pre-conv)	~~0.534~~	0.538		
12. Would you be looked at as a "worrier" if you report the leak? (Pre-conv)		0.843		
7. Is it ever acceptable to not report a potential safety hazard? (Post-conv)			0.843	

Table 15 (continued)

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
13. Can any safety hazard, regardless of how minor, be dismissed as simply an "annoyance"? (Post-conv)			**0.836**	
2. Is your co-worker's comment a reflection of general engineering safety culture? (Post-conv)				0.783*
8. What is your desire to help your co-worker? (Conv)		~~0.425~~		0.621*

Loadings with a strikethrough indicate an item that was double loaded and was removed from one of the factors. Items with an asterisk were not well correlated, and were revised following this study.

Factor 1 – Conventional. Factor one contained items 4 and 11 which were pre-conventional, items 1, 5, and 10 which were conventional, and item 3 which was an M-item. Items 4 and 11 loaded onto multiple factors, which allowed for them to be removed from the first factor. Following the removal of the pre-conventional items, factor one contained most of the conventional items as well as an M-item. The correlation between the conventional items was measured using Cronbach's alpha, and obtained a value of 0.633. Inter-item correlations remained below 0.8, which indicates there were no redundant items. In order for this factor to only contain conventional items, the M-item will be revised in order for it to unload from this factor, and onto its own factor.

Factor 2 – Pre-conventional. Factor two contained items 4, 9, 11, and 12 which were pre-conventional, and item 8 which was conventional. Item 8 loaded onto multiple factors, which allowed for it to be removed from the second factor. This results in only pre-conventional items remaining on the factor. All of the pre-conventional items loaded onto factor two, and obtained a Cronbach's alpha of 0.710, which indicated a strong correlation (Tavakol and Dennick, 2011). Inter-item correlations remained below 0.8, which indicated that there were no redundant items.

Factor 3 – Post-conventional. Factor three contained items 1 and 10 which were conventional, and items 7 and 13 which were post-conventional. Items 1 and 10 loaded onto multiple factors and were able to be removed from factor three. This resulted in factor three containing two of the post-conventional items. The correlation between these items was measured using Cronbach's alpha, and was calculated to be 0.781. This was indicative of a strong relationship (Tavakol and Dennick, 2011). Inter-item correlations remained below 0.8, which specified that there were no redundant items.

Factor 4 – M-items. Factor four contained items 2 and 8, which were post-conventional and conventional respectively. Factor one contains most of the conventional items, so item 8 will be revised to behave more like items 1, 5 and 10. This will result in item 8 loading with the other conventional items. Factor three contains most of the post-conventional items, so item 2 will be revised to behave more like items 7 and 13. This will result in all of the post-conventional items loading together. This results in factor four containing zero items. However, factor one contained item 3, which was an M-item. Following revisions, item 3 should unload from the first factor onto its own factor alone.

Summary. In order to obtain stronger correlations, and correct loading of the items according to their schema, proposed revisions will be made. Moving forward, item 8 will be revised to behave more like items 1, 5 and 10, which will result in the conventional considerations loading together. Item 2 will be revised to behave more like items 7 and 13, which will result in the post-conventional considerations loading together. Item 3 will be revised to behave less like items 1, 5, and 10, which will result in it unloading from the conventional considerations and loading onto a factor by itself. Table 16 summarizes the results obtained from the factor extraction, and the anticipated results following the revision process.

Table 16

Factor extraction results and desired results for dilemma four.

	Name	**Extraction Results**	**Desired Results**	**Revisions**
Factor 1	Conventional	1, 3, 5, 10	1, 5, 8, 10	8
Factor 2	Pre-conventional	4, 9, 11, 12	4, 9, 11, 12	None
Factor 3	Post-conventional	7, 13	2, 7, 13	2
Factor 4	M-item	2, 8	3	3

Dilemma 5. *Appropriateness of data.* The appropriateness of data test revealed that item 15 had an MSA below 0.6. This item, which had an MSA of 0.578 was eliminated, and the process was run again which resulted in successful results for the initial and secondary factor extraction. The secondary analysis gave a KMO value of 0.729, and all remaining items having an MSA value above 0.6. These successful results allowed for the process to move forward with factor extraction.

Factor extraction. Factor extraction was completed by using the principal component analysis approach and an oblique rotation with delta set to zero. Considerations with factor loadings below 0.4 were considered weak and were suppressed from the structure matrix in order to determine where the considerations were most strongly loaded. All the items strongly loaded onto at least one factor. Items 5, 6, 8, and 14 loaded onto more than one factor. Table 17 below shows the results from the factor extraction on dilemma five.

Table 17

Dilemma five considerations and factor loadings.

	Factor			
Consideration number and prompt	**1**	**2**	**3**	**4**
4. How many bolts will need to be undone? (M-item)	0.746			
6. How often does the valve fail to open as expected? (M-item)	0.541	~~0.403~~		
9. What color would the hazardous chemical vapor be if it leaked? (M-item)	0.640			
11. What tools do I need to unbolt the valve? (M-item)	0.803			
1. Are there any health risks associated with a hazardous chemical leak that could impact the local community? (Post-conv)		0.758		
7. What is the possibility of a larger issue, such as an explosion or fire, if a hazardous chemical were to leak from the valve? (Post-conv)		0.818		
8. What are the company's regulations about issues with valve openings? (Conv)	~~0.425~~	0.562*		

Table 17 (continued)

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
10. What impact would a hazardous chemical leak have on the environment? (Post-conv)		**0.807**		
14. Is it ever right to regularly override equipment manually in deviation of established operational procedures? (Post-conv)		0.430	~~-0.488~~	
5. Would the company lose production if the valve doesn't get opened? (Conv)	~~0.479~~		0.437	~~-0.459~~
12. How would a hazardous chemical leak impact the company's image? (Conv)			0.734	
2. Would your co-worker lose confidence in your abilities if you asked for assistance? (Pre-conv)				-0.842
3. Would the engineering supervisor be irritated with you if you asked for help? (Pre-conv)				-0.867
13. Is violating the standard operating procedures grounds for your employer to fire you? (Pre-conv)				-0.521*

Loadings with a strikethrough indicate an item that was double loaded and was removed from one of the factors. Items with an asterisk were not well correlated, and were revised following this study.

Factor 1 – M-items. Factor one contained items 5 and 8 which were conventional, and items 4, 6, 9 and 11 which were M-items. Items 5 and 8 loaded onto multiple factors, and were able to be removed from the first factor. This resulted in factor one only containing M-items. The correlation of these items was measured using Cronbach's alpha, and obtained a value of 0.642. Inter-items correlations remained below 0.8, which indicated that there were no redundant items. The alpha value is not high enough to consider the items as strongly correlated. However, this could be due to the low number of items being correlated (Tavakol and Dennick, 2011). Moving forward, one or two of the M-items will be eliminated since all four are not needed for the purpose of identifying bogus data.

Factor 2 –Post-conventional. Factor 2 contained item 8 which was conventional, items 1, 7, 10, and 14 which were post-conventional, and item 6 which was an M-item. Item 6 loaded onto multiple factors, and was able to be removed from this factor. This resulted in all the post-conventional items and one conventional item loading onto the second factor. In order to remove the conventional item from the factor, it will be revised to behave more like the other conventional items. The correlation of the post-conventional items was measured using Cronbach's alpha, and was determined to be

0.641. Inter-item correlations remained below 0.8, which indicates there were no redundant items.

Factor 3 – Conventional. Factor three contained items 5 and 12 which were conventional, and item 14 which was post-conventional. Item 14 had loaded onto multiple factors, and was able to be removed from the third factor. This resulted in factor three containing most of the conventional items. The correlation between items 5 and 12 was measured using Cronbach's alpha, and was calculated to be 0.499. Inter-item correlations remained below 0.8, which determined that there were no redundant items. The low alpha value is indicative of a weak relationship between the two items (Tavakol and Dennick, 2011). In order for all of the conventional items to load together, and to strengthen the correlation between the conventional items, item 8 will be revised to behave more like items 5 and 12.

Factor 4 – Pre-conventional. Factor four contained items 2, 3 and 13 which were pre-conventional, and item 5 which was conventional. Item 5 loaded onto multiple factors, which allowed for it to be removed from the fourth factor. This resulted in factor four containing all the pre-conventional items. The correlation between these items was measured using Cronbach's alpha. The value was calculated to be 0.657, which was able to be improved by removing item 13. The correlation between items 2 and 3 obtained an alpha value of 0.798. Although eliminating item 13 would result in a stronger correlation between items two and three, item 13 will be kept for multiple reasons. At this point, dilemma five contains three pre-conventional items, and eliminating item 13 would result in two pre-conventional considerations. It would also result in dilemma five containing

12 items. Moving forward, it is not ideal to complete a second factor analysis on a dilemma that contains 12 items and only two pre-conventional considerations. In order to strengthen the correlation between the pre-conventional items, item 13 will be revised to behave more like items 2 and 3.

Summary. In order to obtain stronger correlations, and correct loading of the items according to their schema, proposed revisions will be made. Moving forward, item 13 will be revised to behave more like items 2 and 3, which will result in a stronger correlation between the pre-conventional items. Item 8 will be revised to behave more like items 5 and 12, which will result in the conventional considerations loading onto factor three. M-items will be analyzed to determine which considerations are most relevant to keep, which should also strengthen the correlation between the M-items. Table 18 summarizes the results from the factor extraction, as well the anticipated results following the revision process.

Table 18

Factor extraction results and desired results for dilemma five.

Number	Name	Extraction Results	Desired Results	Revisions
Factor 1	M-items	4, 6, 9, 11	4 or 6 or 9 or 11	Reduce no. considerations
Factor 2	Post-conventional	1, 7, 8, 10, 14	1, 7, 10, 14	None
Factor 3	Conventional	5, 12	5, 8, 12	8
Factor 4	Pre-conventional	2, 3, 13	2, 3, 13	13

Dilemma 6. ***Appropriateness of data.*** The appropriateness of data test revealed that item 15 had an MSA below 0.6. This item, which had an MSA value of 0.540, was eliminated and the process was run again which resulted in successful results. The secondary analysis gave a KMO value of 0.770, and all remaining items having an MSA value above 0.6. These successful results allowed for the process to move forward with factor extraction.

Factor extraction. Factor extraction was completed by using the principal component analysis approach and an oblique rotation with delta set to zero. Considerations with factor loadings below 0.4 were considered weak and were suppressed from the structure matrix in order to determine where the considerations were most strongly loaded. All of the items strongly loaded onto at least one factor. Items 1, 6, 7, and 14 loaded onto more than one factor. Table 19 below shows the results from the factor extraction on dilemma six.

Table 19

Dilemma six considerations and factor loadings.

	Factor			
Consideration number and prompts	**1**	**2**	**3**	**4**
2. Are you concerned that your yearly bonus will be impacted if your company discontinues the use of this chemical? (Pre-conv)	0.75			
6. What is the difficulty and personal time investment it will take for you to find a replacement additive? (Pre-conv)	0.523		~~0.599~~	
11. Is there an opportunity for the company to obtain positive press from eliminating the use of the additive? (M-item)	0.651			
12. Are you concerned about your job security if you should fail to find an appropriate alternative? (Pre-conv)	0.777			
8. What is the potential for negative environmental and human consequences if the additive is eventually proven to be dangerous? (Post-conv)		0.752		
9. What is your own personal exposure to the additive on a regular basis? (Pre-conv)		0.653		

Table 19 (continued)

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
13. Is it ever right to knowingly discharge a chemical that is suspected to be hazardous? (Post-conv)		**0.611**		
14. What level of performance must the replacement additive meet to be a viable replacement? (M-item)		~~0.652~~	0.456	
1. What is the potential for lost production if you discontinue the additive without finding a suitable replacement? (Conv)	~~0.484~~		0.648	
7. How important is it that the government agency in charge of environmental regulations has not issued any ruling on the continued use of this additive? (Conv)			0.664	~~0.402~~
10. Is there any additional time or money it would cost your company to replace the additive? (Conv)			0.829	
3. How important is it that data on the additive leans towards supporting evidence of negative consequences, but it is not conclusive? (Conv)				0.461

Table 19 (continued)

	Factor			
Consideration Number and Prompt	**1**	**2**	**3**	**4**
4. Does it matter that your usage of the chemical is in general at a small scale, especially compared to the total national or global usage? (Post-conv)				**0.796**
5. What types of products is the additive used to make? (M-item)				0.667

Loadings with a strikethrough indicate an item that was double loaded and was removed from one of the factors. Items with an asterisk were not well correlated, and were revised following this study.

Factor 1 – Pre-conventional. Factor one contained items 2, 6, and 12 which were pre-conventional, item 1 which was conventional, and item 11 which was an M-item. Item 1 loaded onto multiple factors and was able to be removed from factor one. Item 11 was eliminated, due to the number of M-items already present for this dilemma. Dilemma six contained three M-items, which is not necessary. This resulted in factor one containing only pre-conventional items. The correlation between items 2, 6, and 12 was measured using Cronbach's alpha, and had a value of 0.664. Inter-item correlations remained below 0.8, which indicated there were no redundant items.

Factor 2 – Post-conventional. Factor two contained item 9 which was pre-conventional, items 8 and 13 which were post-conventional and item 14 which was an M-item. Item 14 loaded onto multiple factors and was able to be removed from the second factor. This resulted in factor two containing most of the post-conventional items, and one pre-conventional item. The correlation between the post-conventional items on factor two was measured using Cronbach's alpha and obtained a value of 0.351. Inter-item correlations remained below 0.8, which indicates there were no redundancies. In order for factor two to contain only post-conventional items, and to strengthen the relationship between the post-conventional items, revisions will be made. Item 9 will be revised to behave more like items 2, 6, and 12, which will result in the pre-conventional items loading together, and item 9 unloading from the second factor. Item 4 will be revised to behave more like items 8 and 13, which will result in the post-conventional items loading together, and obtaining a stronger correlation.

Factor 3 – Conventional. Factor three contained item 6 which was pre-conventional, items 1, 7, and 10 which were conventional, and item 14 which was an M-item. Item 6 loaded onto multiple factors, and was able to be removed. This resulted in factor three containing most of the conventional items, and an M-item. The correlation between items 1, 7, and 10 was measured using Cronbach's alpha, and was determined to be 0.630. Inter-item correlations remained below 0.8, which indicates there were no redundant items. In order for factor three to contain only conventional items, revisions will be made. Item 14 will be revised to behave less like the conventional items, which will result in it unloading from the third factor.

Factor 4 – M-items. Factor four contained items 3 and 7, which were conventional, item 4 which was post-conventional, and item 5 which was an M-item. Item 7 loaded onto multiple factors and was able to be removed from the fourth factor. This resulted in factor four containing a conventional, post-conventional, and M-item. Ideally, the fourth factor will contain the M-items. In order to achieve this, revisions will be made. Item 3 will be revised to behave more like items 1, 7, and 10, which will unload item three from the fourth factor. Item 4 will be revised to behave more like items 8 and 13, which will result in unloading from the fourth factor. Item 14 will be revised to behave more like item 5, which will result in the M-items loading together on factor four.

Summary. In order to obtain stronger correlations, and correct loading of the items according to their schema, proposed revisions will be made. Moving forward, item 9 will be revised to behave more like items 2, 6 and 12, which will result in the pre-conventional considerations loading onto factor one. Item 3 will be revised to behave more like items 1, 7, and 10, which will result in the conventional considerations loading onto factor three. Item 4 will be revised to behave more like items 8 and 13, which will result in the post-conventional considerations loading onto factor two. Item 4 will be revised to behave more like item 5, which will result in the M-items loading onto factor four. Table 20 summarizes the factor extraction results, as well as the anticipated results that will follow the revisions.

Table 20

Factor extraction results and desired results for dilemma six.

	Name	Extraction Results	Desired Results	Revisions
Factor 1	Pre-conventional	2, 6, 11, 12	2, 6, 9, 12	9, Eliminate 11
Factor 2	Post-conventional	8, 9, 13	4, 8, 13	4
Factor 3	Conventional	1, 7, 10, 14	1, 3, 7, 10	3
Factor 4	M-items	3, 4, 5	5, 14	14

Dilemma 7. *Appropriateness of data.* The appropriateness of data test revealed that items 1, 6, 9, and 13 had an MSA below 0.6. These items were eliminated, and the process was run again to reveal that item 4 had an MSA below 0.6. Item 4 was eliminated, and the process was run again which resulted in successful results, however, nine items remained for the dilemma. Ideally, each dilemma will have 12 items, which follows the DIT2 and EERI (Rest et al, 1999a; Zhu et al., 2014). Additionally, no post-conventional considerations remain as a result of the appropriateness of data. For these reasons, the dilemma will be removed from the instrument.

Impact of bogus data. Results from the factor analysis that contained the bogus data (initial) and the factor analysis that did not contain the bogus data (secondary) are discussed in this section to convey the impact of the bogus data. Results for some dilemmas were impacted, while others remained the same. For example, the appropriateness of data from the initial factor extraction on dilemma one only removed item 5, while the secondary factor extraction removed items 5 and 8. The initial factor extraction on dilemma one also resulted in four factors, while the secondary factor extraction resulted in three factors. The results from the initial and secondary factor extraction on dilemma one is shown in Table 21.

Table 21

Initial and secondary factor extraction results on dilemma one

	Initial	**Secondary**
Factor 1	2, 9, 10, 13, 14	2, 6, 7, 9, 10, 13, 14
Factor 2	1, 4, 8, 12, 14	2, 3, 6, 11
Factor 3	3, 11	1, 4, 12, 14
Factor 4	2, 6, 7, 8	

Despite having a different number of factors, the items still appear to be loading similarly. Factor one contains items 2, 9, 10, 13 and 14 for the initial and secondary extraction. Items 1, 4, 12 and 14 loaded together for the initial and secondary extractions, although they loaded onto different factors. Items 3 and 11 loaded together as well. This was common for all of the dilemmas that did not load exactly the same during the initial and secondary extractions.

Other dilemmas, such as dilemmas three and four, were less affected by the bogus data. Items loaded the same for all factors in dilemmas three and four during the initial and secondary factor analysis. The factor extraction results for the initial and secondary extraction on dilemma three is shown in Table 22.

Table 22

Initial and secondary factor extraction results on dilemma three

	Initial	**Secondary**
Factor 1	1, 2, 5, 6, 13	1, 2, 5, 6, 13
Factor 2	3, 9, 10, 12, 13	3, 9, 10, 12, 13
Factor 3	7, 8	7, 8
Factor 4	4, 5, 7	4, 5, 7

All of the items loaded exactly the same between the initial and secondary extractions. Dilemma four had the same results. However, the correlation of the items and factor loading values were different between the initial and secondary extractions.

Despite differences in the number of factors, or the amount of items removed, all dilemmas were not affected by the bogus data in terms of the reliability analysis, and suggestions moving forward. Eliminating the bogus data from the data set did not change the end results for the dilemmas. However, by eliminating the bogus data, the procedure was made more robust since only reliable and complete data was being used.

Conclusions

Data collection. The Engineering Process Safety Research Instrument was implemented in a senior chemical engineering capstone classes across three institutions. There were 237 student responses collected between Rose Hulman Institute of Technology, University of Connecticut and North Carolina State University. This sample size is sufficient to complete a factor analysis on a dilemma basis, but not on the overall instrument. Moving forward the EPSRI could be implemented to a larger sample size that might allow for stronger results, and validation of the overall instrument (Hair et al., 1995).

Bogus data analysis. Unreliable, or bogus data was able to be identified through three tests. The rate-rank score analyzed the consistency of the ranked items, and how they were rated in comparison to the other items. Missing data identified unreliable data by calculating how many responses were omitted from the data set. Repeating data identified unreliable data by analyzing the amount of considerations that were rated the same within a dilemma (Rest et al., 1999a). Following the bogus data analysis, 14 student protocols were purged from the data set. Eliminating the bogus data from the data set ensured that the procedure was robust and that only reliable and complete data was included in the analysis.

Factor analysis. The factor analysis was completed to determine the underlying latent variables for each dilemma, as well as the correlations between the items that loaded together (Devellis, 2012; Pett et al. 2011). Additionally, validation of the EPSRI would result in an instrument for senior chemical engineering students that would produce validated and meaningful results (Briggs and Cheek, 1986). Ideally, each dilemma was supposed to have four factors that contained pre-conventional, conventional, post-conventional considerations, or M-items. While this did occur several times throughout the factor analysis, it was not consistent across the instrument. The correlations between the items that were loading together were measured with Cronbach's alpha and ideally should have been above 0.7 (Tavakol and Dennick, 2011). While this was common with some factor loadings, it was not consistent across the instrument. For these reasons, the instrument was not able to be validated during this initial study.

Moving forward, the large scale validation study will be re-run in the fall semester of 2018 in hopes of obtaining stronger results, and validation of the instrument. Prior to the re-implementation, 22 considerations will be revised. In order to revise these items, considerations that loaded together within a schema will be compared to those that loaded onto different factors. The outlier considerations will be rewritten to reflect the structure of the considerations that had loaded together. Additionally, these considerations will maintain their original theme as to not invalidate the work done from the content validation study.

As a result of the appropriateness of data, one dilemma and six considerations were removed. The reliability testing resulted in the elimination of one consideration, and the revision of 22 considerations. These considerations were revised collaboratively between the researchers. These items were revised instead of being eliminated in order to ensure each dilemma has at least 12 considerations. Eliminating items that had not loaded properly does not ensure validation. Additionally, the factor extraction would have to be re-run on the pool of items following the elimination, which might have resulted in items loading differently than they had during the initial validation study.

Table 23 shows the updated list of dilemmas and considerations following the large scale validation study. Item numbers are shown in the brackets. The EPSRI will be implemented in a senior chemical engineering class in the fall of 2018 to make a second attempt at validating the instrument.

Table 23

Considerations per dilemma following the large scale validation study

	Pre-conventional	**Conventional**	**Post-conventional**	**M-items**	**Total**
Dilemma 1	4 [1, 3, 11, 14]	3 [2, 7, 13]	3 [6, 9, 10]	2 [4, 12]	12
Dilemma 2	4 [2, 6, 9, 14]	4 [3, 5, 8, 10]	4 [1, 4, 12, 15]	3 [7, 11, 13]	15
Dilemma 3	4 [1, 5, 7, 13]	3 [2, 6, 10]	4 [3, 8, 9, 12]	1 [4]	12
Dilemma 4	4 [4, 9, 11, 12]	4 [1, 5, 8, 10]	3 [2, 7, 13]	1 [3]	12
Dilemma 5	3 [2, 3, 13]	3 [5, 8, 12]	4 [1, 7, 10, 14]	4 [4, 6, 9, 11]	14
Dilemma 6	4 [2, 6, 9, 12]	4 [1, 3, 7, 10]	3 [4, 8, 13]	2 [5, 14]	13

Chapter 5

Process Safety Moral Reasoning Think Aloud Study

Overview

The objective of the process safety moral reasoning think aloud study was to ensure the instrument was clear and understandable by its intended audience: senior chemical engineering students. The study was also meant to determine how students approached making process safety decisions. Quantitative methods were used to calculate the students' p-score, N2 score, CDIT score. From these scores, students' moral reasoning could be determined as they made process safety decisions. Using qualitative methods, students' responses were analyzed to determine how students were reasoning through process safety decisions. The qualitative and quantitative methods were compared to determine if their predominant reasoning found from the quantitative methods truly reflected the moral reasoning they were exhibiting in their responses.

The think aloud study took place during the spring semester of 2018. Five senior chemical engineering students participated in the study. Students read through the Engineering Process Safety Research Instrument (EPSRI) and verbalized their thought process as they proceeded through the scenarios. The students' responses were audio recorded and transcribed. The transcriptions were analyzed and coded using provisional coding, which determined pre-conventional, conventional, and post-conventional themes in the responses.

Dilemma seven, which was eliminated in the large scale validation study, was tested in this study, and the responses were analyzed and used in the results section. The think aloud study took place before the large scale validation study was completed and dilemma seven was eliminated from the instrument. However, the results from dilemma seven were important in supporting the conclusions made in this study.

Introduction

Content validation study. The first validation study that the EPSRI underwent was the content validation study. Content experts reviewed the instrument to ensure the dilemmas represented realistic process safety situations, the considerations represented their perceived definitions, and that no content areas were omitted. The original version of the EPSRI contained eight dilemmas, which had 15-17 corresponding considerations. Following the content validation, one dilemma and eleven considerations were eliminated. The version of the EPSPI following the content validations study can be seen in Table 6 in Chapter 3.

Large scale validation study. The large scale validation study was done to ensure that the items were being interpreted as pre-conventional, conventional, post-conventional or M-items. The study also determined the strength of the correlation of the items on the instrument. Following the large scale validation study, one dilemma and seven considerations were eliminated. From the results obtained in the large scale validation study, considerations that were not being interpreted correctly were able to be revised. At the conclusion of the study, 22 considerations were revised. The instrument was not yet validated at the conclusion of this study due to not all of the items being interpreted correctly and weak correlations existing between some of the items.

Overview of methods. This was a mixed methods study that employed quantitative and qualitative methods to answer the four research questions. The quantitative methods were analyzed to determine the moral reasoning schema students represented when making process safety decisions. The qualitative methods were used to determine how students were reasoning through the instrument. The results from both methods were compared to determine if their moral reasoning schema accurately reflected the reasoning students demonstrated as they moved through the instrument.

Quantitative methods. The data obtained from the EPSRI during the protocol was used to determine the students' p-score, N2 score, CDIT score and predominant reasoning which were all adapted from Rest et al. (1997a; 1997b). The p-score and N2 score are able to determine where on the moral reasoning spectrum a student falls, however, a student cannot be classified as pre-conventional, conventional, or post-conventional from these scores. The CDIT score determines if a student was consolidated in their form of reasoning, or if they were transitioning between two forms. Predominant reasoning reflects the schema students resided in the most during the study. Combining the CDIT and predominant reasoning classifies a student into a specific moral schema, which will be explained later in this chapter.

Qualitative methods. The responses from the study were analyzed and coded using provisional coding. Provisional coding is a form of evaluation coding which begins with research-generated codes (Miles, Huberman, and Saldana, 2014). Codes were generated for pre-conventional, conventional, and post-conventional thought processes and used to code the student responses to find themes of moral reasoning. Frequencies of the student codes were generated to determine which form of reasoning was most prevalent in their transcripts. Examples of codes, along with the codes that were used, are provided in this chapter.

Research questions. This study seeks to answer four research questions that are based on Kohlberg's Moral Development Theory, and its application to senior chemical engineering students when making process safety decision. The questions are as follows:

1. How can the p-score and N2 score be applied to understand students' moral reasoning?
2. What schemas of moral reasoning do senior chemical engineering students demonstrate when performing process safety decisions?
3. How do senior chemical engineering students reason through process safety decisions?
4. Do the schemas of moral reasoning students represent truly reflect their moral reasoning process when approaching process safety decisions?

Overview of results. The quantitative results determined that four of the five students were post-conventional and were consolidated in this form of thinking. The fifth student was conventional, but transitioning between conventional and post-conventional reasoning. The p-scores and N2 scores were analyzed on the overall instrument as well as on a dilemma basis. Overall, the students who were post-conventional had higher p-scores and N2 scores than the student who was conventional.

The qualitative results determined that all the students represented mostly post-conventional reasoning. Multiple themes were discovered throughout the responses. Students would often make mention of pre-conventional themes in a dismissive manner such as job security, potential bonuses and personal health. Verbal transitions through two or three levels of moral reasoning were prevalent in the student responses. Lastly, students often did not appear to think about the considerations before making their decisions, and would often use the considerations to support their decision.

When comparing the qualitative and quantitative results on the instrument, it was determined that students' overall predominant reasoning, determined by the quantitative methods, was not fully representative of their reasoning that was shown in the qualitative analysis. Further investigation of the quantitative and qualitative results on a dilemma basis showed the same results although analyzing this comparison on a dilemma basis was more accurate for each student. From this study, it was concluded that students' predominant reasoning was not fully reflective of their reasoning when they moved through the instrument; however, it was informative on which form of reasoning the students resided in the most.

Methods

Data collection. Five senior chemical engineering students participated in the think aloud study. Students were given notice of the opportunity to participate through their student email. One student who participated was involved on a separate project which included knowledge on Kohlberg's Moral Development Theory. However, students' data sets were de-identified prior to the data analysis to avoid unintended bias in the analysis of the results. Students who responded and completed the think aloud protocol were given a $50 Visa gift card at the conclusion of their participation. Proper human subject approval was obtained prior to this study.

Students were given a copy of the EPSRI to read and complete during the think aloud protocol. Students were asked to read the dilemmas, decision choices, and considerations out loud and verbalize their thought process as they moved through the instrument. Students began by reading the dilemma, sharing what they were considering, and indicating what decisions they could make. Students would move on to the decision choices, make a decision, and then verbalize why they made that decision. Students would then proceed to the considerations that accompanied the dilemma and rate them on a scale from one (none) to five (great) in terms of importance towards their overall decision. Students would recall the most important considerations and rank their top four. At the end of each dilemma and set of corresponding considerations, students were asked three questions about their reasoning that were adapted from Sadler and Zeidler (2005). Throughout the think aloud protocol, students were not aware of what was being measured from the instrument. All student responses were audio recorded and transcribed. A researcher was present in the room with the student during the protocol to take notes on student behavior and assist with questions the students may have.

During the first think aloud protocol, the senior member was present, and the junior member was not. During this protocol the senior member of the research team noticed the students would look to the senior member for validation on their responses. This was validated during the second think aloud protocol, which included the senior and junior member of the research team. As a result, the remaining three protocols were completed with the junior member of the research team. The senior member would be present for the first dilemma to ensure any confusion about the instrument was clear before leaving the room for the remainder of the protocol.

Quantitative methods. The scores obtained on the EPSRI during the think aloud protocol were used to determine the students P-score, N2 score, and CDIT score, which were adapted from Rest et al. (1997a; 1997b). The methods for calculating these values will be explored in the following sections.

P-score. The P-score is determined from the ranking of the post-conventional items and was adapted from Rest et al. (1997a). Students are unable to be classified from their p-score, however, a high p-score indicates post-conventional reasoning. In order to calculate the p-score, values are attributed to the post-conventional items that are ranked. If a student ranks a post-conventional item as most important, four points are added to the p-score. If a post-conventional item is ranked as second most important, three points are added to the p-score. This follows for the remaining ranked items (third ranked = 2 points, last ranked = 1 point), and is completed across the seven dilemmas. This value is divided by a base, which is the highest amount of points a student can obtain. On the EPSRI, the base score is 70 points. However, if a student omits a ranking, the base score is adjusted to account for that. For example, if a student does not rank a second item on one of the dilemmas, but does rank a first, third and fourth item, three points are taken away from the base score to adjust for the missing rank to have their new base score be 67.

For this study, the p-score was found on the overall instrument as well as on a dilemma basis. This allowed for better analysis of the students p-scores, as well as a better comparison with the qualitative data later in the study. In order to calculate a p-score on a dilemma basis, points were added up for one dilemma and then divided by a base of ten points, the highest score that was able to be achieved on a dilemma basis. Missing ranks were also accounted for on a dilemma basis, however, there was no missing rankings due to the students completing the instrument during the think aloud protocol.

N2 score. The N2 score is determined from the ranking of the post-conventional items, and the ratings of the pre-conventional and post-conventional items. The N2 score was adapted from Rest et al. (1997a) to fit the EPSRI. Similar to the p-score, students are unable to be classified from their N2 score, however, a high N2 score indicates post-conventional reasoning.

The N2 score is calculated through two parts. The first part of the score is calculated similarly to the p-score, however, missing ranks are not accounted for. In other words, the base score will always be 70 points, regardless of missing data. The second part of the score is calculated from the ratings of the pre-conventional and post-conventional data. A sample calculation is given in the following section that describes how the N2 score is found. Equation one summarizes how the N2 score is found.

$$\left(\overline{x_{post}} - \overline{\mathrm{x}_{pre}}/_{S}\right) \times 3 + P = N2 \qquad \text{[Eq. 1]}$$

The average ratings of the post-conventional and pre-conventional items are represented by $\bar{x}_{post}$ and $\bar{x}_{pre}$, respectively. The difference of these values is divided by the sample standard deviations of the pre-conventional and post-conventional ratings, which is represented by *s* in Equation 1. This value is multiplied by 3 to equalize the two parts of the score, due to the second part of the score (contained within the parenthesis) having about 1/3 the standard deviation of the p-score (Rest, Thoma, Narvaez and Bebeau, 1997a). To find the N2 score, the two scores are combined by adding the P-score to what has been discussed so far.

Table 24 contains the ratings of all the considerations from one of the students responses to dilemma three. Pre-conventional considerations are denoted by an italicized font, and post-conventional considerations are denoted by a bold font. Table 25 shows the necessary variables needed to solve equation one.

Table 24

Ratings of considerations from dilemma three

Consideration	*1*	2	**3**	4	*5*	6	*7*	**8**	**9**	10	11	**12**	*13*
Rating	*5*	1	**2**	1	*1*	4	*1*	**5**	**5**	2	5	**5**	*1*

Table 25

Values to solve equation one obtained from student data

Variable	$\bar{x}_{post}$	$\bar{x}_{pre}$	***s***	***P***
Value	4.25	2	2.03	70

Plugging the values from Table 25 into Equation one gives an N2 score of 73.3, shown below.

$$\frac{4.25 - 2}{2.03} \times 3 + 70 = 73.3$$

The N2 score was calculated on the instrument as well as a dilemma basis. This allowed for a better analysis of the students' N2 score, and determined if the overall N2 score accurately described the students' N2 score on a dilemma basis. The sample calculation given in this section was completed on a dilemma basis, due to the amount of analysis needed to be completed for an overall calculation. In order to find the N2 score on the overall dilemma, the average rating of all of the pre-conventional considerations are subtracted from the average rating of all the post-conventional considerations. This difference is divided by the sample standard deviation of all the pre-conventional and post-conventional ratings. This value is multiplied by three before being added to the students P-score for the overall instrument.

CDIT score and schemas. The CDIT score is determined from all of the item ratings and classifies a student as consolidated or transitional. The CDIT score is the ratio of variance of ratings within schemas to the variance of ratings between schemas. The following sections will describe how to calculate the CDIT score with a sample calculation and will discuss how the CDIT score plays a role in classifying a student.

CDIT score calculation. The CDIT score is calculated through a five-step process, and utilizes the ratings of the pre-conventional, conventional, and post-conventional items. The calculation was adapted from Rest et al. (1997b) to fit the EPSRI. The first step of the calculation is to determine the sum of squares total, which is summarized in Equation 2.

$$\left(SS_{pre} \times A_{pre}\right) + \left(SS_{conv} \times A_{conv}\right) + \left(SS_{post} \times A_{post}\right) = SS \qquad \text{[Eq. 2]}$$

The ratings of each form of reasoning were squared and summed to determine their sum of squares. These are denoted by SS_{pre}, SS_{conv} and SS_{post}. In order to account for the varying number of items in each schema (pre-conventional, conventional, or post-conventional), a multiplier is applied to each schema and is denoted by A_{pre}, A_{conv}, and A_{post}. In order to determine the multiplier for each schema, the number of items for each schema are summed and rounded down to the nearest third. The EPSRI contains 82 pre-conventional, conventional and post-conventional items, which was rounded down to 81. This value represents the adjusted number of items overall, which will be denoted by *A* in future equations. This value is divided by the number of schema, which finds the adjusted number of items per schema. For the EPSRI, this value is 27. The multiplier for each item is found by dividing the adjusted number of items per schema by the actual number of items per schema. For the pre-conventional and conventional schemas, the multiplier is one (27/27), and the post-conventional multiplier is 0.96 (27/28). Using student data from the protocol, Table 26 summarizes the variables needed to solve Equation 2.

Table 26

Variables to solve for sum of squares total

Variable	SS_{pre}	SS_{conv}	SS_{post}
Value	202	326	487

Plugging the values in from Table 26 gives the sum of squares value 997.6, shown below.

$$\left(202 \times \left(\frac{27}{27}\right)\right) + \left(326 \times \left(\frac{27}{27}\right)\right) + \left(487 \times \left(\frac{27}{28}\right)\right) = 997.6$$

Following this step, the correction factor is calculated from the pre-conventional, conventional and post-conventional ratings. This step is summarized in Equation 3.

$$\frac{\left((\sum pre \times A_{pre}) + (\sum conv \times A_{conv}) + (\sum post \times A_{post})\right)^2}{A} = CF \qquad \text{[Eq.3]}$$

This step begins by finding the sum of the pre-conventional, conventional, and post-conventional ratings, which are denoted by Σ_{pre}, Σ_{conv}, and Σ_{post}. The multipliers are applied to each sum in order to account for the different number of items per schema. These values are added and squared before being divided by the adjusted number of items overall, which is denoted by *A*, to determine the correction factor. Table 27 gives the necessary values needed to calculate the correction factor using the same data from the previous step.

Table 27

Variables necessary to calculate the correction factor

Variable	Σ_{pre}	Σ_{conv}	Σ_{post}	A
Value	68	86	113	81

Plugging these values into Equation 3 gives the correction factor value of 853.7, as shown below.

$$\frac{\left(\left((68 \times \left(\frac{27}{27}\right)\right) + \left(86 \times \left(\frac{27}{27}\right)\right) + \left(113 \times \left(\frac{27}{28}\right)\right)\right)^2}{81} = 853.7$$

Following this step, the sum of squares deviation is calculated. This step is summarized in Equation 4.

$$SS - CF = SS_{Dev} \qquad \text{[Eq. 4]}$$

This step finds the difference between the sum of squares found in the first step, and the correction factor found in the second step. Sum of squares is denoted by *SS*, and the correction factor is denoted by *CF*. Using the values that were previously calculated, the sum of squares deviation is found to be 143.9, shown below.

$$997.6 - 853.7 = 143.9$$

Following this step, the sum of squares stage value is calculated. This step is summarized in Equation 5.

$$\frac{\left(\sum pre \times A_{pre}\right)^2 + \left(\sum conv \times A_{conv}\right)^2 + \left(\sum post \times A_{post}\right)^2}{B} - SS_{Dev} = SS_{Stage} \qquad \text{[Eq. 5]}$$

This step begins similarly to the correction factor; however, each schema is individually squared instead of the entire numerator. The sum of the ratings for each schema are found, and then the multiplier is applied. The obtained values from each schema are squared then summed. This value is divided by the adjusted number of items per schema, which was previously calculated to be 27, and is denoted by *B.* The sum of squares deviation found in the previous step is subtracted to obtain the sum of squares stage value. Using the values that were presented in Table 27, and the sum of squares deviation that was calculated in the previous step, the sum of squares stage value is found to be 31.2. This example calculation is shown below.

$$\frac{\left((68 \times \left(\frac{27}{27}\right)\right)^2 + \left(86 \times \left(\frac{27}{27}\right)\right)^2 + \left(113 \times \left(\frac{27}{28}\right)\right)^2}{27} - 143.9 = 31.2$$

Following this step, the calculation to obtain the CDIT score is performed. This step is summarized in Equation 6.

$$\frac{SS_{Stage}}{SS_{Dev}} \times 100 = CDIT \qquad \text{[Eq. 6]}$$

In order to determine the CDIT score, the sum of squares stage is divided by the sum of squares deviation and multiplied by 100. The sum of squares stage is represented by SS_{Stage} and the sum of squares deviation is represented by SS_{Dev}. Using the values that were previously found, the CDIT score is calculated to be 21.7, as shown below.

$$\frac{31.2}{143.9} \times 100 = 21.7$$

The CDIT score was only able to be calculated on the overall instrument. When attempting to calculate the CDIT score on a dilemma basis, unreliable values were obtained. This was due to the calculation not being applicable to a smaller set of data.

Schemas and types. The CDIT score determines if a student is transitional or consolidated. A student who is consolidated will have little variance of ratings within schemas and high levels of variance between schemas. For example, if a student rates all the post-conventional items a five, all the conventional items a three, and all the pre-conventional items a one, there is little variance within each schema, but high levels of variance between schemas. This results in a high CDIT score and indicates consolidation. Students who are transitional will have high levels of variance within schemas, but little variance between schemas. For example, if a student rates post-conventional items either four or five, conventional items threes and fours, and pre-conventional items twos and threes, there is more variance within schemas then there is between schemas. This results in a low CDIT score which insinuates transitional behavior.

Students can be classified as one of six types based on their CDIT score and their highest ranked schema. The highest ranked schema is determined from completing the p-score calculation for all three schemas on the overall instrument. Students can be either pre-conventional dominant, conventional dominant, or post-conventional dominant in their reasoning. The CDIT score determines if students are consolidated in their reasoning, or if they are transitioning between two forms. A CDIT score above 15.705 indicates consolidation and below 15.705 is transitional behavior (Rest, Narvaez, Bebeau and Thoma, 1999b). Table 28 shows the matrix of how students are classified based on these two criteria and was adapted from Rest et al. (1999b).

Table 28

Student types based on CDIT score and predominant reasoning.

	Pre-conventional Dominant	**Conventional Dominant**	**Post-conventional Dominant**
Consolidated	Type 1	Type 4	Type 6
Transitional	Type 2	Type 3	Type 5

The types are systematic based on how students should move through moral reasoning. A student who is consolidated in pre-conventional reasoning is classified as type one. Once this student begins to transition into conventional reasoning, but is still mostly pre-conventional, they will be classified as type two. A student who is mostly conventional in their reasoning, but is transitional either between pre-conventional and conventional, or conventional and post-conventional is classified as type three. A student who is consolidated in conventional reasoning is classified as type four. Once that student becomes post-conventional in their reasoning, but is still transitioning from conventional to post-conventional, they are classified as type five. A student who becomes consolidated in post-conventional reasoning would be classified as type six.

Qualitative methods. The use of qualitative research in engineering education is becoming commonplace, since it is able to provide a perspective to the results that quantitative methods cannot capture (Leydens, Moskal and Pavelich, 2004). Qualitative data allows a deeper insight of the individual's perspectives, and creates complete descriptions of certain situations, compared to quantitative data which is meant to provide numerical descriptors of the data. Qualitative data can be captured through multiple methods, including observations, interviews, and documents. This study was conducted as a think aloud protocol, which has a structure similar to interviews. Think aloud studies and interviews are able to capture participant's perspectives. In this study, the perspective that was being recorded was the students' thought process as they reasoned through different process safety scenarios.

For this study, provisional coding was used to analyze the students' transcripts. Provisional coding occurs when codes are created before reviewing the data, based on what the researchers assume may be present in the data. Provisional coding utilizes the generation of codes from theory, and allows for codes to be adjusted, added to, and taken away as the transcripts are analyzed. Provisional coding is especially useful if the work being done is building off previous research (Miles et al., 2014). Since the EPSRI is based on Kohlberg's Moral Development theory, the research can be built upon based on the findings.

Code book. In order to create the code book, each dilemma and their corresponding considerations were reviewed to create a list of possible codes. Codes were developed to reflect potential responses based on the details given in the dilemmas and the questions that were asked in the considerations. Once a list of codes was generated for a dilemma, they were separated into pre-conventional, conventional and post-conventional lists. This was completed across the seven dilemmas individually before the codes were compared. A master codebook was generated that contained codes that were similar, or codes that could be combined. For example, a code generated for one of the dilemmas was “co-worker’s health and safety,” which was a conventional code. For a different dilemma, a code that was generated was “co-worker’s job security,” which was also a conventional code. When creating the master codebook, these codes were combined to create the “co-worker’s concerns” code. This code represented co-worker’s job security, health and safety, and other concerns as well.

While proceeding through the coding of the dilemmas, provisional coding was performed to adjust, add, and remove codes from the codebook. An original code from the master list was “company safety measures and procedures,” and was a conventional code. However, some students were mentioning company safety culture and company safety improvements, which were not encompassed in the code. There were also post-conventional themes that were seen when a student mentioned company safety, such as improving a procedure by communicating with a manager or supervisor. As a result, this code was split into two codes; “company safety culture,” which was conventional, and “safety communication and practice,” which was post-conventional.

The final version of the master codebook is shown in Table 29. The table shows the three schemas, which codes fall within each specified schema, a description of the code and an example where the code was applied. Most codes were not applicable to all seven dilemmas, so the specified dilemmas are shown in parenthesis.

Table 29

Master codebook

Category	Sub-category	Description	Example
Pre-conventional	Career concerns	Students mention keeping their job, yearly bonuses, or advancing in their career	"Then also caring about my job, that would be concerning."
	Personal image/satisfaction	Students mention others view or opinion of them	"I think that the plant workers… would definitely have a better opinion of you if you chose the option safer for them."
	Personal health/exposure	Personal health, safety, or exposure to chemicals from plant	"…if you were exposed to this at a high volume, it's going to negatively impact yourself…"
	Personal time investment/effort (Found in dilemmas 2, 4, 5, 6 & 7)	Personal amount of time or effort spent on a task	"Even if it takes a month to figure it all out, I would still do it."
	Personal belongings (Found in dilemma 2)	Students mention their personal belongings	"Personal belongings, I would say, that's not really important to me because they can be ... well, most of the time they can be replaced…"
Conventional	Co-worker's concerns	Health, safety, time investment, abilities, and job security of co-workers	"I would feel for the people who not only get exposed to it working every day…"
	Company concerns	Company money, time, image, productivity, and equipment	"That is concerning because you wouldn't want to set your company back…"

Table 29 (continued)

Category	Sub-category	Description	Example
	Company safety culture (Found in dilemmas 1, 2, 4, 5 & 7)	Company safety measures, procedures, and general safety culture	"I wouldn't want to break protocol from how to handle opening a valve, so I would say that that affected me greatly too. I wouldn't want to go against what the company does typically."
	Supervisor perception (Found in dilemmas 1, 3, 4, 6 & 7)	Students mention the opinions or though process their supervisor or boss may have	"I guess it depends on how the manager sees things, because if they want to make the most money possible or if they want to run the safest business possible."
Conventional	Family impacts (Found in dilemmas 2, 3, 5 & 7)	Students mention the impacts of their decision on their family	"…you have to consider your own safety and the safety of your family."
	Government regulations/legal issues (Found in dilemmas 2, 3, 6 & 7)	Student mentions government regulations (ex. EPA)	"And that's when you start to get into the OSHA problems and fines…"
	Contractor's safety (Found in dilemma 3)	Students mention the impact on the workers from a contracted company.	"Chances are it's going to immediately impact the people that were working to load and unload the tanks…"
	Product improvement (Found in dilemma 6)	Student mentions ways in which the product could be improved	"…I would want to know it's improving the product and making it safe."

Table 29 (continued)

Category	Sub-category	Description	Example
	Doing the "right" thing	Students mention making the correct decision	"I knew it was the right thing to do to try to find a replacement…"
	Potential for negative consequences	Students mention possible consequences that accompany a decision	"I think it's important to see that there are negative consequences…"
	Community impacts (Found in dilemmas 1, 2, 3, 5, 6 & 7)	Impacts on health or safety of a community	"I think when it comes to things like that, your duty is less to your company and more to the people in the environment in the surrounding area."
	Environmental impacts (Found in dilemmas 2, 3, 5, 6 & 7)	Impacts made to the environment or eco-system	"…be substantially less dangerous to the environment…"
Post-conventional	Safety Communication and Practice (Found in dilemmas 3, 4, 5 & 7)	Students mention how safety practices could be improved through communication with their teams	"…because maybe if I chose to send a correct report about what happened, that would force my company to improve their handling and transporting procedures."
	Greater good for society (Found in dilemmas 2, 6 & 7)	Making a decision that would benefit everyone	"I guess if this product was like curing cancer, maybe that would affect my decision…"
	Risk assessment (Found in dilemmas 1 & 7)	Students weight the potential risk that a decision may have	"…the replacement schedule, because if people left it too or waited a little bit too long just to not have to do it right away after a month, so it would be useless at that point."

Training process. In order to ensure consistency and quality of coding across the transcripts, a training process was created and used by the members of the research team. The training allowed for equal understanding of the codes by ensuring that similar passages from the transcript were being coded the same. Three of the researchers individually coded dilemma six using the code book. When the individual coding was complete, the codes were combined into one document, which the researchers could review before meeting. The researchers met to discuss any discrepancies in the coding and made adjustments to the codebook.

Coding of remaining dilemmas. Each of the researchers coded four of the remaining six dilemmas in a pattern that allowed for each pair of researchers to code two dilemmas together. When coding the remaining dilemmas, the researchers individually coded the dilemmas using the codebook. When coding was complete, the codes were combined into one document which researchers were able to review before meeting. Each pair of researchers met to discuss any discrepancies they had while coding their two dilemmas and make necessary adjustments to the codebook. Adjustments to the codebook were not made until the changes had been discussed with the other researcher. The process allowed for quality of the interpretation of the data.

Research quality. This section reports the steps that were taken to ensure the quality of the data collected and analyzed during the qualitative portion of this study. The Q3 framework was referenced to ensure high quality data (Walther, Sochacka, and Kellam, 2013; Walther, Pawley, and Sochacka, 2015). This framework is constructed of various forms of validation and reliability; however, this study focuses on the theoretical, procedural, communicative, pragmatic validation, and process reliability. Theoretical validation is concerned with the relationship between the theories being used, and the social reality being studied (Walther et al., 2013; Walther et al., 2015). In order to encompass the social reality during the data collection phase, members of the research team reviewed Kohlberg's Moral Development theory (Kohlberg and Hersh, 1977). During the data analysis phase, the code book was created to align with Kohlberg's Moral Development Theory, and the difference between behavioral ethics and ethics was analyzed to understand the results obtained (Kohlberg and Hersh, 1977; Bazerman and Tenbrunsel, 2011). Procedural validation is concerned with the fit between reality and the theory, and the strategies used to ensure contextual validation (Walther et al., 2013; Walther et al., 2015). Procedural validation was achieved by making modifications to the protocol due to the power dynamic that was observed by the senior member of the research team. While analyzing the data, at least two researchers coded each transcript who met to discuss any discrepancies within the coding. An audit trail was kept to keep track of changes made to the codebook, or data analysis plan during the study. Communicative validation ensures that the data collection process is able to encompass the participant's inter-subjective reality (Walther et al., 2013; Walther et al., 2015). We focused on this form of validation by allowing students to alter their answers, giving students feedback only if they were confused by the protocol, or a phrasing of words, and taking notes on students' behaviors. While analyzing the data, researchers would meet to discuss discrepancies in the codes, and changes to the codebook which were agreed upon by all researchers. Pragmatic validation ensures that the underlying concepts of the study are applicable to the reality within the field (Walther et al., 2013; Walther et al., 2015). The participants for this study were senior chemical engineering students, which represents the demographic of people for whom the EPSRI was created. Additionally, the EPSRI focuses on process safety scenarios and decision making, which is beneficial for the chemical engineering education field. During the data analysis, pragmatic validation was achieved by investigating the underlying themes that would relate to process safety decision making. Process reliability ensures the collection of data is dependable, and independent from random influences (Walther et al., 2013; Walther et al., 2015). Student responses during the protocol were audio recorded and transcribed by an outside source. An audit trail was kept to record all steps that were taken during the study, as well as any changes that were made. As part of the data analysis, each researcher would keep their original version of the coded file, in addition to the combined files, and the final codes that were determined after the researchers met. Each pair of researchers would meet to discuss discrepancies within coding, and adjustments that should be made to the codebook. And audit trial was kept as a record of these changes. High quality data was able to be collected and analyzed through understanding of the theories being used, and reference to a validated framework for research quality.

Results and Discussion.

This section will provide results obtained from the quantitative and qualitative methods to answer the four proposed research questions.

Research question 1. How can the p-score and N2 score be applied to understand students' moral reasoning? In order to answer this research question, the p-score and N2 score were calculated based on the students' responses on the EPSRI. Both scores were calculated on a dilemma basis and on the overall instrument. A summary of the scores and results are provided in the following sections.

***P-score*.** The students' p-score was determined from their ranking of the post-conventional items (Rest et al., 1997a). Table 30 shows the students' p-scores per dilemma and overall.

Table 30

Students' p-scores per dilemma and overall.

	Student 1	**Student 2**	**Student 3**	**Student 4**	**Student 5**
Dilemma 1	40	50	70	40	10
Dilemma 2	70	50	50	70	0
Dilemma 3	50	80	70	70	40
Dilemma 4	50	70	60	60	90
Dilemma 5	100	80	90	90	10
Dilemma 6	90	70	70	20	30
Dilemma 7	100	100	90	70	80
Overall	71.43	71.43	71.43	60.00	37.14

Students are unable to be classified from their p-score; however, a higher p-score insinuates post-conventional reasoning (Rest et al., 1997a). From this, predictions can be made about which students were the most and least post-conventional out of the group. Overall, students one, two and three tied for the highest p-score, which implies that they are the most post-conventional out of the group. However, student one had the highest p-score for four out of the seven dilemmas. Student five had the lowest overall p-score, insinuating they were the least post-conventional out of the group. This is supported by the fact that student five was predominantly conventional, which is discussed later in the chapter. However, student five had the highest p-score for dilemma four. This insinuates that the p-score can determine which student was the most post-conventional overall, but for more accurate results, p-scores should be compared on a dilemma basis.

N2 score. The N2 score is calculated from the ranking of the post-conventional items and the rating of the pre-conventional and post-conventional items (Rest et al., 1997a). The N2 score was calculated on the overall instrument, as well as on a dilemma level. The students' N2 scores for the overall instrument and dilemma level are shown in Table 31.

Table 31

Students' N2 score per dilemma and overall

	Student 1	Student 2	Student 3	Student 4	Student 5
Dilemma 1	45.28	55.40	74.11	44.49	14.32
Dilemma 2	75.39	50.00	51.69	73.86	-0.90
Dilemma 3	49.06	85.30	73.32	75.04	43.42
Dilemma 4	54.13	73.33	62.90	63.97	93.49
Dilemma 5	104.04	82.34	93.57	94.10	10.20
Dilemma 6	93.11	72.16	72.27	22.34	32.77
Dilemma 7	104.28	105.39	95.54	75.04	83.01
Overall	75.22	74.89	74.93	64.16	39.79

Similar to the p-score, students cannot be classified from their N2 score. However, a high N2 score insinuates more post-conventional reasoning (Rest et al., 1997a). The students' N2 score can be used to identify which student is the most post-conventional from the group. Student 1 had the highest N2 score overall and implies they were the most post-conventional student from the group. This differs from the results obtained from the p-score, where students 1 through 3 tied for the highest p-score. The N2 score is calculated from the difference between their average pre-conventional and post-conventional rating and the standard deviation of the pre-conventional and post-conventional ratings. This value is added to the p-score as a correction value. A high N2 score results from a large difference between the pre-conventional and post-conventional ratings and a low standard deviation. Student 1 having the highest N2 score insinuates that they rated the pre-conventional items lower than students 1 or 2.

Student 5 had the lowest N2 score overall, insinuating they were the least post-conventional student of the group. This is supported by student 5 being predominantly conventional, which is explained in the following section. However, student 5 had the highest N2 score for dilemma four, which was the same observation made with student five's p-score. Similar to p-score, the N2 score is able to determine which student was most post-conventional overall; however, it may not be true on a dilemma basis. Analyzing the N2 score on a dilemma basis allows for more accurate results.

Research question 2. What schemas of moral reasoning do senior chemical engineering students demonstrate when performing process safety decisions? In order to answer this research question, the students CDIT score and predominant reasoning were calculated from their responses on the EPSRI. From these results, students can be classified into one of the six types which was previously defined.

CDIT score and predominant reasoning. The CDIT score was able to determine if a student was consolidated in their reasoning or if they were transitioning between two forms of reasoning. The CDIT score compares the variance of ratings within a schema to the variance of rating between schemas. A cutoff value of 15.705 was used to determine if a student was consolidated or transitional (Rest et al., 1999b). If the students' score was above the cutoff, they were consolidated and if it was below the cutoff value, they were transitional. The CDIT score was not able to be calculated on a dilemma basis due to the unreliable results that were obtained.

Predominant reasoning was determined from the rankings of the pre-conventional, conventional and post-conventional items. This was determined from the application of the p-score calculation to each of the schemas. Predominant reasoning reflects the level of reasoning the students resided in most during the think aloud. Combining predominant reasoning and the CDIT score, the students' type can be determined. Table 32 below shows the student CDIT score, whether they were transitional or consolidated, predominant reasoning, and type.

Table 32

Students' CDIT score, predominant reasoning and type.

	Student 1	**Student 2**	**Student 3**	**Student 4**	**Student 5**
CDIT score	30.04	21.70	25.16	29.81	13.79
Consolidated (C) or Transitional (T)	C	C	C	C	T
Predominant Reasoning	Post-conv	Post-conv	Post-conv	Post-conv	Conv
Type	Type 6	Type 6	Type 6	Type 6	Type 3

Students 1 through 4 were type six, or post-conventional consolidated. Student 5 was type three, or conventional, and transitioning between two forms of reasoning. Student 1 was the most consolidated student, followed by student 4, 3 than 2. This differs from the results that were observed from the p-score and N2 score. Students 1 through 3 tied for the highest p-score, followed by student 4. Student 1 obtained the highest N2 score, but was followed by student 3, 2 and 4. This insinuates that while the p-score and N2 score are informative when gauging the post-conventional nature of a student, it is not accurate enough to compare students' consolidation in their reasoning. However, student 5 had the lowest overall p-score and N2 score, which reflects the results shown from the CDIT score. Student five obtained the lowest CDIT score and was also conventional and transitional.

In order to obtain a better understanding of the students CDIT score, a graph is provided below which compares the overall rating of each students' pre-conventional, conventional and post-conventional items. A large distance between each bar and smaller error bars indicate high consolidation.

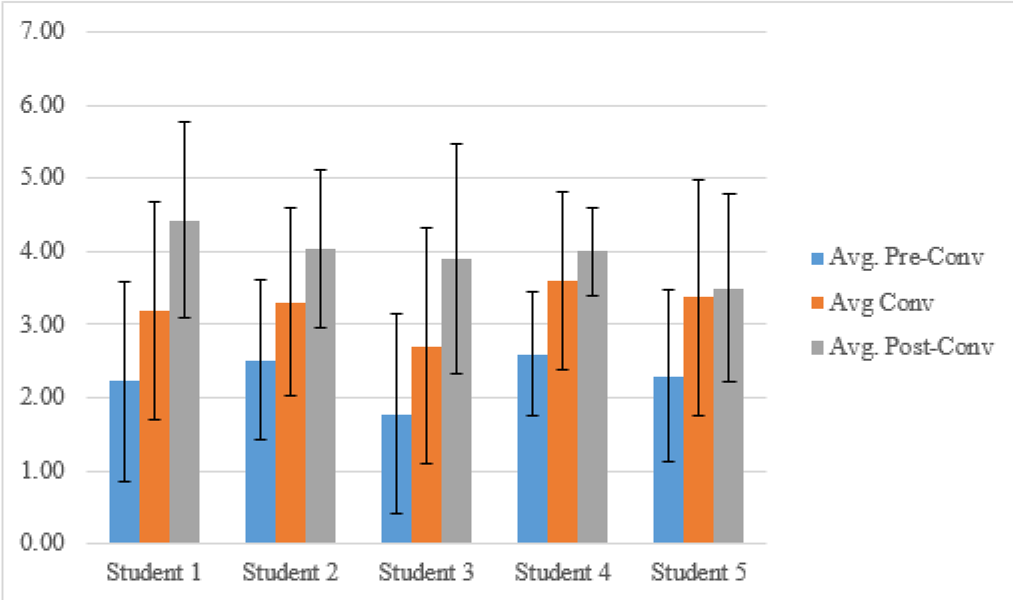

Figure 1. Average ratings of pre-conventional, conventional and post-conventional items

Student 1 has the largest variance between pre-conventional, conventional and post-conventional reasoning which explains why they obtained the highest CDIT score. Student 4 had the second highest CDIT score, even though they had the least variance between conventional and post-conventional reasoning out of the consolidated students. However, student 4 had the least amount of variance within their schemas, which explains why they obtained the second highest CDIT score. Student 3 had similar variance between schemas as student one but obtained the lowest CDIT score out of the consolidated students. This is due to the large amount of variance within their schemas. Student 5, who was transitional, had very little variance between conventional and post-conventional reasoning, and had very high variance within each schema.

Research question 3. How do senior chemical engineering students' reason though process safety decisions? To answer this research question, the students' transcripts were read and analyzed for pre-conventional, conventional and post-conventional codes. The frequencies of the codes were calculated and recorded. The following sections will summarize and give examples of pre-conventional, conventional, and post-conventional codes, and report other themes that were discovered while analyzing the transcripts.

Pre-conventional codes. Pre-conventional reasoning occurs when an individual prioritizes the satisfaction of their needs or wants. Pre-conventional reasoning also encompasses decisions that are made to avoid punishment, or physical consequence (Kohlberg and Hersh, 1977). Students conveyed pre-conventional reasoning if they expressed concerns about their job security, personal health or safety, and their personal image. The following quotes are examples of how students could express concern about their job, health, safety, and image.

- "I think that the plant workers, the people actually that would be responsible for maintaining and changing the lines, would definitely have a better opinion of you if you chose the option safer for them. {Pre-conv - Personal image/satisfaction}"
- "That is most certainly a loss of employment and probably a huge black mark on your resume {Pre-conv - Career concerns}"

- And I feel like I would be much more concerned with the immediate risk to my health than I would be to what repercussions might come later. {Pre-conv - Personal health/exposure}"

Pre-conventional reasoning also occurred if a student conveyed that they did not want to spend much time or effort on a task, or if they wanted to obtain a bonus or promotion. The following quotes are examples of students who expressed concern about their effort, bonuses or career advancements.

- "Again, I think that, I read that as if it's referring to would it be so annoying. Would I have to write up all that paperwork and stuff? {Pre-conv - Personal time investment/effort}"
- "I feel like if anything this might help you get a job after graduation 'cause it's showing that you're taking initiative and actively caring about what the plant is doing and what their safety measures are. {Pre-conv - Career concerns}"

Pre-conventional codes were the least frequent across the instrument for all the students. This finding can be supported by research done by Rest et al. (1999a), who found that moral reasoning increases with age and education. Senior chemical engineering students should not be operating at the pre-conventional level according to the work done by Rest et al. (1999a).

Conventional codes. Conventional reasoning occurs when an individual prioritizes the benefits of people close to them, such as co-workers, family and friends, as well as the needs of their company. Within the EPSRI, students could convey conventional reasoning by expressing concern about their co-worker's health, safety, time investment, and job security, or the time, image, productivity and equipment of their company. The following quotes are examples of how students showed conventional reasoning in their responses.

- "You're not going to be the one changing the lines or working the other chemicals. It's going to be the other employees, so you have to put their needs and safety ahead of any of your own gain that you could get financial from this. {Conv - Co-worker concerns}"
- "An explosion is literally the worst case scenario. Most likely loss of life, millions of dollars in damage and a very big negative impact on the company. {Conv - Company concerns}"

Conventional reasoning also occurs when an individual prioritizes the law or government regulations. In the EPSRI, students would express concern about government regulations such as OSHA or EPA and may have considered the legality of their decisions. Students would also express concern about following the company's safety regulations. Examples of this type of reasoning can be seen in the quotes below.

- "Because if you are breaking the law by not inspecting it as much as the law requires, then that's a big issue right off the bat. That shows negligence. It shows not caring and a bad culture. {Conv - Government regulations/legal issues; Conv - Company safety culture}"

- "I would say, once again, it's not technically illegal or wrong for you to do this, but the ethical implications are there and as soon as the EPA says that there is an issue, then you need to change. {Conv - Government regulations/legal issues}"

Post-conventional reasoning occurs when an individual prioritizes individual rights, justice, the equality of human rights, and respect for humans as individual beings. Those who reason at a post-conventional level follow the law but consider changing the rules depending on the situation. Students were able to convey post-conventional reasoning by expressing conflict over the outputs of their decisions and which would be better for society. The following quotes are examples of conflict students had during the scenarios.

- "Now I'm weighing it on, if you do send volunteers and it all works out, you saved surrounding neighborhoods and the environment from all these bad things that could happen at the expense of, worst case scenario, a couple people who volunteered to be there, even. {Post-conv - The greater good}"
- "If you're making products that have to be that are like actively used to help people and help the environment it might be a little bit of weighing the benefits versus the risks to the environment. {Post-conv - Environmental impact}"

Post-conventional reasoning also prioritizes the safety and health of the environment, as well as people in surrounding communities. At the post-conventional level, individuals consider people who are extraneous to the situation, but may be impacted by the decisions made. The following quotes are examples of the students expressing concern about the surrounding communities, or environment.

- "...'cause if you do have a loss of containment and it does greatly negative impact the environment and surroundings then even if no one did get hurt during the storm or explosion that your plant might have caused, it could negatively impact the quality of life in the area for a long time to come. {Post-conv - Potential for negative consequences}"
- "They're not sure what it would do the environment. And also with a flood coming through, the organic chemicals could actually probably travel much farther than they would if they just got accidentally released normally. {Post-conv - Environmental impact}"

Post-conventional codes occurred most frequently during the think aloud for the students. While it would be ideal for all the students to be operating at this level, it may not be entirely accurate of the students' moral reasoning if they were placed in the actual scenario. Since the students' decisions have no impact, they are working in what is known as the predictive space. Within this space, it is easy for the students to make behavioral forecasting errors or incorrect predictions about how they would behave in the situation (Osberg and Shrauger, 1986).

This can be seen in some of the responses students had toward pre-conventional considerations. Students would often dismiss pre-conventional themes, such as their personal health or safety, job security or a potential bonus. Pre-conventional considerations were rated low on the EPSRI, which is to be expected from senior chemical engineering students. However, the reasoning behind some of these low ratings are telling of behavioral forecasting errors. The following quotes are examples of students dismissing pre-conventional themes.

- "I think ultimately your manager would be a little bit happier which eventually could lead to further promotion or benefits, but that didn't weigh too much on my decision. {Pre-conv - Career concerns}"
- "I wasn't concerned about the bonus or accolades or opportunities for career advancement in this decision {Pre-conv - Career concerns}.

It is promising that a student believes that their job security or safety would not influence them to make an unethical decision; however, they would not understand the full implication of these accolades unless they were actually in the situation. While it is possible that some students would keep the same decision and mindset, others may need the bonus money that accompanies a cheaper design option or may realize the importance of maintaining their job.

Frequency of codes. The frequency of the pre-conventional, conventional, and post-conventional codes were recorded for each student on the overall instrument. All codes were of equal importance to one another. As previously stated, pre-conventional codes were the least frequent for all the students, ranging from 9% to 24%. Post-conventional codes were the most frequent for all the students across the instrument, ranging from 39% to 56%. Table 33 shows the frequency of pre-conventional, conventional and post-conventional codes for the students.

Table 33

Frequency of pre-conventional, conventional and post-conventional codes on the overall instrument.

	Pre-conventional	**Conventional**	**Post-conventional**
Student 1	24%	30%	47%
Student 2	14%	29%	56%
Student 3	17%	34%	49%
Student 4	9%	37%	44%
Student 5	24%	37%	39%

Low frequency of pre-conventional codes was to be expected, as previously stated. Rest et al. (1999a) found in their study that moral reasoning increases with age and education. Senior chemical engineering students should be operating above the pre-conventional level, but the frequencies obtained for post-conventional reasoning might be too high. Rest et al. (1999a) found in their work that a senior undergraduate student should reason at the conventional level. All the students in this study reasoned at the post-conventional level most frequently. However, the students are working in the predictive phase which insinuates that there may have been behavioral forecasting errors in their responses (Osberg and Shrauger, 1986).

Progression in moral reasoning. Kohlberg and Hersh (1977) stated in their theory that the stages of moral reasoning are "hierarchical integrations," meaning that an individual operating at a high level of moral reasoning will still understand and reason through lower levels. According to the frequency of the codes, the students are reasoning post-conventionally through the scenarios, so it was common to see a student reason through lower levels before reaching a high level of moral reasoning. A clear movement from pre-conventional to conventional reasoning was seen when students were asked about their health and safety. In response to the consideration, students would often express concern about their own protection before expressing concern about their co-workers or their company. The quotes below are examples of students showing pre-conventional and conventional reasoning in their responses.

- "Just because it's obviously if you were exposed to this at a high volume, it's going to negatively impact yourself] and chances are if it's technically impacting you it's going to negatively impacting all of your coworkers as well so it's just adding to a very unsafe plant environment. {Pre-conv - Personal health/exposure; Conv - Co-worker concerns}"

Students would also show progression from conventional reasoning to post-conventional reasoning in their responses as well. This was typically a result of a student being asked about the safety of their co-workers. Students would express concern about their co-workers in response to the consideration before moving on to the potential impacts on the environment and surrounding communities. The following quote is an example of a progression from conventional to post-conventional reasoning.

- "That is concerning because you wouldn't want to set your company back; however, it's definitely the right thing to do, especially if the company can be a cause of people's health and safety. {Conv - Company concerns; Post-conv - Doing the "right" thing}"

Occasionally, students would show clear progression through all three levels of reasoning. In these responses, a student may express concern for their personal health or job security before moving on to the safety of their co-workers or the security of the company, but then recognize the impacts of their decisions on surrounding communities or the environment. The following response is an example of a progression from pre-conventional to conventional then post-conventional reasoning.

- …some of the first things I thought were "Do I really want to be around this stuff continually?" I thought of "Well, if we take this month does that slow down productivity and if so, then that will not look good at all if we decide to go with it." And I thought about how it was showing up in animals and other environmental areas and they can't determine if it's bad yet, but it is showing up {Pre-conv - Personal health/exposure; Conv - Company concerns; Post-conv - Environmental impact; Post-conv - Potential for negative consequences}.

Research question 4. Do the schemas of moral reasoning students represent truly reflect their moral reasoning process when approaching process safety decisions? In order to answer this research question, the students' predominant reasoning obtained from the quantitative methods was compared to the frequency of codes obtained from the qualitative methods. Results were compared on an overall basis before further analyzing the discrepancies on a dilemma level. Comparing the results on a dilemma basis allows for a more accurate analysis of the students predominant reasoning and decision-making process.

Overall instrument comparison. From the quantitative results, it was found that students one through four were post-conventional and student 5 was conventional. However, the qualitative results showed that all the students showed mostly post-conventional reasoning across the instrument. The discrepancy between student five's predominant reasoning and moral reasoning in the transcripts indicated that the predominant reasoning is not fully representative of their moral reasoning. This can be explained by Kohlberg's hierarchical integrations in moral reasoning. Hierarchical integrations imply that an individual operating at a higher level of moral reasoning will still understand and reason with lower levels of reasoning (Kohlberg and Hersh, 1977). From the qualitative results, student five can be seen to operate at a post-conventional level. However, student 5 is still using pre-conventional and conventional reasoning, which was represented in their quantitative results.

This can be further investigated by analyzing the students' predominant reasoning and moral reasoning on a dilemma basis. Analyzing the results on a dilemma basis may reveal student five predominantly reasoning at the post-conventional level. Similarly, it may reveal the other students, who were at the post-conventional level overall, operated at lower levels on the dilemmas.

Dilemma level comparison. Predominant reasoning was analyzed on a dilemma basis to determine if students were representing hierarchical integration in their data. Student five, who was predominantly conventional on the overall instrument, was predominantly post-conventional on two out of the seven dilemmas. Students 1, 2 and 4, who were predominantly post-conventional on the overall instrument, were predominantly conventional on at least one of the dilemmas. Student 3 remained post-conventional across the seven dilemmas. These findings prove Kohlberg's observation that an individual operating at a higher level of reasoning will still understand and reason through lower level reasoning. However, the discrepancy between the qualitative and quantitative results for student five on the overall instrument is indicative that the predominant reasoning is not truly reflective of the students' moral reasoning. In order to further investigate this observation, the qualitative results were analyzed on a dilemma basis, and compared to the predominant reasoning on a dilemma basis.

Dilemma one comparison. Table 34 shows the predominant reasoning for each student on dilemma one, as well as the frequency of pre-conventional, conventional and post-conventional codes. Due to the low number of considerations per dilemma, it was common for a student to obtain the same ranking score for two levels of reasoning. The ranking score determined predominant reasoning.

Table 34

Predominant reasoning and frequency of codes for dilemma one

	Predominant Reasoning	Frequency of Pre-conv codes	Frequency of Conv Codes	Frequency of Post-conv codes
Student 1	Conv	17.9%	35.7%	46.4%
Student 2	Conv & Post	13.2%	35.9%	50.9%
Student 3	Post	18.2%	36.4%	45.4%
Student 4	Conv	10.0%	46.7%	43.3%
Student 5	Conv	33.3%	16.7%	50.0%

Students 1, 4 and 5 were predominantly conventional on dilemma one, however, students 1 and 5 showed post-conventional reasoning most frequently in their transcripts. In fact, student 5 showed conventional reasoning the least out of the three schemas for dilemma one. Student 2 tied between conventional and post-conventional predominant reasoning on dilemma one, however, they had about 15% more post-conventional codes than conventional codes in their transcripts. Students 3 and 4 had predominant reasoning that represented their moral reasoning in the transcripts. For dilemma one, predominant reasoning reflected moral reasoning three times out of the possible five, including student 2 who tied for two forms of predominant reasoning.

Dilemma two comparison. Table 35 shows predominant reasoning for each student on dilemma two. Additionally, frequencies of pre-conventional, conventional and post-conventional codes are provided as well. Due to the low number of considerations per dilemma, it was common for a student to obtain the same ranking score for two levels of reasoning. The ranking score determined predominant reasoning. Similarly, it was possible for the same amount of codes to be obtained across multiple levels of reasoning, which resulted in the same frequency of codes.

Table 35

Predominant reasoning and frequency of codes for dilemma two.

	Predominant Reasoning	**Frequency of Pre-conv codes**	**Frequency of Conv Codes**	**Frequency of Post-conv codes**
Student 1	Post	12.5%	47.5%	40.0%
Student 2	Conv & Post	15.9%	34.1%	50.0%
Student 3	Post	10.7%	46.4%	42.9%
Student 4	Post	21.7%	39.15%	39.15%
Student 5	Conv	33.3%	66.7%	0.0%

Students 1, 3 and 4 were predominantly post-conventional on dilemma two. However, students 1 and 3 showed mostly conventional reasoning in their responses to the scenarios. Student 4 showed the same amount of conventional and post-conventional reasoning in their responses. Student 2 was predominantly conventional and post-conventional on dilemma two, however, they had about 15% more post-conventional codes in their transcripts than conventional codes. Student 5, who was predominantly conventional on dilemma two, showed mainly conventional codes and no post-conventional codes. Including the students who tied for predominant reasoning or frequency of codes, predominant reasoning reflected the students' moral reasoning three out of the possible five times on dilemma two.

Dilemma Three Comparison. Table 36 shows predominant reasoning for each student on dilemma three, as well as the frequency of pre-conventional, conventional and post-conventional codes. It was possible for the same amount of codes to be obtained for multiple schemas, which results in the same frequency.

Table 36

Predominant reasoning and frequency of codes for dilemma three.

	Predominant Reasoning	**Frequency of Pre-conv codes**	**Frequency of Conv Codes**	**Frequency of Post-conv codes**
Student 1	Post	36.1%	25.0%	38.9%
Student 2	Post	22.0%	24.4%	53.6%
Student 3	Post	18.8%	43.8%	37.5%
Student 4	Post	20.0%	20.0%	60.0%
Student 5	Conv	16.7%	16.7%	66.7%

Students 1 through 4 were predominantly post-conventional on dilemma three. However, only students 1, 2 and 4 showed mostly post-conventional reasoning in their responses to the scenarios. Student 3 showed mostly conventional reasoning in their responses, despite them being predominantly post-conventional. Student 5, who was predominantly conventional, showed mostly post-conventional codes in their responses. In fact, they had 50% more post-conventional codes than conventional codes. Predominant reasoning was reflective of moral reasoning for three out of the five students on dilemma three.

Dilemma four comparison. Table 37 shows the students' predominant reasoning on dilemma four, as well as the frequency of pre-conventional, conventional and post-conventional codes that appeared in their responses. Due to the low number of considerations per dilemma, it was possible for students to tie between two levels of predominant reasoning.

Table 37

Predominant reasoning and frequency of codes for dilemma four.

	Predominant Reasoning	Frequency of Pre-conv codes	Frequency of Conv Codes	Frequency of Post-conv codes
Student 1	Conv & Post	15.0%	30.0%	55.0%
Student 2	Post	13.5%	13.5%	73.0%
Student 3	Post	32.0%	24.0%	44.0%
Student 4	Post	22.5%	35.0%	42.5%
Student 5	Post	16.7%	33.3%	50.0%

Students two through five were predominantly post-conventional. These students also showed mainly post-conventional reasoning in their responses as well. For these four students, their predominant reasoning was reflective of their moral reasoning for this dilemma. Student 1 tied between conventional and post-conventional reasoning as their predominant form of reasoning. However, they showed mainly post-conventional reasoning in their transcripts. In fact, they had 15% more post-conventional codes than conventional codes, despite them showing the same amount of conventional and post-conventional reasoning in their quantitative results. Excluding student 1, predominant reasoning represented the form of moral reasoning that was most common for all of the students on dilemma four.

Dilemma five comparison. Table 38 shows students' predominant reasoning for dilemma five, as well as the frequency of pre-conventional, conventional, and post-conventional codes that appeared in the transcripts. It was possible for the same amount of codes to be obtained for multiple schemas, which resulted in the same frequency.

Table 38

Predominant reasoning and frequency of codes for dilemma five.

	Predominant Reasoning	**Frequency of Pre-conv codes**	**Frequency of Conv Codes**	**Frequency of Post-conv codes**
Student 1	Post	40.0%	12.0%	48.0%
Student 2	Post	19.4%	41.7%	38.9%
Student 3	Post	17.6%	11.8%	70.6%
Student 4	Post	26.5%	23.5%	50.0%
Student 5	Pre	42.9%	14.3%	42.9%

Students 1 through 4 were predominantly post-conventional, and student 5 was predominately pre-conventional. Students 1, 3 and 4 all showed mostly post-conventional codes in their responses, as shown by their predominant reasoning. However, student 2 showed mostly conventional reasoning in their responses despite being predominantly post-conventional. Student 5 showed the same amount of pre-conventional and post-conventional reasoning in their responses, even though they were just predominantly pre-conventional on dilemma five. Including student 5 who had the same percentage of pre-conventional and post-conventional codes, predominant reasoning reflected the students reasoning that appeared most in their responses for four of the five students.

Dilemma six comparison. Table 39 shows the students' predominant reasoning on dilemma six, as well as the frequency of pre-conventional, conventional and post-conventional codes that appeared in their transcripts.

Table 39

Predominant reasoning and frequency of codes for dilemma six.

	Predominant Reasoning	**Frequency of Pre-conv codes**	**Frequency of Conv Codes**	**Frequency of Post-conv codes**
Student 1	Post	17.5%	22.5%	60.0%
Student 2	Post	14.0%	37.2%	48.8%
Student 3	Post	12.8%	35.9%	51.3%
Student 4	Conv	16.3%	55.8%	27.9%
Student 5	Conv	14.3%	57.1%	28.6%

Students 1 through 3 were predominantly post-conventional for dilemma six, and students 4 and 5 were predominantly conventional. Students 1 through 3 showed mostly post-conventional reasoning in their responses, which is representative of their predominant reasoning. Similarly, students 4 and 5 showed mostly conventional reasoning in their responses, which is representative of their predominant reasoning. For all five students, predominant reasoning reflected the form of reasoning used most in their responses.

Dilemma seven comparison. Table 40 shows the students' predominant reasoning on dilemma seven, as well as the frequency of the pre-conventional, conventional, and post-conventional codes that appeared in the transcripts.

Table 40

Predominant reasoning and frequency of codes for dilemma seven.

	Predominant Reasoning	Frequency of Pre-conv codes	Frequency of Conv Codes	Frequency of Post-conv codes
Student 1	Post	28.6%	32.1%	39.3%
Student 2	Post	4.4%	17.8%	77.8%
Student 3	Post	12.5%	37.5%	50.0%
Student 4	Post	16.2%	35.1%	48.7%
Student 5	Post	0.0%	40.0%	60.0%

All the students were predominantly post-conventional on dilemma seven. Additionally, all the students showed mostly post-conventional reasoning in their responses. Predominant reasoning reflected the form of reasoning used most often for all five students on dilemma seven.

Key takeaways. Students' predominant reasoning and frequency of codes were compared on a dilemma level to determine if predominant reasoning was truly reflective of the students' moral reasoning. For some dilemmas, such as one or two, predominant reasoning was not reflective of multiple students' moral reasoning. For example, student 1 was predominantly post-conventional on dilemma two, but showed mostly conventional reasoning in their responses. Student 1's response to a post-conventional consideration about the company's responsibility to locate its facilities in areas where negative impacts to the surrounding community are minimized is shown below.

- "...It is important for the company to make sure that companies surrounding the plant aren't affected by our mistakes…{Conv - Company concerns}"

Their response to a post-conventional consideration was conventional, because they showed concern about their company instead of the impacts it could have on surrounding communities. However, student 1 rated this consideration as "great," which translated to a five out of five. Even though this student was clearly reasoning conventionally on this consideration, the quantitative analysis would classify the student as post-conventional based on their rating of the consideration.

However, other dilemmas were not as problematic. For dilemmas four, six and seven, the students' predominant reasoning reflected the form of reasoning they used the most for all five students. Predominant reasoning was reflective of moral reasoning for at least three students on all the dilemmas. Across the seven dilemmas and five students, the quantitative results represented the qualitative results on 28 out of the 35 comparisons. Discrepancies between the quantitative and qualitative results could be a result of a few issues.

For instance, students would often dismiss pre-conventional themes such as job security, bonuses, and health, which was previously discussed. However, these responses would still be coded as pre-conventional, since they were mentioning pre-conventional themes. While high frequencies of pre-conventional codes were not common, high frequencies of conventional codes were. Similar to how a student would dismiss pre-conventional themes, students would occasionally dismiss conventional themes as well. For example, a student may dismiss the company's image, however, their response would still be coded as conventional. These instances of dismissing a consideration while still reasoning through the consideration on different lines of moral reasoning could have led to discrepancies between the quantitative and qualitative data.

Students would also use the considerations to support the initial decision they had made on the scenario. The options that accompanied each dilemma were not meant to reflect “right” or “wrong” answers, and the considerations were meant to create an argument for either decision. However, students rarely changed their initial response and would often use considerations to support their decisions. In turn, a student might have been reasoning conventionally, but was using post-conventional considerations to support their decision. As a result, this would lead to discrepancies in the comparative data.

Analyzing these results on a dilemma basis proved to be beneficial due to the accuracy of review that could be completed for each student. On the overall instrument, student 5’s quantitative and qualitative results did not align. However, further investigation showed alignment of the results for several of the dilemmas. Similarly, the students who showed alignment on the overall instrument had discrepancies between their results on a dilemma basis, with the exception of student 4. These results show that predominant reasoning is not truly reflective of the students’ moral reasoning when faced with a process safety decision. Predominant reasoning is useful in understanding the type of reasoning students are portraying the most, but it does not show the full spectrum of the students’ moral reasoning throughout their decision making process. Qualitative methods should be applied to fully depict how a student reasons through process safety decisions.

Conclusions

The objective of this study was to ensure clarity of the EPSRI and to determine how senior chemical engineering students morally reason through process safety decisions. Five senior chemical engineering students participated in a think aloud protocol in which they read through the EPSRI out loud and verbalized their decision-making process. The students' responses were audio recorded and transcribed.

This was a mixed methods study that employed quantitative and qualitative methods. Scores from the EPSRI were used to calculate the students' p-score, N2 score, CDIT score and predominant reasoning. Students were able to be classified as one of six types based on their predominant reasoning and CDIT scores. Students' responses were analyzed for pre-conventional, conventional, and post-conventional themes using provisional coding. Frequency of codes were calculated and recorded for the overall instrument, as well as on a dilemma basis.

The N2 score and p-score are meant to represent the level of students' post-conventional reasoning. A student who obtains a high p-score or N2 score is considered more post-conventional than a student that would obtain a lower score. Students are unable to be classified from these scores, however, they should represent which student was operating the most at a post-conventional level. Student 5 who was conventional obtained the lowest quantitative scores, which was well representative of their predominant reasoning.

Student types are determined from their predominant reasoning and CDIT score. Predominant reasoning reflects the schema that was ranked the highest, and the CDIT score determined if a student is consolidated in their level of reasoning or if they are transitioning between two forms. Four students were type six, or post-conventional consolidated. Student 5 was type three, or conventional and transitional. Student 5 was transitioning between the conventional and post-conventional levels.

Student responses were analyzed for pre-conventional, conventional, and post-conventional themes. All levels of moral reasoning were prevalent in the student responses; however, post-conventional codes were most common for all of the students across the instrument. This is a promising theme; however, undergraduate students should be operating at a conventional level, according to Rest *et al.* (1999b). These results may be due to the students operating in the predictive phase, which insinuates that they may be making behavioral forecasting errors.

The quantitative and qualitative results were compared to determine if the students' predominant reasoning was reflective of their moral reasoning. On the overall instrument, predominant reasoning reflected the form of reasoning that occurred most for four of the students. Student 5, who was predominantly conventional on the overall instrument, showed mostly post-conventional codes in their transcripts. Further analysis was completed by comparing the results on a dilemma basis. Similar results were obtained, but analyzing the results on a dilemma basis allowed for a more accurate interpretation. Overall, it was found that predominant reasoning is not reflective of student moral reasoning, but it does reflect the type of reasoning the students apply most often.

This study provided an understanding of how senior chemical engineering students approach and reason through process safety decisions. The results from this study have shown that moving forward, students should be taught about the full implications of their decisions in order to push them toward true post-conventional behavior.

References

[1] American Institute of Chemical Engineers, "Process Safety Boot Camp" *American Institute of Chemical Engineers Academy*. [Online]. Available: https://www.aiche.org/academy/courses/ch900/process-safety-boot-camp [Accessed: September 5, 2018].

[2] J.A. Baker et al., "The Report of The BP U.S. Refineries Independent Safety Review Panel," Baker Panel Report, January, 2007.

[3] K. Bannigan, and R. Watson, "Reliability and Validity in a Nutshell," *Journal of Clinical Nursing,* vol. 18, no. 23, pp. 3237-3243, 2009.

[4] M.H. Bazerman, and D. Chugh, "Bounded Awareness: Focusing Failures in Negotiation," in *Negotiation Theory and Research,* New York: Psychology Press, 2006, pp. 7-26.

[5] M.H. Bazerman, and A.E. Tenbrunsel, *Blind Spots: Why We Fail to Do What's Right and What to Do About It*, Princeton: Princeton Univ Press, 2011.

[6] M.H. Bazerman, A.E. Tenbrunsel, and K. Wade-Benzoni, "Negotiating with Yourself and Losing: Making Decisions with Competing Internal Preferences," *Academy of Management Review,* vol. 23, no. 2, pp. 225-241, 1998.

[7] S.R. Briggs, and J.M. Cheek, "The Role of Factor Analysis in the Development and Evaluation of Personality Scales," *Journal of Personality,* vol. 54, no. 1, pp. 106-148, 1986.

[8] U. Bruyere, M. Fox, C. Watson, "Process Safety Fundamentals Training for First Level Leaders," *Institution of Chemical Engineers Symposium Series*, vol. 155, pp. 529-535, 2009.

[9] A.D. Carvalho Neto, and I Correa, "Tridimensional Training Matrix to Improve Process Safety Performance and Assure All Organization Involved in PSM Implementation" presented at AIChE Spring Meeting and Global Congress on Process Safety, San Antonio, TX, 2017.

[10] CCPS. *Risked Based Process Safety Overview.* New York, NY: Library of Congress Cataloging-in-Publication Data, 2014.

[11] J. Champion, S. Van Geffen, and L. Borrousch, "Reducing Process Safety Events: An Approach Proven by Sustainable Results," *Process Safety Progress,* vol. 36, no. 4, pp. 326-337, 2017.

[12] M. Chu, S. Li, and F. Yang, "Using Factor Analysis to Train Students of an Engineering Major to Innovate," *World Transactions on Engineering and Technology Education*, vol. 12, no. 2, pp. 181-185, 2014.

[13] *Criteria for Accrediting Engineering Programs, 2018-2019* [Online]. Available; http://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2018-2019/#GC3

[14] S.J. Dee, B.L. Cox, and R.A. Ogle, "Process Safety in the Classroom: The Current State of Chemical Engineering Programs at US Universities," *Process Safety Programs,* vol. 34, no. 4, pp. 314-319, 2015.

[15] R.F. Devellis, *Scale Development: Theory and Applications.* 3rd ed. Thousand Oaks: Sage Publications, 2012.

[16] M.C. Dillon. (2013, April 19). *WVU Offers First-of-its-kind Process Safety Boot Camp to Chemical Engineering Students* [Press Release]. Available: http://wvutoday-archive.wvu.edu/n/2013/04/19/wvu-offers-first-of-its-kind-process-safety-boot-camp-to-chemical-engineering-students.html

[17] D. Grubbe, “Ethics in Sustainability and Engineering,” presented at the 2018 Engineering Education for Sustainable Development, Glassboro, NJ, 2018.

[18] O. Ha, S. Brown, N. Pitterson, “An Exploratory Factor Analysis of Statistics Concept Inventory Data from Practicing Civil Engineers,” *International Journal of Engineering Education*, vol. 33, no. 1, pp. 236-246, 2017.

[19] J.F. Hair Jr., R.E. Anderson, R.L. Tatham, and W.C. Black, “Multivariate Data Analysis with Readings,” 4th ed. Englewood Cliffs: Prentice Hall, 1995.

[20] B. Johnston, *Integrated Chemistry I* [Online]. Available: https://ocw.mit.edu/courses/chemical-engineering/10-490-integrated-chemical-engineering-i-fall-2006/

[21] D. Kahneman, “A Perspective of Judgement and Choice: Mapping Bounded Rationality,” *American Psychologist,* vol. 58, no. 9, pp. 697-720, 2003.

[22] “Key Lessons from the ExxonMobil Baton Rouge Refinery Isobutane Release and Fire,” Chemical Safety Board, Washington, D.C., Rep. 2016-02-I-LA, 2016.

[23] L. Kohlberg, and R.H. Hersh, “Moral Development: A Review of the Theory,” *Theory into Practice*, vol. 16, no. 2, pp. 53-59, 1977.

[24] J.A. Leydens, B.M. Moskal, and M.J. Pavelich, “Qualitative Methods Used in the Assessment of Engineering Education,” *Journal of Engineering Education*, vol. 93, no. 1, pp. 65-72, 2004.

[25] McDowell, and C. Newell, “Measuring Health: A Guide to Rating Scales and Questionnaires,” 2nd ed. New York: Oxford University Press Inc, 1996.

[26] G. McRae, *Integrated Chemistry II* [Online]. Available: https://ocw.mit.edu/courses/chemical-engineering/10-491-integrated-chemical-engineering-ii-spring-2006/

[27] M.B. Miles, A.M. Huberman, J. Saldana, 2014. *Qualitative Data Analysis: A Methods Sourcebook,* 3rd ed. Washington DC: Sage Publication, 2011.

[28] T.M. Osberg, and J.S. Shrauger, "Self-Prediction: Exploring the Parameters of Accuracy," *Journal of Personality and Social Psychology,* vol.51, no. 5, pp. 1044-1057, 1986.

[29] M.A. Pett, N.R. Lackey, J.J Sullivan, *Making Sense of Factor Analysis*, Thousand Oaks: Sage, 2011.

[30] *Recommended Practice 754: Process Safety Indicators for the Refining and Petrochemical Industries,* American Petrochemical Institute, Washington, D.C., 2010.

[31] J. Rest, L. Edwards, and S. Thoma, 'Designing and Validation a Measure of Moral Judgement: Stage Preference and Stage Consistency Approaches," *Journal of Educational Psychology* vol. 89, no. 1, pp. 5-28, 1997b.

[32] J. Rest, D. Narvaez, M. Bebeau, and S. Thoma, "A neo-Kohlbergian approach: The DIT and schema theory," *Educational Psychology Review,* vol. 11, no. 4, pp. 291-324, 1999b.

[33] J. Rest, D. Narvaez, and S. Thoma, "DIT2: Devising and Testing a Revised Instrument of Moral Judgement," *Journal of Educational Psychology,* vol. 91, no. 4, pp. 644-659, 1999a.

[34] J. Rest, S. Thoma, D. Narvaez, and M. Bebeau, "Alchemy and Beyond: Indexing the Defining Issues Test," *Journal of Educational Psychology,* vol. 89, no. 3, pp. 498-507, 1997a.

[35] J. Schonrock-Adema, M. Heijne-Penninga, E.A. Van Hell, J. Cohen-Schotanus, "Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example," *Medical Teacher,* vol. 31, no. 6, pp. 226-232, 2009.

[36] T.D. Sadler, and D.L. Zeidler, "Patterns of Informal Reasoning in the Context of Socioscientific Decision Making," *Journal of Research in Science Teaching*, vol. 42, no. 1, pp. 112-138, 2005.

[37] D.C. Shallcross, "Safety Education through Case Study Presentations," *Education for Chemical Engineers*, vol. 8, no. 1, pp. 12-30, 2013.

[38] D.C. Shallcross, "Safety Shares in the Chemical Engineering Classroom," *Education for Chemical Engineers*, vol. 9, no. 4, pp. 94-105, 2014.

[39] K.E. Stanovich, and R.F. West, "Individual Differences in Reasoning: Implications for the Rationality Debate," *Behavioral & Brain Sciences*, vol. 23, no. 5, pp. 645-665, 2000.

[40] "T2 Laboratories, Inc. Runaway Reaction," Chemical Safety Board, Washington, D.C., Rep. 2008-3-I-FL, 2009.

[41] M. Tavakol, R. Dennick, "Making Sense of Cronbach's Alpha," *International Journal of Medical Education*, vol. 2, pp. 53-55, 2011.

[42] A.E. Tenbrunsel, and D.M. Messick, "Ethical Fading: The Role of Self-Deception in Unethical Behavior," *Social Justice Research,* vol. 17, no. 2, pp. 223-236, 2004.

[43] J. Walther, A.L. Pawley, and N.W. Sochacka, “Exploring Ethical Validation as a Key Consideration in Interpretive Research Quality Paper,” in ©American Society for Engineering Education, 2015. doi: 10.18260/p.24063.

[44] J. Walther, N. Sochacka, and N.M. Kellam, “Quality in Interpretive Engineering Education Research: Reflections on an Example Study,” *Journal of Engineering Education*, vol. 102, no. 4, pp. 626-659, 2013.

[45] J.A. Woodzicka, and M. LaFrance “Real versus Imagined Gender Harassment,” *Journal of Social Issues*, vol. 57, no. 1, pp. 15-30, 2001.

[46] J.T. Youn, and S.A. Choi, “Factor Analysis for Women in Engineering Education Program to Increase the Retention Rate of Female Engineering Students,” *International Journal of Applied Engineering Research*, vol. 11, no. 8, pp. 5657-5663, 2016.

[47] Q. Zhu, et al., “The Development of an Instrument Assessing Individual Ethical Decision-making in Project-based Design Teams: Integrating Quantitative and Qualitative Methods” in ©American Society for Engineering Education, 2014. doi: https://peer.asee.org/23130