CREATING A DATA ANALYSIS PLAN FOR THE COMMUNICATION AND TEAMWORK SKILLS ASSESSMENT TOOL TO MEASURE THE IMPACT OF HIGH FIDELITY INTERPROFESSIONAL EDUCATION

by © Matthew Brake

A practicum report submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of
Master of Nursing
School of Nursing
Memorial University of Newfoundland

August 2018

St. John's, Newfoundland and Labrador

Abstract

Background: The impact of high fidelity interprofessional education (HF-IPE) on fostering teamwork and communication among undergraduate nursing, medicine, and pharmacy students has not been well established. The Communication and Teamwork Skills (CATS) assessment tool is one research instrument that could be used to measure the impact of HF-IPE on teamwork in undergraduate health sciences students. **Purpose:** The purpose of this research practicum project was to demonstrate advanced nursing competencies by participating in the data analysis phase of the research process and developing a data analysis plan for the CATS. **Methods:** Four methods were used to accomplish the purpose of the practicum including: conducting a comprehensive literature review; consulting with a statistician and a nurse researcher; developing the data analysis plan with SPSS codebooks, and testing the plan using a fictitious data set. **Results:** The data analysis plan developed for this practicum project was implemented successfully to analyze, summarize, interpret and display fictitious quantitative data from the CATS. The Paired t-test was selected as an appropriate statistical measure to determine differences between groups' mean scores. Methods to organize, analyze and visually display the data are recommended including a high and low closed chart, bar graphs, and tables. **Conclusion:** This practicum project demonstrated the achievement of advanced nursing competencies by developing a data analysis plan that could be used to guide the analysis of the quantitative data collected using the CATS assessment tool.

Key Words: data analysis plan, communication and teamwork, high fidelity simulation, interprofessional undergraduate education

Acknowledgements

I would first like to thank my practicum supervisor, Dr. Sandra MacDonald of the Nursing Faculty at Memorial University of Newfoundland. Her compassion, patience, and guidance throughout this whole process made it an enjoyable and rewarding experience and one I will not forget. Her commitment to research and teaching was evident in every interaction and her desire to see others succeed did not go unnoticed or unappreciated.

I would also like to thank Joanne Smith-Young, Research Coordinator at the School of Nursing at Memorial University of Newfoundland. Her ability to provide encouragement and different perspectives on research methodologies was truly appreciated. I will not forget her ability to emphasize the importance of looking at problems through a critical lens as it is a crucial component of the research process.

Finally, I would like to thank my wife, my children, and my friends for their encouragement, support, and guidance throughout this entire Master of Nursing program. My wife, Janet, was a driving force over the last three years and I was only able to accomplish this task because of her unwavering commitment and support. My children, George and Meredith, motivate me in a way that they will never understand. Without my children I would not be where I am today.

Table of Contents

Appendices

Introduction

Interprofessional education (IPE) is an effective teaching and learning strategy that is used to foster and develop teamwork and communication skills in both academic (Speakman, 2016; World Health Organization, 2010) and professional settings (Weaver et al., 2010; Van Schaik, Plant, Diane, Tsang, & O'Sullivan, 2011). However, the impact of high fidelity interprofessional education (HF-IPE) on fostering communication and teamwork skills at the undergraduate level has not been well established. Angelini (2011) believed the current uniprofessional nature of academic curriculums has led to attitudes of professional hierarchy and a sense of professional competitiveness, which can ultimately disrupt effective teamwork behaviours. While undergraduate health science curriculums have traditionally included a variety of clinical and simulation exercises, many of these curricula are uniprofessional in their program delivery (Angelini, 2011; Garbee et al., 2012; Garbee et al., 2013). Therefore, it is critical to promote and evaluate new and innovative approaches to IPE, such as the use of high fidelity simulation as a teaching and learning strategy that could potentially foster positive communication and teamwork skills in academic and clinical practice settings.

The World Health Organization (WHO) (2010) asserted that fostering teamwork and communication skills is crucial to ensuring positive patient outcomes, such as patient safety and quality care. In order to achieve this goal, the WHO recommends that teaching and learning strategies that focus on IPE should be initiated at the undergraduate health sciences level and also be supported within clinical practice settings. IPE initiated at the undergraduate level can lead to a positive interprofessional teamwork environment. A positive teamwork environment and effective communication skills can potentially increase patient safety, decrease clinical mistakes, increase patient satisfaction, decrease nursing turnover, and decrease mortality

(Manser, 2009; Sorbero, Farley, Mattke, & Lovejoy, 2008; Weaver et al., 2010; Zangaro & Soeken, 2007). Conversely, ineffective teamwork and poor communication can potentially lead to clinical inefficiencies, an increased waste of clinical supplies, delayed procedures, procedural errors, poorer patient outcomes, and dissatisfaction among team members (Aebersold, Tschannen, & Sculli, 2013; Lingard et al., 2004; Mazzocco et al., 2009).

While there is a growing research database surrounding HF-IPE at the undergraduate level (Dillon et al., 2009; Garbee et al., 2012; Garbee et al., 2013; Jakobsen et al., 2018; Paige et al., 2014; Reese, Jeffries, & Engum, 2010; Smithburger, Kane-Gill, Kloet, Lohr, & Seybert, 2013; Tofil et al., 2014) the consensus within the research community is that further research is needed to measure the effectiveness of HF-IPE to foster communication and teamwork skills. Exposing health sciences students to HF-IPE early in their undergraduate curriculums could lead to effective collaborative practices when they enter the workforce (Dillon, Noble, & Kaplan, 2009; Garbee et al., 2012; Garbee et al., 2013; King et al., 2014). However, further research in this area will require the use of valid and reliable instruments and the creation of data analysis plans for each instrument. One such instrument is the Communication and Teamwork Skills (CATS) assessment tool, which could be used to measure the impact of HF-IPE on communication and teamwork behaviors.

Purpose of Practicum Project

The purpose of this practicum project was to participate in the data analysis phase of the research process by creating a data analysis plan for the quantitative data obtained from the CATS assessment tool (Frankel, Gardner, Maynard, & Kelly, 2007) being used in the study titled "*Measuring the Effectiveness of High Fidelity Simulation in Interprofessional Education to Foster Teamwork Among Undergraduate Nursing, Medicine and Pharmacy Students*"

(MacDonald et al., 2016). This practicum project provided an opportunity to develop advanced nursing practice skills by taking part in nursing clinical, leadership, research, and collaborative activities. Specifically, this practicum project provided an opportunity to consult with MacDonald et al.'s research team to create a data analysis plan for the quantitative data collected from one of the research instruments being used in that study. The MacDonald et al. research team has used the data analysis plan created for this practicum project to analyze the data collected from the CATS assessment tool to measure the effectiveness of HF-IPE to foster communication and teamwork skills.

The objectives for this practicum project included:

1. Demonstrate advanced nursing practice competencies through clinical, leadership, research, and collaborative activities.
2. Analyze, synthesize, and interpret nursing research knowledge as it relates to quantitative data analysis.
3. Analyze and synthesize nursing research knowledge as it relates to the CATS assessment tool, communication and teamwork behaviors, and high fidelity interprofessional education, into a comprehensive literature review.
4. Create a data analysis plan for the CATS assessment tool that is congruent with the objectives of the research study.
5. Analyze quantitative data collected using the CATS and interpret those results.
6. Identify patterns within the data analysis of the quantitative data, and identify why those patterns are important to nursing research.
7. Disseminate the findings of the practicum.

Methodology

Four primary methods were used to successfully achieve the purpose and learning objectives outlined for this practicum including: a comprehensive literature review with literature summary tables (Appendix A and B); consultations with key informants; the creation of the data analysis plan including proof of concept report (Appendix C); and the creation of the SPSS codebooks (Appendix D). The comprehensive literature review focused on teamwork assessment tools that measured communication and teamwork behaviors in undergraduate education and the evaluation of high fidelity interprofessional education, to garner a greater understanding of the topic of interest while also identifying gaps and limitations within the research.

Three consultations were conducted with key informants to ensure this practicum project developed an appropriate data analysis plan for MacDonald et al.'s (2016) study. Consultations occurred via face-to-face, emails, and in telephone conversations with Dr. Variyath a statistician and a faculty member with the mathematics department at MUN; a graduate mathematics student at the Statistics Help Centre at MUN, and Joanne Smith-Young a member of the MacDonald et al. research team and Research Coordinator at MUNSON Nursing Research Unit. The development of the data analysis plan included the creation of two SPSS codebooks to organize and analyze the data set. The data analysis plan was tested using a fictitious data set and recommendations for implementing and evaluating the data analysis plan were discussed. This practicum project demonstrated the achievement of advanced nursing competencies by developing a data analysis plan that can be used to guide the analysis of the quantitative data collected using the CATS assessment tool. Important components of each of these methods will be discussed and integrated into this final practicum report.

Summary of Literature Review

Search Methodology

The search of the literature included searching the databases of CINAHL (2006 to January 2018), PubMed (2006 to January 2018), and Google Scholar (2006 to January 2018). Key words and phrases used while searching those databases included: communication and teamwork skills, CATS, teamwork, high fidelity interprofessional education, interprofessional education, simulation, undergraduate students, and data analysis plan. The parameters from 2006 to 2018 were set to reflect current, relevant research, while also incorporating all research that has been published on the CATS assessment tool. A lateral search was conducted using the "similar article" function present on CINAHL, PubMed, and Google Scholar. An additional lateral search was also conducted searching reference lists of collected articles. A final lateral search was completed using Google to collect gray literature focused on these areas of interest. Once a relevant article was identified, the abstract was scanned for key words and phrases. If applicable, the entire article was reviewed and critiqued. A total of 52 articles/resources were part of the comprehensive literature review, which included 25 research studies, eight systematic/literature reviews, eight reports, seven textbook sources, and four grey literature sources.

Questions used to guide the literature review included:

1. Has the CATS assessment tool been used to measure communication and teamwork behaviours among undergraduate nursing, medicine, and pharmacy students?
2. Has the CATS assessment tool been used to measure teamwork behaviours in HF-IPE?

3. Is there a relationship between HF-IPE and communication and teamwork behaviours in nursing, medicine, and pharmacy students?

Communication and Teamwork Skills Assessment Tool

Frankel et al. (2007) created the CATS assessment tool to observe and document the communication and teamwork skills of healthcare teams in the real world and in simulated settings. The CATS assessment tool has been used to assess the communication and teamwork skills displayed by a broad range of healthcare professionals and undergraduate students in nursing, medicine, social work, and respiratory therapy (Frankel et al., 2007; Garbee et al., 2012; Garbee et al., 2013; Hughes et al., 2014; Joshi, Hernandez, Martinez, AbdelFattah, & Gardner, 2017; Passauer-Baierl, Baschnegger, Bruns, & Weigl, 2014; Smithburger et al., 2013). The CATS assessment tool focuses on directly observing communication and teamwork behaviours while quantitatively gathering data on the quality of the observed behaviours. Frankel et al. wanted to develop a quantitative assessment tool that focused on how often and how well particular communication and teamwork behaviours were performed.

The CATS assessment tool investigates four domains of teamwork behaviour: situational awareness, coordination, communication, and cooperation. Within these four domains, there are 21 behaviour markers that are assessed by a trained observer, including three behaviour markers that are only scored if a crisis situation arises. Specific behaviour marker scores need to be combined to determine each respective domain score. For example, the coordination domain is comprised of the following behaviour markers: briefing, verbal plan, verbalize expected outcomes, debriefing, and establish event manager. Behavior markers are scored on the basis of how often an event occurs and the quality of the team's communication and teamwork behaviours.

Each time a behaviour is observed it produces a raw data score as either "*Good*" = 1 point; "*Variable in Quality*" = 0.5 points, or "*Expected but Not Observed*" = 0 points, under the appropriate behaviour marker. The raw data under each behaviour marker is subsequently used to determine raw scores for each of the four domains, and as an overall score. The raw data collected using the CATS assessment tool is initially calculated into non-weighted total scores. The non-weighted total scores need to be further calculated into weighted total scores. The weighted total score out of 100 is calculated for each individual behaviour marker, each domain, and as an overall score. The weighted total scores can be used to compare team performances either between teams, or pre and post an intervention, or across two different testing conditions such as HF-IPE and low fidelity interprofessional education (LF-IPE). The data collected using this tool is considered ratio level data.

Frankel et al. (2007) believed the CATS assessment tool was appropriate to assess communication and teamwork skills in a variety of healthcare settings. The CATS assessment tool provides a unique opportunity for a trained observer to evaluate team behaviour without focusing on individual behaviour or performance, as behaviours are catalogued and analyzed from an overall-team perspective only. From a research standpoint, it appears that the CATS assessment tool can be used to gather data in a variety of settings, with an overall goal that focuses on understanding the communication and teamwork skills of a given team.

In total, only six research studies were retrieved that used the CATS assessment tool as part of their data collection methods. These studies have been performed in both professional settings (Hughes et al., 2014; Joshi et al., 2017; Passauer-Baierl et al., 2014) and academic settings (Garbee et al., 2012; Garbee et al., 2013; Smithburger et al., 2013). Within professional settings, research using the CATS tool has focused on designing interprofessional programs to improve teamwork skills (Hughes et al., 2014), investigating the teamwork skills within stable and dynamic teams (Joshi et al., 2017), and assessing teamwork and communication skills within an operating room setting (Passauer-Baierl et al., 2014). Only Joshi et al. used the CATS to assess teamwork in a HF-IPE scenario within a professional setting. Furthermore, these studies made no inferences regarding HF-IPE experiences during health sciences undergraduate education and whether or not early HF-IPE could lead to positive benefits when professionals enter the workforce. This would suggest there is a need for appropriate research instruments to measure the impact of HF-IPE within academic settings.

Within academic settings, researchers believe that HF-IPE is an effective teaching and learning strategy for undergraduate education and it has been shown to enhance the development of effective communication and teamwork skills that students can use when they enter the workforce (Garbee et al., 2012; Garbee et al., 2013). Smithburger et al. (2013) also determined that HF-IPE sessions over time can lead to a statistically significant improvement in communication and teamwork scores. The difficulty associated with comparing these research studies is the IPE teams consisted of different health sciences students from different disciplines. There is a need for more research to measure the impact of HF-IPE in undergraduate health sciences education.

It is clear from this comprehensive literature review that there is a lack of research using the CATS assessment tool to measure the impact of HF-IPE with undergraduate health sciences students and no studies were found with teams exclusively using nursing, medicine, and pharmacy. Furthermore, no research has been conducted examining communication and teamwork behaviours in high or low fidelity simulation using the CATS assessment tool. This would indicate the need for further research using the CATS in this area. Please refer to the comprehensive literature review in Appendix A for an expanded description of the research conducted using the CATS assessment tool and an analysis of the strengths and limitations of the CATS assessment tool.

Communication and Teamwork in High Fidelity

The majority of studies related to communication and teamwork in high fidelity simulation in academic settings focused primarily on self-perception of communication and teamwork behavior (Dillon et al., 2009; Jakobsen et al., 2018; King et al., 2014; Paige et al., 2014; Reese et al., 2010). These studies did identify increases in self-perception and confidence in communication and teamwork behaviours as they related to HF-IPE. However, using self-perception as a form of data collection could be considered a limitation due to the fact that overestimation or underestimation of abilities can occur (Havyer et al. 2016; Paige et al., 2014). While self-reporting does provide insightful information surrounding how participants feel regarding their HF-IPE experiences, it does not provide any concrete evidence surrounding their knowledge acquisition, communication and teamwork behaviours. None of these studies measured the long-term impact that HF-IPE participation can have on both communication and teamwork behaviours.

Paige et al. (2014) completed a HF-IPE study with health science students that included observed team behaviour scores, but the CATS was not used for that study. Those observed behaviour scores were completed using a data collection tool that was specifically designed by the researchers to measure operating-room teamwork. Paige et al. determined that HF-IPE led to statistically significant gains ($p < 0.001$) in all subscales of the team behaviour assessment tool. Paige et al. asserted that HF-IPE can be a feasible and effective method of education delivery that can have an immediate impact on participants' perceived and observed teamwork behaviours. HF-IPE is often the preferred environment for high-stakes medical training as they provide a safe space where teamwork skills and task-orientated skills can be performed (Hunt, Fiedor-Hamilton, & Eppich, 2008; Scheckel, 2016). Benefits from participating in HF-IPE include: increasing knowledge, improving patient outcomes, increasing skill competency, and increasing appropriate clinical behaviours (Cook et al., 2011).

It is clear that there is a lack of research focused on assessing HF-IPE using objective-based, observer-focused, data collection instruments such as the CATS assessment tool. Objective measurement tools that analyze data collected on observed teamwork behaviours - such as the CATS assessment tool - could help document a more precise result surrounding the impact of HF-IPE on communication and teamwork behaviours. Please refer to the literature review in Appendix A for an expanded description of the research conducted surrounding HF-IPE and communication behaviours in academic settings.

Barriers to Implementing High Fidelity Interprofessional Education

HF-IPE can be a feasible and effective method of education delivery that can have an immediate impact on participants' perceived and observed teamwork behaviours, however, there are few studies that measure the impact on behaviour and it is difficult to infer whether those changes would transfer to real-life clinical settings. Van Schaik et al. (2011) believed that while HF-IPE can be beneficial for participants, limitations and barriers exist surrounding the implementation of these programs including: difficulty coordinating the participant's schedules, high cost for set up and maintenance of the human patient simulators, and difficulty in recreating real-life work environments. Van Schaik et al. made reference to these limitations as they related to HF-IPE and working professionals, but these limitations are also present when planning for HF-IPE in undergraduate health science curricula (Lapkin, Levett-Jones, & Gilligan, 2012).

Newton et al. (2015) believed IPE in academic settings is often limited by a lack of flexibility in undergraduate curricula, limited shared free time across various academic disciplines, resource constraints, space constraints, economic constraints, and a lack of faculty development regarding IPE. Conversely, others believed HF-IPE can be feasible given a large enough target population and the proper teaching environment (Jakobsen et al., 2018; Paige et al., 2014). A large-scale cost benefit analysis surrounding HF-IPE within health science academic programs could provide vital information regarding whether or not HF-IPE is a cost effective endeavour within these undergraduate programs.

Summary of Consultations

Polit and Beck (2017) asserted that consultations with experts in a particular area are an integral part of the research design process. Consultations occurred with Dr. Variyath, a statisticians and a faculty member with the mathematics department at MUN and a graduate student at the Statistics Help Centre at MUN. Consultations also took place with the Research Coordinator at MUNSON, Nursing Research Unit. These consultations were considered a vital part of the data analysis plan process, as these experts can often play an integral role in ensuring the statistical tests chosen are congruent with the research questions being asked (Planter, 2011; Simpson, 2015). Consultations were completed as part of this practicum project to ensure that the data analysis plan was developed properly while also ensuring the data analysis plan effectively answered the research questions.

Consultation with Statisticians

It is believed that statisticians can assist with determining a thorough statistical analysis plan that can help control for confounding variables (Chasan-Taber, 2014; Polit & Beck, 2017).

Dr. Variyath, a statistician and faculty member with the mathematics department at MUN was identified as a person of interest due to his experience in quantitative data analysis. The consultation with Dr. Variyath was vital to ensure the research questions being asked would be properly addressed within the data analysis plan. Prior to meeting Dr. Variyath, Simpson's (2015) decision tree was utilized to determine what inferential statistic test would be appropriate to use, given the context of the research question and the data collected. Please refer to Appendix E for a diagram outlining the path along Simpson's decision tree.

The Simpson's (2015) decision tree identified the Paired t-test as the most appropriate test, given the context of the data collected and the research questions being asked. Dr. Variyath also agreed that the Paired t-test was most appropriate test given the context of this project and the assumptions being made regarding the data collected. In order to obtain further assurance that the statistical methods chosen were correct given the context of the research design and the research questions, consultations took place with the staff at the Statistics Help Centre at MUN where it was confirmed that a Paired t-test would provide the intended results. Having two separate individuals with statistic expertise confirm that the Paired t-test was the appropriate test for this data analysis plan provided reassurance that the data analysis plan for this practicum project would produce the intended results.

Consultation with Nurse Researcher

Joanne Smith-Young, the Research Coordinator at MUNSON Nursing Research Unit was also consulted as an expert in the field who could provide valuable information regarding the data analysis of the CATS assessment tool, while also providing insight regarding how to properly set up an SPSS codebook. Consultations with Joanne focused on various topics including the limitations of the CATS assessment tool and its ability to guide the collection of data, and SPSS generalities as it related to codebook writing. Joanne helped to confirm that the SPSS codebooks created as part of the data analysis plan for this practicum project were correct and would produce the desired results. There was a lot of discussions regarding whether or not to create a single codebook instead of two codebooks, as a way of limiting potential manual transcription errors.

Joanne also played a vital role in pinpointing potential weaknesses within the SPSS codebooks and potential limitations within the data analysis plan. For example, Joanne pointed out that having to manually transcribe data from one codebook to another could lead to a transcription error. These human errors could influence the data and lead to incorrect results. This form of transcription error was also discussed in the literature. A duplicate data entry methodology - where two people enter the data electronically and discrepancies are flagged and corrected - would be ideal when performing data entry to prevent manual transcription errors. Wahi, Parks, Skeate, and Goldin (2008) asserted that duplicate data entry can decrease transcription errors when compared to single data entry, but operational constraints are a major limitation when trying to implement this practice. Similarly, within this practicum project a duplicate data entry method would not be possible. As stated previously, it was clear that the data analysis plan using two codebooks may increase transcription errors and the potential for errors must be taken into consideration when transcribing the data.

During the consultation process, Joanne provided valuable information regarding the context of the research study and her insights into the CATS assessment tool. From the consultation with Joanne, there was increased clarity surrounding MacDonald et al.'s (2016) research study and the role this practicum project would have as it related to the CATS assessment tool. Joanne also provided valuable information regarding SPSS, how to write codebooks to meet their desired outcomes, and how to be cognizant of potential weaknesses within a data analysis plan or within a SPSS codebook.

Consultation Impact on Practicum Project

Both consultations were instrumental in the development of the content for this practicum project. The statistician and mathematics graduate student provided reassurance that the Paired t-test would produce the intended results as part of the data analysis plan. Joanne provided clarity surrounding the variables being measured in the research study, the application of the CATS assessment tool to collect and analyze communication and teamwork data, and how to create SPSS codebooks to meet the needs of the CATS assessment tool. While the majority of those discussions focused on creating a single codebook as opposed to using two codebooks, it was determined that given the context of the data and the research questions being asked, two codebooks allowed for an easier process with regards to organizing and analyzing the data. However, the fact remained that when there are multiple junctures where manual transcription is necessary, there is the potential for transcription errors. Strategies used to mitigate these potential errors can be found in Appendix C as part of the data analysis plan.

Summary of Data Analysis Plan

The data analysis plan developed for this practicum project focused on the analysis of ratio level data that would be collected using the CATS assessment tool. The data analysis plan guides the evaluation of communication and teamwork behaviours of nursing, medicine, and pharmacy students working within an interprofessional team during a clinical simulation. As discussed in the consultation section, two different SPSS codebooks were created for this data analysis plan. The first codebook is used to input the raw data collected using the CATS assessment tool and compute the non-weighted total scores, which would subsequently be used to calculate the weighted total scores for each of the 21 behaviour markers, the four domains, and as an overall score. The second codebook is used to organize the weighted total scores into their respective high fidelity and low fidelity scenarios, which will allow for the statistical analysis to occur. Within the second codebook, the weighted total scores will be separated for all 26 variables in order to allow for analysis of the data using the Paired t-test.

Due to the vast differences in research methodologies previously used with the CATS assessment tool, there is no consensus regarding how to analyze data collected using this tool. Previous statistical analysis methods have included such tests as: ANOVA with Bonferroni (Smithburger et al., 2013), chi-square and Fisher's exact (Hughes et al., 2014), Independent sample t-test (Joshi et al., 2017), and Paired sample t-test (Garbee et al., 2012; Garbee et al., 2013). Since no consensus was present, the data analysis plan and statistical analysis methods chosen for this practicum project had to be curated to meet the specific needs of the research questions being asked. Through consultations with the statistician and the Statistics Help Centre at MUN, it was deemed appropriate that the Paired t-test would produce the desired results given the context of the research question and the scope of this practicum project. This appropriateness was reinforced by the fact that two previous studies (Garbee et al., 2012; Garbee et al., 2013) used the same statistical analysis test to analyze data collected using the CATS assessment tool.

Proof of Concept

It was decided for the purpose of this practicum project that a fictitious data set would be generated and used to test "proof of concept" based on the research design of MacDonald et al., (2016) which compares participation in a HF-IPE scenario with participation in a LF-IPE scenario. Data for a fictitious sample of seven (n = 7) teams was generated and scored using the CATS assessment tool. The fictitious data set was used to ensure the SPSS codebooks were designed correctly and produced the desired results as if real collected data were to be inputted into the files. Since the data entered was fictitious, it would be imprudent to draw inferences regarding what the results could mean as it related to communication and teamwork behaviours and HF-IPE, or relate these results back to the findings in the literature review. Please see Appendix C for a full report on the creation of the CATS data analysis plan and proof of concept exercise. This section of the practicum report will not discuss the specific implications of the findings from the fictitious data set, rather it will only discuss how similar findings could be interpreted if the collected data produced similar results.

As stated previously, it was determined through consultations and the use of a decision tree created by Simpson (2015) that the Paired t-test would be an appropriate test to compare the communication and teamwork scores as observed in the HF-IPE to the scores in the LF-IPE. It is important to note that within this practicum project it was assumed that the data used was normally distributed, the groups were equal, and participants were randomly assigned to different teams. For a more detailed description surrounding the data generation, the data input, the equations generated for this practicum project, and how potential transcription errors were mitigated during the data analysis process, please refer to Appendix C. For a more detailed explanation regarding the SPSS codebooks please refer to Appendix D.

An alternate hypothesis and null hypothesis were established as part of this practicum project to guide the analysis of the fictitious data. The null hypothesis would postulate that HF-IPE and LF-IPE would produce the same scores when assessed using the CATS assessment tool. The alternate hypothesis would postulate that HF-IPE would produce a higher quality score using the CATS assessment tool when compared to the LF-IPE scores. The confidence intervals were set at 95% with the level of significance having a *p* value < 0.05. A *p* value set at this significance level would mean that the likelihood of the differences detected between the scores would emerge due to chance only 5% of the time (Knapp, 2016). A *p* value significance level set at less than 0.05 and confidence intervals set at 95% are considered the standard parameters used for many research studies (Polit & Beck, 2017). The *p* value provides valuable information that allows for either a rejection or acceptance of the null hypothesis and alternate hypothesis.

The proof of concept of the data analysis plan was completed to demonstrate how a fictitious data set from the CATS could be analyzed and visually displayed using a high and low closed chart, bar graphs, as well as tables to display group mean scores, standard deviations and statistical significance. The Paired t-test was successfully used to determine whether or not there were statistically significant differences between HF-IPE and LF-IPE scores in the 21 behaviour markers, the four domains, and the overall score. Any significant differences between these scores would indicate a significant difference between the communication and teamwork skills displayed in the HF-IPE as compared to the LF-IPE.

Unfortunately, the small sample size of the fictitious data set created for this practicum project (n = 7) may have had a significant influence on the standard deviation of the values and thus may have affected the volatility of the data. Knapp (2016) believed that while the Paired t-test can be completed on any sample size, for it to be considered robust the sample size should be greater than 30. Future research using a larger sample size could produce results that would be considered more robust than this fictitious data set.

If this data analysis plan was applied to a real collected data set, it could provide researchers with a good direction to analyze the quantitative data collected using the CATS assessment tool. That analysis could help to determine whether or not communication and teamwork behaviors are displayed differently in a HF-IPE scenario as compared to a LF-IPE scenario.

Discussion and Interpretation of the Plan

If a real data set produced similar results to this fictitious data set, it would be clearly evident that participation in HF-IPE produces stronger communication and teamwork behaviour for nursing, medicine, and pharmacy students, as compared to participation in LF-IPE.
This practicum project data analysis plan and proof of concept exercise demonstrated that given a collected data set, the SPSS files created could be successful in organizing and analyzing the quantitative data collected using the CATS assessment tool.

Advanced Practice Competencies

From an advanced nursing practice perspective, this practicum project has provided opportunities to perform tasks to develop skills within each of the four advanced practice nursing competencies. The following sections will provide examples where tasks performed within this practicum project fall within each competency.

Clinical

From a clinical competency perspective, the results obtained from using the data analysis plan for this practicum project as part of MacDonald et al.’s (2016) research study could provide an opportunity to incorporate new nursing knowledge into the development of future nursing curriculum. The Canadian Nurses Association (2008) believed using new nursing knowledge to guide program and policy development was an advanced nursing practice clinical competency. The findings from the MacDonald et al. study may reveal HF-IPE could be used to guide future undergraduate program development to ensure HF-IPE opportunities are provided to undergraduate nursing, medicine, and pharmacy students. The data analysis plan created for this practicum project provided a roadmap for the research team of MacDonald et al., to analyze the data that could be part of a driving force for policy change to influence the design of future undergraduate health sciences curriculum to include HF-IPE.

Clinical competency was also demonstrated within this practicum project by completing a comprehensive literature review that identified and assessed research trends as they related to HF-IPE and health sciences students. This practicum project has also produced results that helped to identify limitations and gaps within the literature. These limitations and gaps could be considered as a starting point for future research that focuses on the impact of HF-IPE in nursing, medicine, and pharmacy undergraduate programs to foster teamwork and communication skills.

Performing a comprehensive literature review on a research tool not commonly used in undergraduate education provided a unique opportunity to use clinical judgment and decision-making to extrapolate those findings and apply them to the instrument used within this practicum project. While this practicum project focused on the creation of a data analysis plan, learning how to use the research tool to collect the data was also integral to fully understanding how to create the data analysis plan for the tool. This practicum project provided an opportunity to contribute to enhancing nursing knowledge surrounding the CATS and HF-IPE, which would allow for the future advocacy for interprofessional activities both within academic and professional settings. This advocacy could potentially lead to direct improvements in care within a broad range of clinical settings.

Leadership

From a leadership competency perspective, this practicum project provided an avenue to take initiative to partake in a stream of the Master of Nursing program that is not common among nursing graduate students. Choosing to perform a research-based practicum project allowed for leadership to unfold within promotion of this stream as a viable option for future students, and to uniquely contribute to the growing database of practicum projects completed by Master of Nursing students at MUN. Other leadership competencies emerged within the consultation phase of this practicum, during critical discussions with nurse researchers related to the benefits and limitations of selecting a valid and reliable data collection instrument. Those discussions allowed for a critique of the literature surrounding the CATS and a sharing of knowledge that contributed to a greater understanding of the CATS and how it could be used within nursing research. Having developed an enhanced understanding of research instruments and data analysis plans has promoted leadership competencies with regards to advocating for the use of valid and reliable instruments for nursing research and the development of data analysis plans to guide the research process. Leadership competencies were also developed by enhancing knowledge of the research conducted with the tool, thus increasing confidence in sharing that knowledge regarding how the CATS assessment tool could be used within a variety of nursing research studies.

Research

The Canadian Nurses Association (2008) believed that being an active participant in the generation and utilization of nursing research was central to advanced nursing practice. This practicum project provided ample opportunities to perform tasks that could be considered advanced nursing practice from a research competency perspective. Collaboration with mathematics faculty and experienced nurse researchers was an integral component of this practicum project. Research competency was displayed by developing the data analysis plan, analyzing a fictitious data set, and creating a data analysis report based on the fictitious data set. While the fictitious data was generated with the sole purpose of proving the plan could work, the analysis of the data did show that the plan, including the SPSS codebooks, and the suggested method of presenting the findings could be successfully used to analyze data from the CATS assessment tool. The data analysis report created for this practicum project further demonstrated the research competency of interpreting data and how it relates within the context of teamwork and communication skills within HF-IPE.

The comprehensive literature review provided another avenue to complete advanced nursing practice within the research competency. Conducting the literature review allowed for a thorough critique of the previous literature on the topic. The focus of the review was on the quality of information, the content of previous research, limitations within previous research, and gaps within the literature. The literature review also provided an opportunity for interpretation of research findings as they related to the CATS assessment tool, and confirmation of the limited number of research articles using the CATS to measure communication and teamwork skills in undergraduate HF-IPE. The literature summary tables (Appendix B) provide a good example of critiquing literature for the purpose of gathering a greater understanding of the topic of interest.

Interpretation of data was demonstrated during the data analysis component of the practicum project while completing the data analysis report. The SPSS codebooks yielded statistical data that had to be categorized regarding significance level, and interpreted to determine what the findings showed with regards to teamwork and communication skills.

Despite the data set being fictitious, the interpretation process would have been the same if a real data set had been used.

Dissemination is also an integral part of the research process and considered an advanced nursing practice skill. A PowerPoint presentation to select faculty and peers provided an opportunity to disseminate information relating to the comprehensive literature review, the creation of the data analysis plan, and interpreting the results within the data analysis report. The focus of that presentation was on the need for data analysis plans in nursing research. Again, despite using fictitious data, the dissemination component of this practicum project would be identical if real collected data had been used, except the presentation would be directed at the research team and faculty.

Consultation and Collaboration

Collaboration and consultation are considered integral skills that advanced practice nurses demonstrate and utilize in their nursing practice (Canadian Nurses Association, 2008). Advanced nursing practice within this competency was achieved by performing timely and appropriate consultations with statisticians and nurse researchers. Telephone conversations, face-to-face interactions, and email correspondences were all modalities used to complete consultations with individuals who were identified as experts in their respective fields, and who would contribute greatly to the success of this practicum project. Collaboration was demonstrated with the sharing of knowledge with the research coordinator regarding the CATS assessment tool as it related to MacDonald et al.'s (2016) study and this practicum project. Interpersonal relationships are integral to the consultation and collaboration process and this practicum project provided many opportunities to develop these relationships in a way that provided productive interactions while maintaining professional boundaries.

Conclusion

The purpose of this practicum project was to create a comprehensive and thorough data analysis plan that could guide research as it related to communication and teamwork behaviour of nursing, medicine, and pharmacy students as they take part in HF-IPE and LF-IPE scenarios. It is evident that this proof of concept exercise produced the desired results with the data analysis plan having the ability to successfully organize, summarize, and analyze CATS assessment tool data using the appropriate statistical methodology. This practicum project also generated experiences and skills that could be considered advanced nursing practices within the clinical, leadership, research, and collaboration competencies. This practicum project demonstrated how integral a data analysis plan is to research design to ensure the methodologies chosen are congruent with the research questions being asked, while also using appropriate statistical methods to achieve the desired information.

References

Aebersold, M., Tschannen, D., & Sculli, G. (2013). Improving nursing students' communication skills using crew resource management strategies. *Journal of Nursing Education, 52*(3), 125-130. doi:10.3928/01484834-20130205-01

Angelini, D. J. (2011). Interdisciplinary and interprofessional education: What are the key issues and considerations for the future? *The Journal of Perinatal and Neonatal Nursing*, *25*(2), 175-179. doi:10.1097/JPN.0b013e318212ee7a

Canadian Nurses Association. (2008). *Advanced nursing practice: A national framework.* Retrieved from https://www.cna-aiic.ca/~/media/cna/page-content/pdf-en/anp_national_framework_e.pdf

Chasan-Taber, L. (2014). Data analysis plan. *Writing dissertation and grant proposals: Epidemiology, preventive medicine and biostatistics* (pp.179-195). New York, NY: CRC Press.

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. *JAMA, 306*(9), 978-988. doi:10.1001/jama.2011.1234

Dillon, P. M., Noble, K. A., & Kaplan, L. (2009). Simulation as a means to foster collaborative interdisciplinary education. *Nursing Education Perspectives, 30*(2), 87-90. Retrieved from https://journals.lww.com/neponline/

Frankel, A., Gardner, R., Maynard, L., & Kelly, A. (2007). Using the Communication and Teamwork Skills (CATS) assessment to measure health care team performance. *Joint Commission Journal on Quality and Patient Safety, 33*(9), 549-558. doi:10.1016/S1553-7250(07)33059-6

Garbee, D., Paige, J., Bonanno, L., Rusnak, V., Barrier, K., Kozmenko, L., . . . Nelson, T. (2012). Effectiveness of teamwork and communication education using an interprofessional high-fidelity human patient simulation critical care code. *Journal of Nursing Education and Practice, 3*(3), 1-12. doi:10.5430/jnep.v3n3p1

Garbee, D., Paige, J., Barrier, K, Kozmenko, V., Kozmenko, L., Zamjahn, J., . . . Cefalu, J. (2013). Interprofessional teamwork among students in simulated codes: A quasi-experimental study. *Nursing Education Perspective, 34*(5), 339-344. doi:10.5480/1536-5026-34.5.339

Havyer, R. D., Nelson, D. R., Wingo, M. T., Comfere, N. I., Halvorsen, A. J., McDonald, F. S., & Reed, D. A. (2016). Addressing the interprofessional collaboration competencies of the Association of American Medical Colleges: A systematic review of assessment instruments in undergraduate medical education. *Academic Medicine, 91*(6), 865-888. doi:10.1097/ACM.0000000000001053

Hughes, K. M., Benenson, R. S., Krichten, A. E., Clancy, K. D., Ryan, J. P., & Hammond, C. (2014). A crew resource management program tailored to trauma resuscitation improves team behavior and communication. *Journal of the American College of Surgeons, 219*(3), 545-551. doi:10.1016/j.jamcollsurg.2014.03.049

Hunt, E. A., Fiedor-Hamilton, M., & Eppich, W. J. (2008). Resuscitation education: Narrowing the gap between evidence-based resuscitation guidelines and performance using best educational practices. *Pediatric Clinics of North America, 55*(4), 1025-1050. doi:10.1016/j.pcl.2008.04.007

Jakobsen, R. B., Gran, S. F., Grimsmo, B., Arntzen, K., Fosse, E., Frich, J. C., & Hjortdahl, P. (2018). Examining participant perceptions of an interprofessional simulation-based trauma team training for medical and nursing students. *Journal of Interprofessional Care, 32*(1), 80-88. doi:10.1080/13561820.2017.1376625

Joshi, K., Hernandez, J., Martinez, J., AbdelFattah, K., & Gardner, A. K. (2017). Should they stay or should they go now? Exploring the impact of team familiarity on interprofessional team training outcomes. *The American Journal of Surgery*. 1-7. doi:10.1016/j.amjsurg.2017.08.048

King, S., Carbonaro, M., Greidanus, E., Ansell, D., Foisy-Doll, C., & Magus, S. (2014). Dynamic and routine interprofessional simulations: expanding the use of simulation to enhance interprofessional competencies. *Journal of Allied Health, 43*(3), 169-175. Retrieved from http://www.asahp.org/journal-of-allied-health/

Knapp, H. (2016). *Practical Statistics for Nursing Using SPSS.* Los Angeles, CA: SAGE Publications.

Lapkin, S., Levett-Jones, T., & Gilligan, C. (2012). A cross-sectional survey examining the extent to which interprofessional education is used to teach nursing, pharmacy and medical students in Australian and New Zealand universities. *Journal of Interprofessional Care, 26*(5), 390-396. doi:10.3109/13561820.2012.690009

Lingard, L., Espin, S., Whyte, S., Regehr, G., Baker, G. R., Reznick, R., . . . Grober, E. (2004). Communication failures in the operating room: An observational classification of recurrent types and effects. *Quality and Safety in Health Care*, *13*(5), 330-334. doi:10.1136/qhc.13.5.330

MacDonald, S., Manual, A., Dubrowski, A., Bandrauk, N., Law, R., & Curran, V. (2016). *Measuring the effectiveness of high fidelity simulation in interprofessional education to foster teamwork among undergraduate nursing, medicine and pharmacy students.* Teaching and Learning Proposal, Memorial University of Newfoundland.

Manser, T. (2009). Teamwork and patient safety in dynamic domains of healthcare: A review of the literature. *Acta Anaesthesiologica Scandinavica*, *53*(2), 143-151. doi:10.1111/j.1399-6576.2008.01717.x

Mazzocco, K., Petitti, D. B., Fong, K. T., Bonacum, D., Brookey, J., Graham, S., . . . Thomas, E. J. (2009). Surgical team behaviors and patient outcomes. *The American Journal of Surgery*, *197*(5), 678-685. doi:10.1016/j.amjsurg.2008.03.002

Newton, C., Bainbridge, L., Ball, V., Baum, K. D., Bontje, P., Boyce, R. A., . . . Wood, V. (2015). The health care team challenge™: Developing an international interprofessional education research collaboration. *Nursing Education Today, 35*(1), 4-8. doi:10.1016/j.nedt.2014.07.010

Paige, J. T., Garbee, D. D., Kozmenko, V., Yu, Q., Kozmenko, L., Yang, T., . . . Swartz, W. (2014). Getting a head start: High-fidelity, simulation-based operating room team training of interprofessional students. *Journal of the American College of Surgeons*, *218*(1), 140-149. doi:10.1016/j.jamcollsurg.2013.09.006

Passauer-Baierl, S., Baschnegger, H., Bruns, C., & Weigl, M. (2014). Interdisciplinary teamwork in the OR: Identification and measurement of teamwork in the operating room. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen, 108*(5), 293-298. doi:10.1016/j.zefq.2013.05.007

Planter, A. T. (2011). Writing a data analysis plan. In W. Pequegnat, E. Stover, & C. A. Boyce (Eds.). *How to write a successful research grant application: A guide for social and behaviour scientists* (2nd ed.) (pp. 283-298). Boston, MA: Springer US.

Polit, D. F., & Beck, C. T. (2017). *Nursing research: Generating and assessing evidence for nursing practice* (10th ed.). Philadelphia, PA: Walters Kluwer/Lippincott Williams & Wilkins.

Reese, C. E., Jeffries, P. R., & Engum, S. A. (2010). Learning together: Using simulations to develop nursing and medical student collaboration. *Nursing Education Perspectives, 31*(1), 33-37. Retrieved from https://journals.lww.com/neponline/

Scheckel, M. (2016). Designing courses and learning experiences. In D. M. Billings & J. A. Halstead (Eds.). *Teaching in nursing: A guide for faculty* (pp. 159-185). St. Louis, MI: Elsevier.

Simpson, S. H. (2015). Creating a data analysis plan: What to consider when choosing statistics for a study. *The Canadian Journal of Hospital Pharmacy, 68*(4), 311-317. Retrieved from https://www.cjhp-online.ca/

Smithburger, P., Kane-Gill, S., Kloet, M., Lohr, B., & Seybert, A. (2013). Advancing interprofessional education through the use of high fidelity human patient simulators. *Pharmacy Practice, 11*(2), 61-65. Retrieved from https://www.pharmacypractice.org/

Sorbero, M. E., Farley, D. O., Mattke, S., & Lovejoy, S. L. (2008). *Outcome measures for effective teamwork in inpatient care.* Santa Monica: CA. Rand Corporation.

Speakman, E (2016). Interprofessional Education and Collaborative Practice. In D. M. Billings & J. A. Halstead (Eds.). *Teaching in nursing: A guide for faculty* (pp. 186-196). St. Louis, MI: Elsevier.

Tofil, N. M., Morris, J. L., Peterson, D. T., Watts, P., Epps, C., Harrington, K. F., . . . White, M. L. (2014). Interprofessional simulation training improves knowledge and teamwork in nursing and medical students during internal medicine clerkship. *Journal of Hospital Medicine*, *9*(3), 189-192. doi:10.1002/jhm.2126

Van Schaik, S. M., Plant, J., Diane, S., Tsang, L., & O'Sullivan, P. (2011). Interprofessional team training in pediatric resuscitation: A low-cost, in situ simulation program that enhances self-efficacy among participants. *Clinical Pediatrics*, *50*(9), 807-815. doi:10.1177/0009922811405518

Wahi, M. M., Parks, D. V., Skeate, R. C., & Goldin, S. B. (2008). Reducing errors from the electronic transcription of data collected on paper forms: A research data case study. *Journal of the American Medical Informatics Association, 15*(3), 386-389. doi:10.1197/jamia.M2381

Weaver, S. J., Rosen, M. A., DiazGranados, D., Lazzara, E. H., Lyons, R., Salas, E., . . . King, H. B. (2010). Does teamwork improve performance in the operating room? A multilevel evaluation. *The Joint Commission Journal on Quality and Patient Safety*, *36*(3), 133-142. Retrieved from http://www.jointcommissionjournal.com

World Health Organization. (2010). *Framework for action on interprofessional education and collaborative practice*. Retrieved from http://www.who.int/hrh/resources/framework_action/en/

Zangaro, G. A., & Soeken, K. L. (2007). A meta-analysis of studies of nurses' job satisfaction. *Research in Nursing and Health*, *30*(4), 445-458. doi:10.1002/nur.20202

Appendix A

Literature Review

The following literature review is a discussion of the current body of research surrounding the use of the Communication and Teamwork Skills (CATS) assessment tool to measure the impact of high fidelity interprofessional education (HF-IPE) as a teaching and learning strategy with undergraduate health science students. The purpose of this literature review is to analyze, synthesize, and interpret nursing research knowledge as it relates to the development of a data analysis plan for the quantitative data collected using the CATS assessment tool developed by Frankel, Gardner, Maynard, and Kelly (2007). Questions used to guide this literature review included: (1) Has the CATS been used to measure communication and teamwork behaviours among undergraduate nursing, medicine, and pharmacy students? (2) Has the CATS been used to measure teamwork behaviours in HF-IPE? and (3) Is there a relationship between HF-IPE and communication and teamwork behaviours in undergraduate nursing, medicine, and pharmacy students? In particular this literature review will analyze and synthesize nursing research knowledge as it relates to measuring communication and teamwork behaviours in high fidelity interprofessional simulation education, into a comprehensive literature review that will be used to inform the development of a data analysis plan for the CATS assessment tool.

Context of Literature Review

Interprofessional education (IPE) is an effective teaching and learning strategy that is used to foster and develop teamwork and communication skills in both academic (Speakman, 2016; World Health Organization, 2010) and professional settings (Weaver et al., 2010; Van Schaik, Plant, Diane, Tsang, & O'Sullivan, 2011), however the impact of high fidelity simulation in interprofessional education to foster teamwork has not been well established. Angelini (2011) believed the current, uniprofessional nature of academic curricula can lead to attitudes of professional hierarchy and a sense of professional competitiveness, which can ultimately disrupt effective teamwork behaviours. While undergraduate health science curriculums have traditionally included a variety of clinical and simulation exercises, many of these curriculums are uniprofessional in their program delivery (Angelini, 2011; Garbee et al., 2012; Garbee et al., 2013). Therefore, it is critical to promote and evaluate new and innovative approaches to interprofessional education, such as the use of high fidelity simulation to effectively foster communication and teamwork skills within the clinical setting.

The World Health Organization (2010) asserts that fostering teamwork and communication skills within the clinical setting is crucial to ensuring positive patient outcomes such as patient safety and quality care. In order to achieve this goal, it is recommended that teaching and learning strategies focus on interprofessional education be initiated at the undergraduate health sciences level and be supported in the clinical practice setting. Interprofessional education initiated at the undergraduate level can lead to a positive interprofessional teamwork environment. A positive teamwork environment and effective communication skills can potentially increase patient safety, decrease clinical mistakes, increase patient satisfaction, decrease nursing turnover, and decrease mortality (Manser, 2009; Sorbero, Farley, Mattke, & Lovejoy, 2008; Weaver et al., 2010; Zangaro & Soeken, 2007). Conversely, ineffective teamwork and poor communication can potentially lead to clinical inefficiencies, an increased waste of clinical supplies, delayed procedures, procedural errors, poorer patient outcomes, and dissatisfaction among team members (Aebersold, Tschannen, & Sculli, 2013; Lingard et al., 2004; Mazzocco et al., 2009).

It has been well established in the literature that IPE can effectively foster and develop self-perceived improvements in teamwork and communication skills in both academic and practice settings, but the impact of HF-IPE on communication and teamwork in undergraduate education has not been well established. High fidelity simulations have been used extensively within uniprofessional undergraduate programs in nursing, medicine, and pharmacy. Research suggests that health sciences students must be exposed to HF-IPE early in their undergraduate curriculums if this is to lead to effective collaborative practices when they enter the workforce (Dillon, Noble, & Kaplan, 2009; Garbee et al., 2012; Garbee et al., 2013; King et al., 2014).

While there is a growing research database surrounding HF-IPE and undergraduate health science students (Dillon et al., 2009; Garbee et al., 2012; Garbee et al., 2013; Jakobsen et al., 2018; Paige et al., 2014; Reese, Jeffries, & Engum, 2010; Smithburger, Kane-Gill, Kloet, Lohr, & Seybert, 2013; Tofil et al., 2014) researchers agree that more research is needed to further understand the role HF-IPE can play in fostering communication and teamwork behaviours among health science students from different professions. This would support the need for valid and reliable instruments that can measure the impact of HF-IPE on communication and teamwork behaviours.

Description of Search Methods

The search of the literature included searching the databases of CINAHL (2006 to January 2018), PubMed (2006 to January 2018) and Google Scholar (2006 to January 2018). Key words and phrases used while searching those databases included: communication and teamwork skills, CATS, teamwork, high fidelity interprofessional education, interprofessional education, simulation, undergraduate students, and data analysis plan. The parameters from 2006 to 2018 were set to reflect current, relevant research, while also incorporating all research that has been published on the CATS assessment tool since it was created in 2007. A lateral search was conducted using the 'similar article' function present on CINAHL, PubMed, and Google Scholar. An additional lateral search was also conducted searching reference lists of collected articles. A final lateral search was completed using Google to collect gray literature focused on these areas of interest. Once a relevant article was identified, the abstract was scanned for key words and phrases. If applicable, the entire article was reviewed and critiqued. A total of 52 resources were reviewed including 25 research studies, eight systematic/literature reviews, eight reports, seven textbook sources, and four grey literature sources. The following is a discussion of the themes arising from the review of the literature related to HF-IPE and the CATS.

Themes Arising from the Literature

Analysis of the literature revealed research to support two general themes related to the benefits of high fidelity simulation in interprofessional education and the impact of high fidelity simulation on communication and teamwork behaviors. Review of the literature on the CATS assessment tool revealed the tool has been used to gather data in a variety of research and clinical settings, with an overall goal of understanding communication and teamwork behaviours. A review of the literature also revealed that the majority of the data analysis plans for the CATS were for analysis of quantitative data.

Benefits of High Fidelity Simulation in Interprofessional Education

High fidelity simulation (HF) is a practice-based teaching and learning strategy that consists of simulating a real clinical environment using advanced human patient simulators to create a high degree of realism, interactivity, and responsiveness. High fidelity simulations provide students with a safe learning environment where they can apply their critical thinking skills to a practical situation (Cook et al., 2011). Interprofessional education is an education-based teaching and learning strategy that consists of students from two or more different professions or areas of academia coming together to form a team, with a purpose of learning from each other, improving future collaborative practices, and improving the care provided to healthcare recipients (Newton et al., 2015; World Health Organization, 2010). HF-IPE consists of a combination of high fidelity simulation and interprofessional education where teams of two or more students come together to learn using a human patient simulator to create a realistic, and interactive learning environment.

HF-IPE can provide an immersive, hands-on experience in a non-threatening learning environment (Jeffries, Swoboda, & Akintade, 2016). Benefits from participating in HF-IPE include: increasing knowledge, improving patient outcomes, increasing skill competency, and increasing appropriate clinical behaviours (Cook et al., 2011). HF-IPE is often the preferred environment for high-stakes medical training as they provide a safe space where teamwork skills and task-orientated skills can be performed (Hunt, Fiedor-Hamilton, & Eppich, 2008; Scheckel, 2016). It is important to note that while research indicates that HF-IPE is an appropriate teaching and learning strategy for health science curriculums, there are few studies that measure impact of HF-IPE on team behaviour, and it is difficult to infer whether those changes would transfer to real-life clinical settings.

Participating in IPE can help to break down real or perceived barriers among healthcare team members, improve cohesiveness among team members, and can be instrumental in building mutual respect among team members (Jeffries, Swoboda, & Akintade, 2016). Scherer, Myers, O'Connor, and Haskins (2013) determined that simulation-based IPE was more beneficial to knowledge acquisition, preparedness for collaboration, professional identity, and understanding roles and responsibilities, when compared to uniprofessional simulation-based education. These researchers along with Aliner et al. (2014) agree that immersive experiences like HF-IPE, could be used to bridge the gap between traditional uniprofessional education curriculum and the interprofessional collaborative practices that are needed in real-life clinical settings.

This research shows that HF-IPE can have a positive impact on undergraduate students' understanding of the complexity surrounding communication and teamwork behaviours. It is an appropriate teaching and learning strategy to improve cohesiveness among team members, increase knowledge, improve patient outcomes, increase skill competency, and provide a safe space where communication and teamwork skills can be fostered.

Impact of Simulation on Communication and Teamwork Behaviours

The key to the success of HF-IPE is to create engaging experiences that accurately reflect a real life clinical situation. Evidence suggests that HF-IPE with post licensure health care professionals can have a positive impact on their perceived communication and teamwork skills (King et al., 2014), but it is not clear whether this same impact is seen in undergraduate students. Reese et al. (2010) investigated self-perception of role in nursing and medical students during an HF-IPE experience and determined that students perceived participation in the HF-IPE as benefiting the development of their team collaboration skills. Other research studies reported positive improvements in the student's confidence and perception of communication skills (Jakobsen et al., 2018; Paige et al., 2014). However, neither of these studies measured the impact of participation on both communication and teamwork behaviours.

Other studies report similar findings of the positive impact of HF-IPE on collaboration, communication and teamwork. Dillon et al. (2009) also measured students' perception of the impact of HF-IPE after they participated in an interprofessional mock-code simulation for nursing and medical students. That study assessed perceptions of collaboration from a quantitative and qualitative perspective. Dillon et al. noted that after participation in the HF-IPE, both nursing and medical students reported the experience was beneficial and that HF-IPE should be a part of future education curriculums for nursing and medicine. Stewart, Kennedy, and Cuene-Grandidier (2010) and Tofil et al. (2014) also reported on the positive benefits of HF-IPE to enhance professional identity and role awareness within an interprofessional situation.

While the majority of research collected for this literature review focused on self-reporting as a form of data collection, Paige et al. (2014) completed a HF-IPE study with health science students that included observed behaviour scores. Those observed behaviour scores were completed using a data collection tool that was specifically created for operating-room teamwork assessments, as this was the environment where the HF-IPE was designed to take place. Paige et al. determined that HF-IPE led to statistically significant gains ($p < 0.001$) in all subscales of their observed team behaviour assessment tool. Paige et al. asserted that HF-IPE can be a feasible and effective method of education delivery that can have an immediate impact on participants' perceived and observed teamwork behaviours, however, it is difficult to infer whether those changes would transfer to real-life clinical settings.

Murdoch, Bottorff, and McCullough (2014) performed a systematic literature review that focused on best practices surrounding simulation within IPE as it relates to students within nursing programs. Murdoch et al. believed that a wide variety of simulation techniques – including high fidelity patient simulation – offered benefits to nursing students surrounding IPE practices. However, Murdoch et al. also believed that future research is needed to develop valid and reliable evaluation tools to measure the success of IPE within academic settings. While it is postulated that HF-IPE targeting undergraduate nursing, medicine, and pharmacy students could be beneficial for future real-life scenarios, no research has been completed to determine if these experiences will ultimately lead to increased skills when entering the workforce.

A review of this literature revealed that the primary method of data collection when focusing on the benefits of HF-IPE among undergraduate health sciences programs is self-efficacy and self-perceptions via self-reporting. There is limited evidence focusing on other data collection methods such as measuring the impact of HF-IPE on communication and teamwork behaviours through observation of behaviours in a simulated setting. Despite the positive findings related to the previous research, there are limitations within the research on HF-IPE that must be acknowledged.

Limitations of High Fidelity Simulation Research

When looking at the previous research focused on HF-IPE, the limitations must be noted and taken into consideration when evaluating the strengths of the reported findings. Convenience samples (Dillon et al., 2009; King et al., 2014) and small sample sizes (Dillon et al., 2009; Reese et al., 2010; King et al., 2014; Paige et al., 2014) were limitations within the previous literature that could interfere with the generalizability of the findings. It should also be noted that some studies (Tofil et al., 2014) used non-validated assessment tools to collect their data, thus the generalizations of their results might also be limited.

It was also evident from the research gathered for this literature review that most researchers tailored HF-IPE simulations to fit the specific learning needs of their participants. Only Paige et al. (2014) asserted that their simulation was a standardized scenario. Each research team appeared to develop their own scenario and their expected outcomes. This could be considered a major limitation due to the fact that research-specific tailored scenarios may be difficult to compare across research studies. While the basic concepts of HF-IPE are the same within the research gathered for this literature review, the intricate differences in scenarios could contribute to confounding variables that may impact the generalizability of the results from each study. A standardized simulation scenario used in future research surrounding HF-IPE could be beneficial, as it could provide a consistent data collection environment that would be easier to compare across different research studies.

The lack of variety surrounding different interdisciplinary teams may also be considered a limitation due to its poor reflection of real-life scenarios. Within the research gathered for this literature review, the IPE team members were primarily nursing and medical students (Dillon et al., 2009; Jakobsen et al., 2018; Paige et al., 2014; Reese et al., 2010; Stewart et al., 2010; Tofil et al., 2014). In real-life, interprofessional interactions would not be limited to only nursing and medical professionals, but also include a variety of other health disciplines. King et al. (2014) believed HF-IPE should not be limited to nursing and medical students, but should also include other health sciences students, such as respiratory therapy and nursing aides. A team consisting of only nursing and medical students may differ greatly from a team that is comprised of students from a variety of disciplines such as nursing, medicine, respiratory therapy, physical therapy, and pharmacy. While having interprofessional teams consisting of the same student populations may increase the ability to compare results across different research studies, it limits the generalizability of results to real-life situations. More research is needed focusing on interprofessional teams of health science students from a variety of health disciplines to garner a greater understanding regarding how HF-IPE affects communication and teamwork skills for a variety of health disciplines.

Self-reporting as a primary method of data collection (Dillon et al., 2009; King et al., 2014; Jakobsen et al., 2018; Reese et al., 2010; Tofil et al., 2014) could also be considered a limitation of the previous research due to the fact that overestimation or underestimation of abilities can occur (Havyer et al. 2016; Paige et al., 2014). While self-reporting does provide insightful information surrounding how participants felt regarding their HF-IPE experiences, it does not provide any concrete evidence surrounding their knowledge acquisition, skills, or teamwork behaviours. Objective measurement tools that analyze data collected on observed teamwork behaviours - such as the CATS assessment tool - could help document more precise result surrounding the impact of HF-IPE on communication and teamwork. While limitations are present within the previous research designs, researchers have also outlined barriers to implementing HF-IPE.

Barriers to Implementing HF-IPE

HF-IPE can be a feasible and effective method of education delivery that can have an immediate impact on participants' perceived and observed teamwork behaviours, however, there are few studies that measure the impact on behaviour and it is difficult to infer whether those changes would transfer to real-life clinical settings. Van Schaik et al. (2011) believed that while HF-IPE can be beneficial for participants, limitations and barriers exist surrounding the implementation of these programs including: difficulty coordinating the participant's schedules, high cost for set up and maintenance of the human patient simulators, and difficulty in recreating real-life work environments. Van Schaik et al. made reference to these limitations as they related to HF-IPE and working professionals, but these limitations are also present when planning for HF-IPE in undergraduate health science curricula (Lapkin, Levett-Jones, & Gilligan, 2012).

Newton et al. (2015) believed IPE in academic settings is often limited by a lack of flexibility in undergraduate curricula, limited shared free time across various academic disciplines, resource constraints, space constraints, economic constraints, and a lack of faculty development regarding IPE. Conversely, others believed HF-IPE can be feasible given a large enough target population and the proper teaching environment (Jakobsen et al., 2018; Paige et al., 2014). A large-scale cost benefit analysis surrounding HF-IPE within health sciences academic programs could provide valuable information regarding whether or not HF-IPE is a cost effective endeavour within undergraduate programs.

The Communication and Teamwork Skills Assessment Tool

Frankel et al. (2007) created the CATS as an instrument to measure communication and teamwork skills of healthcare professionals in real world and simulated settings. The CATS assessment tool was designed to reach a broad range of healthcare professionals, and focused on directly observing behaviour while quantitatively gathering data. Frankel et al. wanted to develop an assessment tool that focused on quantitative - how often - and qualitative - how well - particular skills were performed, while also having an opportunity to provide feedback to a team as a whole. It is important to note that feedback given at the end of a CATS assessment focuses only on team communication behaviours. The tool does not collect data in such a way that allows the feedback to pinpoint specific examples or specific team member's behaviours.

The CATS assessment tool is based upon crisis resource management behaviour-based markers, which have been used in other non-medical professions, such as aviation (Frankel et al., 2007). The CATS assessment tool investigates four domains of team behaviour: situational awareness, coordination, communication, and cooperation. Within these domains, there are 21 behaviour markers that are assessed by a trained observer. However, three of these behaviour markers are only observed and scored if the scenario also involves a crisis situation.

The scores under each behaviour marker are weighted, depending on the quality of the behaviour observed. A trained observer will place a tick in the appropriate box given a specific behaviour. A behaviour viewed as "*good*" scores 1 point, a behaviour viewed as "*variable in quality*" scores 0.5 points, and an "*expected but not observed*" behaviour scores 0 points. A weighted score out of 100 is then calculated for each individual behaviour marker, each domain, and as an overall score. Teams are scored on the basis of how often an event occurs and the quality of their communication and teamwork behaviours (Seelandt et al., 2014). The CATS assessment tool provides a unique data analysis opportunity because researchers can focus on the overall score, the score within a specific domain, or the score of a specific behaviour marker or a group of behaviour markers.

Frankel et al. (2007) believed that the CATS assessment tool is appropriate to assess communication and teamwork skills in a variety of healthcare settings. The CATS assessment tool provides a unique opportunity for a trained observer to evaluate team behaviour without focusing on individual behaviour or performance, as behaviours are catalogued and analyzed from an overall-team perspective, not individual behaviour monitoring. From a research perspective, it appears that the CATS assessment tool can be used to gather data in a variety of research settings, with the overall goal focusing on understanding the communication and teamwork skills of a given team.

Since Frankel et al. (2007) there has been some evidence to suggest that the CATS is a valid and reliable assessment tool to measure communication and teamwork behaviours in professional practice environments (Hughes et al., 2014; Joshi, Hernandez, Martinez, AbdelFattah, & Gardner, 2017; Passauer-Baierl, Baschnegger, Bruns, & Weigl, 2014). There has also been limited research using the CATS assessment tool to assess communication and teamwork behaviours among teams of interprofessional undergraduate health science students (Garbee et al., 2012; Garbee et al., 2013; Smithburger et al., 2013). This would suggest a need for more research in this area.

Hughes et al. (2014) noted that the CATS assessment tool is an important resource when evaluating and designing an interprofessional education program that focuses on teamwork skills among working professionals. Hughes et al. used the CATS assessment tool to pinpoint specific aspects of teamwork skills, which were addressed via an education program. After participation in the program the teams showed statistically significant improvements (p = <0.05) in their CATS assessment scores post-education delivery (Hughes et al., 2014). Joshi et al. (2017) took a different perspective on teamwork research by using the CATS assessment tool to investigate whether stable or dynamic team structures have an impact on teamwork communication skills. Joshi et al. determined that both dynamic and stable teams can experience positive benefits from taking part in repeated exposure to simulated scenarios. Passauer-Baierl et al. (2014) used the CATS to assess interprofessional teamwork skills within an operating room. It is clear that the CATS assessment tool has been used to assess communication and teamwork skills from a variety of perspectives in different professional environments. However, not all studies focused on HF-IPE scenarios (Hughes et al., 2014; Passauer-Baierl et al., 2014). Furthermore, these studies made no inferences regarding HF-IPE experiences during health sciences undergraduate education and whether or not it could lead to positive benefits when students enter the workforce.

The CATS assessment tool has been used to assess interdisciplinary teams comprised of various undergraduate health science students. Smithburger et al. (2013) used the CATS assessment tool to assess the communication and teamwork skills of teams comprised of pharmacy, medicine, nursing, social work, and physician assistant students. Smithburger et al.

argued that HF-IPE sessions allowed for a statistically significant improvement ($p < 0.01$) in communication teamwork scores when assessed using the CATS assessment tool. Furthermore, Smithburger et al. also determined the inter-rater reliability of the CATS assessment scores were high among different evaluators, which is congruent with previous research findings (Garbee et al., 2013).

Other research has focused on quasi-experimental, pre-test/post-test research designs that investigate how student teams develop and retain communication and teamwork skills over time (Garbee et al., 2012; Garbee et al., 2013). Teams were comprised of nursing, nurse anesthesia, medicine, and respiratory therapy students (Garbee et al., 2013) or medicine, nurse anesthesia, nursing, and physical therapy students (Garbee et al., 2012). Garbee et al. (2012) used the CATS assessment tool to show that participation in HF-IPE had a positive impact on participants' communication and teamwork skills and this improvement was retained after six months. Conversely, Garbee et al. (2013) noted the retention of these skills was not evident when re-evaluated after a five-month hiatus. Despite these conflicting results, researchers believed HF-IPE is an effective teaching and learning strategy for undergraduate education and it has been shown to enhance the development of effective communication and teamwork skills that they can use when students enter the workforce (Garbee et al., 2012; Garbee et al., 2013). It is clear from this literature review that the CATS assessment tool is an appropriate instrument to measure the impact of HF-IPE on undergraduate nursing, medicine, and pharmacy students' teamwork and communication behaviours.

Strengths and Limitations of the CATS

When looking at the previous research that used the CATS assessment tool, the limitations must be noted and taken into consideration when evaluating the strengths of the reported findings. Small sample sizes, attrition of participants, and convenience samples are all factors that could be considered limitations within the previous research and could interfere with the generalizability of the results (Garbee et al., 2012; Garbee et al., 2013; Smithburger et al., 2013). Garbee et al. (2013) believed that scheduling conflicts among different academic programs was one of the greatest contributors to small samples sizes and attrition in their research study. Furthermore, the variation in interprofessional team members may also limit the ability to compare findings between studies. The different academic backgrounds, program expectations, and previous knowledge of students from different disciplines, may all be contributing factors to confounding variables that could negatively impact the validity of the results. While using a variety of team members may increase the generalizability of results to real-life scenarios, more research is necessary to strengthen the claims of the previous research findings. Despite the lack of research and potential limitations within the research gathered using the CATS assessment tool, it does appear evident that the CATS assessment tool can be used in a variety of professional and academic settings while focusing on different aspects of communication and teamwork skills. However, the strengths and limitations of the assessment tool itself must be taken into account.

Many literature reviews and systematic reviews have focused on communication assessment tools and have analyzed the benefits and limitations of the CATS assessment tool. Rosen et al. (2010) believed the CATS assessment tool allowed for a thorough assessment of teamwork, by assessing the quantity of behaviours, the quality of behaviours, and assessing behaviours from a whole-team perspective. Using trained observers to directly observe behaviour is considered one of the strengths of the CATS, as self-assessment can often lead to an overestimation or underestimation of skills and abilities (Havyer et al. 2016; Paige et al., 2014).

Havyer et al. completed a systematic review of communication assessment tools and determined that the CATS assessment tool does appear to have content validity, response process and internal structure validity, while also having a high degree of inter-rater reliability. This high-degree of inter-rater reliability is also congruent with previous research findings (Garbee et al., 2013; Smithburger et al., 2013). Havyer et al. recommended that the CATS assessment tool should be used when assessing interprofessional collaboration within undergraduate medical education because it aligns with the interprofessional collaboration competencies that are set forth by the Association of American Medical Colleges. It is important to note that while there are strengths surrounding the CATS assessment tool and the previous research findings do seem positive regarding the CATS assessment tool's ability to evaluate communication and teamwork skills, there are differences of opinions regarding the validity, reliability, and limitations of the CATS.

Some researchers believe that the CATS assessment tool has not undergone enough rigorous statistical analysis to determine its validity or reliability regarding measuring communication and teamwork skills (Rehim, DeMoor, Olmsted, Dent, & Parker-Raley, 2017; Sanfey, McDowell, Meier, & Dunnington, 2011; Seelandt et al., 2014; Van Schaik et al., 2011). Furthermore, Havyer et al. (2016) based their arguments surrounding validity and inter-rater reliability only on two studies (Frankel et al., 2007; Garbee et al., 2013), which could be considered not sufficient evidence to make such determinations. Feasibility of using the CATS assessment tool is also a concern due to the financial requirements necessary when training observers to collect data (Havyer et al, 2016). Rosen et al. (2010) believed that since the CATS assessment tool only collects data using a tick-sheet format, it might be difficult to debrief participants and discuss specific situations that may have happened during a scenario.

Furthermore, some researchers argued that the CATS assessment tool does not effectively determine if behaviours are being performed correctly, appropriately, or effectively, it merely focuses on the frequency in which behaviours are being performed or not performed (Flowerdew, Brown, Vincent, & Woloshynowych, 2012; Hughes et al., 2014; Rosen et al., 2010). It is clear that there is a difference of opinion within the academic world surrounding the strengths and limitations of the CATS assessment tool.

It is also important to note that some researchers have modified the CATS assessment tool to create a new assessment tool that meets their specific research needs (Weaver et al., 2010). Caution has to be noted in this case due in part to the fact that the CATS assessment tool is not widely considered a robust and validated assessment tool. Creating different tools based on non-validated tools can put the validity of research findings into question. More research is needed using the CATS assessment tool to determine its validity and reliability before other tools can be created using the CATS assessment tool as a guideline.

Weller et al. (2011) believed there is a lack of robust assessment tools that focus on teamwork skills within a multidisciplinary setting. It is evident from this literature review there is no consensus regarding the strengths and the limitations of the CATS assessment tool. This conflict in information only reinforces the assertion that more research is necessary to further understand the validity and reliability of the CATS assessment tool. It does appear evident that the CATS assessment tool can be used in a variety of IPE environments, including high fidelity simulations targeting undergraduate students. Gaps in the literature provide unique opportunities for future research studies to garner a greater understanding surrounding communication and teamwork skills during HF-IPE among undergraduate students.

Gaps in the Literature

This literature review has pinpointed gaps in the literature surrounding the CATS assessment tool and HF-IPE within academic settings. Lapkin et al. (2012) believed there is no evidence to suggest at what point in a student's undergraduate academic program they should start IPE. Research collected for this literature review focused on students in the latter parts of their academic programs. The rationalization for only including students who are nearly finished their academic programs is because senior students have the existing knowledge and confidence to take part in IPE scenarios (Stewart et al., 2010). It is the assumption that novice students would not have the skills or knowledge required to participate in HF-IPE, but no research has been completed investigating such assumptions. Furthermore, no longitudinal studies have been completed looking at HF-IPE throughout a student's undergraduate program and beyond into their professional practice. Stewart et al. asserted that long term follow up studies are necessary to determine the lasting effects of HF-IPE within education programs.

No research could be found that focused on HF-IPE that consists of nursing, medicine, and pharmacy students. Furthermore, no research could be found that investigated the differences between communication and teamwork skills used and acquired during a HF-IPE experience when compared to other educational experiences. Only one study was retrieved that focused on high fidelity versus low fidelity (Cheng et al., 2015) but it did not focus on IPE. Furthermore, Cheng et al.'s meta-analysis included articles that focused on both undergraduate students and working professionals. Masiello (2012) asserted that simulation approaches to team learning have not been used effectively, but this assertion did not solely focus on simulations at an academic level. It is evident that more research is needed in these areas of interest.

These various gaps in the literature provide an opportunity for this practicum project to collect valuable information surrounding the relationship between HF-IPE and communication and teamwork skills among nursing, medicine, and pharmacy students. Even though some researchers have concerns regarding the validity of the CATS assessment tool, further research using this tool is necessary to provide greater insight into the assessment tool's validity. The lack of research surrounding the CATS assessment tool should not intimidate future researchers from using this assessment tool. Researchers should continue to use the CATS assessment tool to increase the breadth and depth of academic knowledge surrounding the CATS assessment tool. In order to achieve success when using a relatively new research tool, a data analysis plan is necessary to ensure proper steps are taken during the research process.

A Data Analysis Plan for the Communication and Teamwork Assessment Tool

Within any research study, a data analysis plan is integral to the research process as it is a way to convince the intended audience that a comprehensive plan is in place to analyze the data once collected. The data analysis plan acts as a road map to guide the research study from planning, to implementation, to evaluation of the data, and interpretation of the results (Planter, 2011; Simpson, 2015). A data analysis plan will also outline your plan to answer your research questions in a clear and concise manner (Chasan-Taber, 2014; Planter, 2011). Without a properly detailed data analysis plan, it would be difficult to determine if specific research findings have any validity or importance to the research question. Furthermore, a good data analysis plan can allow a researcher to transform quantitative data into a descriptive explanation, discussing the meaning of the information and why this information is important.

One of the first steps in any data analysis plan is to properly outline the research question, the proposed hypothesis or hypotheses of the research study, and the specific aims of the research study (Chasan-Taber, 2014; Feldman, 2014). Understanding the distinction between such basic research concepts as variable, value, independent variable, and dependent variable, are also important starting points to any data analysis plan (Simpson, 2015). When looking at the CATS assessment tool, each of the 21 observable behaviour markers are the variables, whereas the values would be the number of ticks within each respective three-point check system. The options within the CATS assessment tool provides flexibility to the researcher, which allows them to focus their data analysis plan on either the overall score, a specific domain score, or each of the 21 behaviour marker scores. This flexibility allows the researcher to modify their data analysis plan to meet their specific research question needs.

Within any data analysis plan, the researcher must also be cognizant of what is considered a dependent and independent variable. The dependent variable is considered the variable of interest, as its results are directly influenced by the manipulated variable, which is also called the independent variable (Polit & Beck, 2017). When looking at the CATS assessment tool and its role in the proposed practicum project, the dependent variable would be the CATS assessment scores for each team and the independent variable would be the HF-IPE and LF-IPE simulations. Specifically, the project would focus on whether or not a team's communication and teamwork skills – as measured by the CATS assessment tool – is dependent on taking part in HF-IPE or LF-IPE.

It is also important to determine if a data analysis plan should focus on descriptive statistics, inferential statistics, or both (Simpson, 2015). Descriptive statistics focus on merely describing and summarizing data sets, whereas inferential statistics focus on examining the relationship among variables and making inferences based on these relationships (Kellar & Kelvin, 2013; Polit & Beck, 2017). Descriptive statistics traditionally are analyzed using univariate analyses, whereas inferential statistics may incorporate univariate or multivariate analyses (Chasan-Taber, 2014). Simpson (2015) believed that when choosing the appropriate statistical analysis for a data analysis plan, certain questions need to be answered. A researcher must understand what they are trying to determine from their research, what the design of their study is, and what level of measurement they are using to collect data. Consultations are another important component of the data analysis plan. Statisticians, faculty members, or research coordinators are all individuals that can be consulted during the creation of a data analysis plan to ensure the planned statistical analysis is congruent with the research question being asked (Planter, 2011; Simpson, 2015).

When using inferential statistics, the primary objective is to determine the *p* value, which looks at the probability that the observed results are due to chance (Polit & Beck, 2017). Traditionally, a *p* value of less than 0.05 is considered statistically significant (Polit & Beck, 2017). In simple terms, if a *p* value is less than 0.05 it means that the findings of the study would be the result of random chance less than 5 out of 100 times. The lower the *p* value, the lower the likelihood the findings are a direct result of chance (Polit & Beck, 2017).

Chasan-Taber (2014) asserted that a data analysis plan should also take into account confounding variables, and outline how to control these variables. Confounding variables are variables outside of the parameters of the research design that may inadvertently affect the results of the proposed research (Polit & Beck, 2017). Consultation with a statistician should be utilized as a way to determine a thorough statistical analysis plan that can help control for confounding variables (Chasan-Taber, 2014; Polit & Beck, 2017). It is important to note that sample size can also play a vital role in controlling for confounding variables. Small sample sizes may prohibit a researcher from performing a multivariate regression model to control for confounding variables, due to the limited amount of data collected (Chasan-Taber, 2014). An adequate sample size is needed to control for these confounding variables.

A power analysis can be completed as part of a data analysis plan to determine the sample size and the number of participants needed given the context of the research question and the statistical analysis requirements of the proposed research design (Planter, 2011; Polit & Beck, 2017). A proper power analysis will also limit the probability of committing a type II error, which happens when researchers assert no relationship exists among variables, when in fact a relationship does exist (Chasan-Taber, 2014; Polit & Beck, 2017) Statistical software such as G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) can be utilized to conduct a power analysis to help determine an adequate sample size for the study.

A data analysis plan should also outline what statistical software will be used to complete the statistical analysis, to ensure the software can successfully complete the statistical analysis required (Centers for Disease Control and Prevention, 2013; Planter, 2011). Each statistical software program that is included in the data analysis plan should also include the version of the software. Planter (2011) believed this information was integral for the reader to determine if the statistical software chosen is appropriate given the context of the research question and the intended data analysis plan that has been outlined.

A data analysis plan not only outlines specifically what type of data is being collected and how the data will be analyzed, the plan should also outline how to visually represent the results of the data. Simpson (2015) believed that how the data is visually displayed is directly related to what type of data is collected. Pie graphs, bar graphs, histograms, box plots, scatter plots, and tables are just a few examples of how to display data once it has been collected and analyzed (Simpson, 2015). Ultimately, if the analysis of the data cannot be understood due to poor visual representation, then the information - no matter how important - will never be successfully disseminated and absorbed by the intended audience.

Planter (2011) believed with any data analysis plan, there must be flexibility to modify the plan to ensure it meets all the requirements of the research study, in case those requirements change throughout the research process. A good data analysis plan contributes to the reliability and validity of any research study and acts as a blue print to follow throughout the research process. While this literature review produced only one peer-reviewed article (Simpson, 2015) that focused on developing data analysis plans, various textbook material and gray literature was retrieved that provided valuable information on the topic (Chasan-Taber, 2014; Centers for Disease Control and Prevention, 2013; Planter, 2011; Polit & Beck, 2017). More peer-reviewed research focusing on data analysis plans for health related instruments may provide greater insight and guidance for future researchers when preparing their own research proposals and formulating their own data analysis plans.

Conclusion

In summary, this paper has provided a robust and comprehensive literature review to answer the following questions: (1) Has the CATS been used to measure communication and teamwork behaviours among undergraduate nursing, medicine and pharmacy students? (2) Has the CATS been used to measure teamwork behaviours in HF-IPE? and (3) Is there a relationship between HF-IPE and communication and teamwork behaviours in undergraduate nursing, medicine, and pharmacy students? Research has shown that the CATS assessment tool has been used extensively to measure communication and teamwork behaviours in post-licensure healthcare professionals. However, there are few studies that use the CATS assessment tool to measure HF-IPE and communication and teamwork behaviours in nursing, medicine, and pharmacy undergraduate students. More research is needed in this area.

One of the themes that emerged from the literature review related to the benefits of HF-IPE in undergraduate health science education programs, however it is difficult to generalize those findings because of the intricate differences between the experiences that could contribute to confounding variables that may impact the generalizability of the results. Although current research indicates there are many benefits to participating in HF-IPE including increased knowledge, improved patient outcomes, increased skill competency, and increased appropriate clinical behaviours, there is no research on the impact of HF-IPE on communication and teamwork behaviours among teams of interprofessional undergraduate nursing, medicine, and pharmacy students. Although this literature review helped to answer the second question, it could not answer the questions surrounding the CATS assessment tool and whether or not there is a relationship between HF-IPE and communication and teamwork behaviours with nursing, medicine, and pharmacy undergraduate students. The majority of the research studies focused on post-licensure health care professionals. This would indicate there is a need for further research in the area of undergraduate HF-IPE.

This literature review helped to inform the development of a comprehensive data analysis plan for the CATS assessment tool. Although there were gaps in the literature related to the CATS assessment tool in undergraduate education, there were also opportunities identified for future research. In relation to the development of a data analysis plan for the CATS, previous research studies show that the flexibility of the CATS allows the researcher to modify the instrument to meet their specific research question needs. Therefore, the data analysis plan for the CATS should include correlation of the items in the CATS to the research question and modifications as needed. A power analysis could be completed as part of the data analysis plan to determine the number of subjects needed given the context of the research question and the statistical analysis requirements of the proposed research design. It is recommended that the SPSS data analysis package be used to analyze the inferential data from the CATS. The data analysis plan for the CATS should also include how the data will be displayed e.g. pie graphs, bar graphs, histograms, and / or tables. It is also recommended that consultations with a statistician occur when developing the data analysis plan for the CATS.

This literature review has determined that the CATS assessment tool is an appropriate tool to analyze the impact of HF-IPE on communication and teamwork behaviours among undergraduate nursing, medicine, and pharmacy students. The majority of authors agreed however, that more research is needed on the CATS assessment tool to determine whether or not it accurately measures the impact of HF-IPE on communication and teamwork behaviours in undergraduate education. This literature review formed the basis for the evidence used to develop a data analysis plan for the CATS as a research instrument to measure the impact of HF-IPE on communication and teamwork behaviours in undergraduate nursing, medicine, and pharmacy students.

References

Aebersold, M., Tschannen, D., & Sculli, G. (2013). Improving nursing students' communication skills using crew resource management strategies. *Journal of Nursing Education*, *52*(3), 125-130. doi:10.3928/01484834-20130205-01

Aliner, G., Harwood, C., Harwood, P., Montague, S., Huish, E., Ruparelia, K., . . . Antuofermo, M. (2014). Immersive clinical simulation in undergraduate health care interprofessional education: Knowledge and perceptions. *Clinical Simulation in Nursing*, *10*(4), E205-E216. doi:10.1016/j.ecns.2013.12.006

Angelini, D. J. (2011). Interdisciplinary and interprofessional education: What are the key issues and considerations for the future? *The Journal of Perinatal and Neonatal Nursing*, *25*(2), 175-179. doi:10.1097/JPN.0b013e318212ee7a

Centers for Disease Control and Prevention. (2013). *Creating an analysis plan.* Retrieved from https://www.cdc.gov/globalhealth/healthprotection/fetp/training_modules/9/creating-analysis-plan_pw_final_09242013.pdf

Chasan-Taber, L. (2014). Data analysis plan. *Writing dissertation and grant proposals: epidemiology, preventive medicine and biostatistics* (pp.179-195). New York: NY. CRC Press.

Cheng, A., Lockey, A., Bhanji, F., Lin, Y., Hunt, E. A., & Lang, E. (2015). The use of high-fidelity manikins for advanced life support training - A systematic review and meta-analysis. *Resuscitation*, *93,* 142-149. doi:10.1016/j.resuscitation.2015.04.004

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., . . . Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. *JAMA, 306*(9), 978-988. doi:10.1001/jama.2011.1234

Dillon, P. M., Noble, K. A., & Kaplan, L. (2009). Simulation as a means to foster collaborative interdisciplinary education. *Nursing Education Perspectives, 30*(2), 87-90. Retrieved from https://journals.lww.com/neponline/

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/bf03193146

Feldman, H. (2014). Data analysis plans [PowerPoint slides]. Retrieved from https://catalyst.harvard.edu/pdf/biostats/HSPHAnalysisPlanBW2014.pdf

Flowerdew, L., Brown, R., Vincent, C., & Woloshynowych, M. (2012). Development and validation of a tool to assess emergency physicians' nontechnical skills. *Annals of Emergency Medicine*, *59*(5), 376-385. doi:10.1016/j.annemergmed.2011.11.022

Frankel, A., Gardner, R., Maynard, L., & Kelly, A. (2007). Using the Communication and Teamwork Skills (CATS) assessment to measure health care team performance. *Joint Commission Journal on Quality and Patient Safety, 33*(9), 549-558. doi:10.1016/S1553-7250(07)33059-6

Garbee, D., Paige, J., Bonanno, L., Rusnak, V., Barrier, K., Kozmenko, L., . . . Nelson, T. (2012). Effectiveness of teamwork and communication education using an interprofessional high-fidelity human patient simulation critical care code. *Journal of Nursing Education and Practice, 3*(3), 1-12. doi:10.5430/jnep.v3n3p1

Garbee, D., Paige, J., Barrier, K, Kozmenko, V., Kozmenko, L., Zamjahn, J., . . . Cefalu, J. (2013). Interprofessional teamwork among students in simulated codes: A quasi-experimental study. *Nursing Education Perspective*, *34*(5), 339-344. doi:10.5480/1536-5026-34.5.339

Havyer, R. D., Nelson, D. R., Wingo, M. T., Comfere, N. I., Halvorsen, A. J., McDonald, F. S., & Reed, D. A. (2016). Addressing the interprofessional collaboration competencies of the Association of American Medical Colleges: A systematic review of assessment instruments in undergraduate medical education. *Academic Medicine*, *91*(6), 865-888. doi:10.1097/ACM.0000000000001053

Hughes, K. M., Benenson, R. S., Krichten, A. E., Clancy, K. D., Ryan, J. P., & Hammond, C. (2014). A crew resource management program tailored to trauma resuscitation improves team behavior and communication. *Journal of the American College of Surgeons*, *219*(3), 545-551. doi:10.1016/j.jamcollsurg.2014.03.049

Hunt, E. A., Fiedor-Hamilton, M., & Eppich, W. J. (2008). Resuscitation education: Narrowing the gap between evidence-based resuscitation guidelines and performance using best educational practices. *Pediatric Clinics of North America*, *55*(4), 1025-1050. doi:10.1016/j.pcl.2008.04.007

Jakobsen, R. B., Gran, S. F., Grimsmo, B., Arntzen, K., Fosse, E., Frich, J. C., & Hjortdahl, P. (2018). Examining participant perceptions of an interprofessional simulation-based trauma team training for medical and nursing students. *Journal of interprofessional care*, *32*(1), 80-88. doi:10.1080/13561820.2017.1376625

Jeffries, P. R., Swoboda, S. M., & Akintade, B. (2016). Teaching and learning using simulations. In D. M. Billings & J. A. Halstead (Eds.). *Teaching in nursing: A guide for faculty* (pp.304-323). St. Louis, MI: Elsevier.

Joshi, K., Hernandez, J., Martinez, J., AbdelFattah, K., & Gardner, A. K. (2017). Should they stay or should they go now? Exploring the impact of team familiarity on interprofessional team training outcomes. *The American Journal of Surgery*. 1-7. doi:10.1016/j.amjsurg.2017.08.048

Kellar, S. P., & Kelvin, E. A. (2013). *Munro's statistical methods for health care research* (6th ed.). New York, NY: Wothers Kluwer Health | Lippincott Williams & Wilkins.

King, S., Carbonaro, M., Greidanus, E., Ansell, D., Foisy-Doll, C., & Magus, S. (2014). Dynamic and routine interprofessional simulations: expanding the use of simulation to enhance interprofessional competencies. *Journal of Allied Health*, *43*(3), 169-175. Retrieved from http://www.asahp.org/journal-of-allied-health/

Lapkin, S., Levett-Jones, T., & Gilligan, C. (2012). A cross-sectional survey examining the extent to which interprofessional education is used to teach nursing, pharmacy and medical students in Australian and New Zealand universities. *Journal of Interprofessional Care*, *26*(5), 390-396. doi:10.3109/13561820.2012.690009

Lingard, L., Espin, S., Whyte, S., Regehr, G., Baker, G. R., Reznick, R., . . . Grober, E. (2004). Communication failures in the operating room: An observational classification of recurrent types and effects. *Quality and Safety in Health Care*, *13*(5), 330-334. doi:10.1136/qhc.13.5.330

Manser, T. (2009). Teamwork and patient safety in dynamic domains of healthcare: A review of the literature. *Acta Anaesthesiologica Scandinavica*, *53*(2), 143-151. doi:10.1111/j.1399-6576.2008.01717.x

Masiello, I. (2012). Why simulation-based team training has not been used effectively and what can be done about it. *Advances in Health Sciences Education, 17*(2), 279-288. doi:10.1007/s10459-011-9281-8

Mazzocco, K., Petitti, D. B., Fong, K. T., Bonacum, D., Brookey, J., Graham, S., . . . Thomas, E. J. (2009). Surgical team behaviors and patient outcomes. *The American Journal of Surgery*, *197*(5), 678-685. doi:10.1016/j.amjsurg.2008.03.002

Murdoch, N. L., Bottorff, J. L., & McCullough, D. (2014). Simulation education approaches to enhance collaborative healthcare: A best practices review. *International Journal of Nursing Education Scholarship, 10*(1), 307-321. doi: 10.1515/ijnes-2013-0027

Newton, C., Bainbridge, L., Ball, V., Baum, K. D., Bontje, P., Boyce, R. A., . . . Wood, V. (2015). The health care team challenge™: Developing an international interprofessional education research collaboration. *Nursing Education Today, 35*(1), 4-8. doi:10.1016/j.nedt.2014.07.010

Paige, J. T., Garbee, D. D., Kozmenko, V., Yu, Q., Kozmenko, L., Yang, T., . . . Swartz, W. (2014). Getting a head start: High-fidelity, simulation-based operating room team training of interprofessional students. *Journal of the American College of Surgeons*, *218*(1), 140-149. doi:10.1016/j.jamcollsurg.2013.09.006

Passauer-Baierl, S., Baschnegger, H., Bruns, C., & Weigl, M. (2014). Interdisciplinary teamwork in the OR: Identification and measurement of teamwork in the operating room. *Zeitschrift* *für Evidenz, Fortbildung und Qualität im Gesundheitswesen, 108*(5), 293-298. doi:10.1016/j.zefq.2013.05.007

Planter, A. T. (2011). Writing a data analysis plan. In W. Pequegnat, E. Stover, & C. A. Boyce (Eds.). *How to write a successful research grant application: A guide for social and behaviour scientists* (2nd ed.) (pp. 283-298). Boston, MA: Springer US.

Polit, D. F., & Beck, C. T. (2017). *Nursing research: Generating and assessing evidence for nursing practice* (10th ed.). Philadelphia, PA: Walters Kluwer/Lippincott Williams & Wilkins.

Reese, C. E., Jeffries, P. R., & Engum, S. A. (2010). Learning together: Using simulations to develop nursing and medical student collaboration. *Nursing Education Perspectives*, *31*(1), 33-37. Retrieved from https://journals.lww.com/neponline/

Rehim, S. A., DeMoor, S., Olmsted, R., Dent, D. L., & Parker-Raley, J. (2017). Tools for assessment of communication skills of hospital action teams: A systematic review. *Journal of Surgical Education*, *74*(2), 341-351. doi:10.1016/j.jsurg.2016.09.008

Rosen, M. A., Weaver, S. J., Lazzara, E. H., Salas, E., Wu, T., Silvestri, S., . . . King, H. B. (2010). Tools for evaluating team performance in simulation-based training. *Journal of Emergencies, Trauma and Shock*, *3*(4), 353-359. doi:10.4103/0974-2700.70746

Sanfey, H., McDowell, C., Meier, A. H., & Dunnington, G. L. (2011). Team training for surgical trainees. T*he Surgeon*, *9*, S32-S34. doi:10.1016/j.surge.2010.11.018

Scheckel, M. (2016). Designing courses and learning experiences. In D. M. Billings & J. A. Halstead (Eds.). *Teaching in nursing: A guide for faculty* (pp. 159-185). St. Louis, MI: Elsevier.

Scherer, Y. K., Myers, J., O'Connor, T. D., & Haskins, M. (2013). Interprofessional simulation to foster collaboration between nursing and medical students. *Clinical Simulation in Nursing, 9*(11), e497-e505. doi:10.1016/j.ecns.2013.03.001

Seelandt, J. C., Tschan, F., Keller, S., Beldi, G., Jenni, N., Kurmann, A., . . . Semmer, N. K. (2014). Assessing distractors and teamwork during surgery: developing an event-based method for direct observation. *BMJ Quality and Safety*, *23*(11), 918-929. doi:10.1136/bmjqs-2014-002860

Simpson, S. H. (2015). Creating a data analysis plan: What to consider when choosing statistics for a study. *The Canadian Journal of Hospital Pharmacy*, *68*(4), 311-317. Retrieved from https://www.cjhp-online.ca/

Smithburger, P., Kane-Gill, S., Kloet, M., Lohr, B., & Seybert, A. (2013). Advancing interprofessional education through the use of high fidelity human patient simulators. *Pharmacy Practice, 11*(2), 61-65. Retrieved from https://www.pharmacypractice.org/

Sorbero, M. E., Farley, D. O., Mattke, S., & Lovejoy, S. L. (2008). *Outcome measures for effective teamwork in inpatient care*. Santa Monica: CA. Rand Corporation.

Speakman, E (2016). Interprofessional Education and Collaborative Practice. In D. M. Billings & J. A. Halstead (Eds.). *Teaching in nursing: A guide for faculty* (pp. 186-196). St. Louis, MI: Elsevier.

Stewart, M., Kennedy, N., & Cuene-Grandidier, H. (2010). Undergraduate interprofessional education using high-fidelity paediatric simulation. *The Clinical Teacher*, *7*(2), 90-96. doi:10.1111/j.1743-498X.2010.00351.x

Tofil, N. M., Morris, J. L., Peterson, D. T., Watts, P., Epps, C., Harrington, K. F., . . . White, M. L. (2014). Interprofessional simulation training improves knowledge and teamwork in nursing and medical students during internal medicine clerkship. *Journal of Hospital Medicine*, *9*(3), 189-192. doi:10.1002/jhm.2126

Van Schaik, S. M., Plant, J., Diane, S., Tsang, L., & O'Sullivan, P. (2011). Interprofessional team training in pediatric resuscitation: A low-cost, in situ simulation program that enhances self-efficacy among participants. *Clinical Pediatrics*, *50*(9), 807-815. doi:10.1177/0009922811405518

Weaver, S. J., Rosen, M. A., DiazGranados, D., Lazzara, E. H., Lyons, R., Salas, E., . . . King, H. B. (2010). Does teamwork improve performance in the operating room? A multilevel evaluation. *The Joint Commission Journal on Quality and Patient Safety*, *36*(3), 133-142. Retrieved from http://www.jointcommissionjournal.com

Weller, J., Frengley, R., Torrie, J., Shulruf, B., Jolly, B., Hopley, L., . . . Paul, A. (2011). Evaluation of an instrument to measure teamwork in multidisciplinary critical care teams. *BMJ Quality and Safety*, 20(3), 216-222. doi:10.1136/bmjqs.2010.041913

World Health Organization. (2010). Framework for action on interprofessional education and collaborative practice. Retrieved from http://www.who.int/hrh/resources/framework_action/en/

Zangaro, G. A., & Soeken, K. L. (2007). A meta-analysis of studies of nurses' job satisfaction. *Research in Nursing and Health*, *30*(4), 445-458. doi:10.1002/nur.20202

Appendix B
Literature Summary Tables

Bold and italicized texts are my impressions and critiques of the research study.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Garbee et al., (2012) **Type of study**: Quasi-experimental, pre-test/post-test design. ***Moderate design. No randomization limits the design strength.*** **Objective**: Evaluate the efficacy and retention of teaching team-based competencies to interprofessional student teams using high-fidelity simulation. **Ethics:** Explicit approval was outlined in article and informed consent obtained. This literature summary table will only focus on CATS assessment results.	**Sample:** n=35 took part in fall session. n=25 took part in spring session. ***Small sample size, poor retention of participants between sessions.*** Convenience sample of senior level students in medicine, nurse anesthesia, nursing, and physical therapy. **Setting:** Simulation centre at 1 health centre in USA, sessions took place 6 months apart. ***Single setting limits generalizability.***	**Methodology:** Teams consisted of 2 medical students, 2 nursing students, 2 nursing anesthesia students, and 2 physical therapy students. Took part in different 2 standardized scenarios in fall and spring sessions. 4 simulations in total. **Independent Variable:** Hi-fidelity simulation for interprofessional teams. **Dependent Variable:** Teamwork behaviour as measured by observer evaluations. **Instruments:** Communication and Teamwork Skills assessment tool. **Analysis:** Paired-sample t-tests to compare mean item and subscale scores between fall and spring sessions and between scores after each scenario. ***Proper statistical methods chosen.*** Statistical significance was set at $p < 0.05$.	Fall 2009: Statistically significant increase in CATS subscales of Situational Awareness, Cooperation, and Communication. No statistical significant increase in Coordination subscale. Spring 2010: No statistically significant increase in any subscales of CATS. Retention: CATS assessment scores not statistically different from scenario 2 in fall to scenario 1 in spring. CATS assessment scores significantly improved in all subscales from scenario 1 in fall to scenario 2 in spring. **Conclusion:** Conflicting results regarding statistical significance of findings surrounding CATS and retention of skills. Researchers believed their results showcase that HF-IPE can improve observed competencies, with retention over time.	**Strengths:** Standardized simulations used help control for potential confounding variables. Participants from a variety of disciplines does help with generalizability of results. Equal team structure makes it easier to compare within and between groups. **Limitations:** Low retention rate between sessions (lost 28% of participants). Small sample size limits generalizability. Convenience sample may not represent the general population of interest. Quasi-experimental design may lead to non-equivalent groups due to lack of randomization, limits generalizability and increases threats to internal validity. **Critique:** Despite conflicting results regarding significant findings and limitations, I believe the methodology for this study was sound. A similar design with a larger sample size may produce findings that are more indicative of the general population.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Garbee et al., (2013) **Type of study**: Quasi-experimental, pre-test/post-test design. This was chosen to ensure all willing participants can participate, instead of excluding individuals due to randomization. ***Moderate design. No randomization limits the design strength.*** **Objective**: Evaluate the efficacy of using crisis resource management principles and HF-IPE for training students. **Ethics:** Explicit approval was outlined in article and informed consent obtained from participants. This literature summary table will only focus on CATS assessment results.	**Sample:** n=52 in fall session, n=40 returned for spring session. ***Small sample size, poor retention of participants between sessions.*** Convenience sample of students from undergraduate nursing and respiratory therapy and graduate medicine and nurse anesthesia. Respiratory therapy student were the only junior students on the team ***Junior member may not have skills or confidence to participate in HF-IPE sessions. Team members of varying ability may interfere with results.*** **Setting:** Simulation centre at 1 health centre in USA, sessions took place 5 months apart. ***Single setting limits generalizability.***	**Methodology:** Teams consisted of at least 1 but not more than 2 students from each discipline. ***No standardized team make-up could limit the ability to compare results between teams.*** Took part in different 2 standardized scenarios in fall and spring sessions. 4 simulations in total. **Independent Variable:** Hi-fidelity simulation for interprofessional teams. **Dependent Variable:** Teamwork behaviour as measured by observer evaluations. **Instruments:** Communication and Teamwork Skills assessment tool. **Analysis:** Paired t-tests used to compare mean scores between scenarios. ***Proper statistical methods chosen.*** Statistical significance was set at p value < 0.05.	Fall 2009: CATS scores saw significant increases in all four subscales: Coordination, Situational Awareness, Cooperation, and Communication. Spring 2010: Only significant increases in Situational Awareness and Cooperation subscales. Retention: Mean observer scores were not significantly different between simulations 2 in Fall 2009 to simulation 1 to Spring 2010. CATS scores significantly improved in all subscales from scenario 1 in fall to scenario 2 in spring. **Conclusion:** HF-IPE can be effective at improving teamwork and communication skills among students participating in IP teams.	**Strengths:** Standardized simulations used help control for potential confounding variables. Participants from a variety of disciplines does help with generalizability of results. **Limitations:** Low retention rate between sessions (lost 23% of participants). Convenience sample may not represent the general population of interest. Quasi-experimental design may lead to non-equivalent groups due to lack of randomization, limits generalizability and increases threats to internal validity. Different team structures limit ability to compare results between groups. **Critique:** Despite conflicting results regarding the significant findings and limitations I believe the methodology for this study was sound. A similar design with a larger sample size and similar team structures may produce findings that are more indicative of the general population.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Smithburger, Kane-Gill, Kloet, Lohr, & Seybert, (2013) **Type of study**: Quasi-experimental, repeated measures design. **Objective**: To determine if HF-IPE is an accepted and effective approach to improving communication and teamwork skills for students from various disciplines. **Ethics:** Institutional Review Board approval explicitly outlined. No mention of informed consent, but participants volunteered to participate. ***Study should explicitly discuss informed consent.***	**Sample:** All students were senior level. Consisting of pharmacy, medicine, nursing, social work and physician assistants. Participants volunteered to participate. n = 8. ***Small sample size, only one IP team. Limits generalizability of results.*** ***This convenience sample may limit generalization of results as certain subjects are not part of the sample.*** **Setting:** Took place in single facility. With students recruited from same school. ***Single setting limits generalizability.***	**Methodology:** One IP team took part in HF scenarios. Four simulation scenarios occurred once a week over a four-week span. ***No reference to standardization of simulations. Limits ability to generalize results.*** **Independent Variable:** High fidelity simulation for interprofessional teams. **Dependent Variable:** Teamwork behaviour as measured by observer evaluations. **Instruments:** Communication and Teamwork Skills assessment tool. **Analysis:** ANOVA with Bonferroni was used to compare CATS scores between different sessions. ***Proper statistical methods chosen.*** Statistical significance was set at p value < 0.05.	CATS scores statistically improved from session 1 to 2 ($p=0.01$) from session 2 to 3 ($p=0.035$) and from session 1 to 4 ($p=0.001$). No significance between session 3 to 4 ($p=0.07$). Inter-rater reliability (0.85) was high among independent evaluators. **Conclusion:** Using HF-IPE can improve student's teamwork and communication skills. This source of teaching modality should be encouraged into education curriculums.	**Strengths:** Observers trained in CATS, study also assessed inter-rater reliability and determined high level of agreement between scores. Simple design is easy to replicate. Variety of healthcare students from various disciplines may make results more generalizable to real-world situations. **Limitations:** Small sample size only lead to one IP team, unable to compare results to another team. Non-randomization of participants limits generalization of results. Convenience sample may not represent the general population of interest. **Critique:** This study is simple in its design which allows for easy replication. While the results are promising, a larger sample size would be necessary to generate results that may be generalizable to the target population.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Joshi et al., (2017) **Type of study**: Between subjects and within subjects repeated measures design. **Objective**: Determine the degree to which team familiarity affected team competencies including teamwork, team clinical effectiveness, and team situation awareness. **Ethics:** Study had approval by the Institutional Review Board. ***No discussion surrounding informed consent. Should be discussed in article.***	**Sample:** 1st year residents (general surgery and emergency medicine) recruited and randomly assigned to stable or dynamic team. n = 46 participants randomly split into 8 stable teams and 7 dynamic teams. ***Randomization of participants is a strength.*** ***Only included medical residents from different disciplines, not a true interprofessional team. Cannot generalize results to real-life IPE situations.*** **Setting:** Participants took part in all simulations over an 8-hour span during their orientation day. 3 simulations were programmed into three separate simulation rooms. Took part in all scenarios in the same sequence. ***Same sequence of simulations by all participants. Can help control for confounding variables.***	**Methodology:** A series of HF simulations were created and members had to complete the simulation as part of either a stable team or a dynamic team. The stable team had the same team members for every simulation. The dynamic team had fluctuating team members during each simulation. Debriefing took part after each simulation. **Independent variable:** Stable team or dynamic team condition during HF simulation. **Dependent variable:** Teamwork behaviour. **Instrument:** Communication and Teamwork Skills assessment tool. **Analysis:** Paired sample t-tests used to assess changes in teamwork behaviour between different simulations. Independent samples t-test were used to assess differences in stable and dynamic group results. Statistical significance was set at $p < 0.05$. ***Appropriate statistical test chosen given the context of the research study.***	Both stable teams and dynamic teams showed a statistically significant improvement in their team scores from simulation 1 to simulation 3. Stable teams did perform better than the dynamic teams, but these differences were not statistically significant ***No breakdown was given regarding domain scores, only discussed overall team scores. Difficult to make inferences regarding different domains of the CATS assessment tool*** **Conclusion:** Regardless of team structure, simulation training can have a positive benefit on teamwork behaviour. Simulation training should be utilized more within professional settings to ensure healthcare professionals can hone their critical teamwork skills.	**Strengths:** Multiple trained observers help determine inter-rater reliability. Randomization between stable and dynamic teams. Fair sample size. **Limitations:** Only using medical residents and not including any other professionals may limit generalizability of results. Limited information regarding CATS scores besides the overall team scores. **Critique:** Despite not including other professions, the results seem to indicate that stable and dynamic teams can benefit from HF simulations. However, these results are not generalizable. Future research should focus on HF-IPE with stable and dynamic teams to determine if these benefits also exist for different IP teams.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Hughes et al., (2014) **Type of study**: Observational, Pre-test/post-test design. **Objective**: Evaluate the effectiveness of a team-building process in resuscitation of trauma to determine if Crisis Resource Management (CRM) education can improve teamwork and communication among working professionals. **Ethics** Ethics board approval was explicitly outlined within the study. ***Only observation so informed consent was not necessarily needed.***	**Sample:** n=25 pre-CRM resuscitations were observed and scored using the CATS. n=38 post-CRM resuscitations were observed and scored using the CATS. ***Large amount of data collected pre and post education.*** **Setting:** Took place in a single hospital setting. ***Education was tailored to meet the identified weaknesses of one department, future education programs have to be tailored in a similar fashion.*** ***Cannot generalize results to other stable and dynamic teams.*** ***Single setting limits generalizability.***	**Methodology:** A steering committee created a crisis resource management program that included didactic classroom education. Pre-post assessments were completed to determine if the education sessions could have benefits to teamwork and communication skills among working professional responding to trauma resuscitations. The CATS used to assess teamwork behaviour and identify weaknesses. These weaknesses would be incorporated into an education program and the group would be tested again. **Independent variable:** Education sessions focused on CRM. **Dependent variable:** CATS assessment scores during traumatic resuscitations. **Instruments:** Communication and Teamwork Skills assessment tool. **Analysis:** Chi-square and Fisher's exact test were used to determine if there was a significance difference between scores pre and post education. ***Proper statistical methods chosen.*** No reference was made to what *p* value was considered statistically significant.	Significant improvements (p <0.001) were noted in the behaviour markers of briefing, verbalizing plan of care, establishing a team leader, assigning roles, using names, verbal update-think aloud, closed loop, cross monitoring, ask for help from team as needed, request external resources as needed, and giving patient summary to trauma personal. No information was discussed regarding the 4 sub-categories scores or overall scores pre to post education. **Conclusion:** CRM education can improve teamwork and communication skills among working professionals taking part in trauma resuscitation.	**Strengths:** Observers trained in the CATS assessment. Large number of assessment scores collected pre and post education. **Limitations:** Limited generalization of results due to the tailored nature of education program. Only one observer trained may limit reliability of scoring. No discussion of who made up resuscitation team members. No discussion regarding pre/post scores for four domains and overall CATS scores. **Critique:** Only study reviewed that focused on tailoring education to meet weaknesses identified by CATS, and then showed improvement in those weaknesses post education. Despite limitations, the information is promising in asserting that education programs can be tailored to improved weaknesses in communication and teamwork skills that are identified by a CATS assessment.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Reese, Jeffries, & Engum, (2010) **Type of study**: Exploratory, descriptive design with simple exposure. **Objective**: Investigate the use of HF-IPE as a way to support collaboration between nursing and medical students. **Ethics:** Approval was received from the Institutional Review Board. Stated that researchers explicitly obtained informed consent.	**Sample: S**enior year nursing and medical students. Convenience sample of n=13 medial students and n=15 nursing students. ***Small sample and convenience sample do not allow for generalizability of results.*** **Setting:** Participants recruited from single university setting. ***Single setting limits generalizability.***	**Methodology:** 4 students involved in each simulation (2 nursing and 2 medicine). A single simulation was designed for this study. Simulation was based on code blue scenario. **Independent variable:** HF-IPE. **Dependent variable:** Self-confidence, perceptions, satisfaction, and collaboration results. **Instrument:** Simulation Design Scale. Satisfaction and Self-Confidence Scale. ***Strong tools due to previously established reliability.*** Collaboration scale developed by researchers. ***No established validity or reliability for this tool.*** **Analysis:** Descriptive statistics to investigate open ended question themes. Independent samples t-test assess differences in nursing and medical scores related to educational design, self-confidence, satisfaction, and collaboration. ***Proper statistical methods chosen.*** No reference was made to what *p* value was considered statistically significant.	Both groups believed independent problem solving, appropriate feedback, timely feedback, was accomplished with HF-IPE. No significant differences were noted between nursing and medical students on all measured scores. Themes emerged for qualitative analysis: 1) Interaction with other disciplines were perceived as beneficial. 2) Simulation reflected well on real-life situations and providing a safe learning environment. 3) Being part of a code simulation was beneficial to their experiences when preparing for real-life code situations. 4) Fear and uncertainly of role during simulation was common among participants. **Conclusion:** Findings support the evidence that HF simulations can support a student's education.	**Strengths:** Focused on both qualitative and quantitative data. Most data collection instruments had previously established reliability. **Limitations:** Small sample size and nature of participant recruitment do not allow for generalizability of results. Single HF simulation experience does not allow for generalizations to all HF-IPE experiences. Only two disciplines included in IP teams, not reflective of real life code blue scenarios. **Critique:** The results are promising that HF-IPE has perceived benefits from the participants perspectives. More research is needed that focuses on larger sample sizes, IP teams with a variety of healthcare students, and objective measures to evaluate performance in these HF-IPE scenarios.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Jakobsen et al., (2018) **Type of study**: Post-intervention study design. ***Design does not allow for controlling of external variables or allow for generalization of results. Poor design.*** **Objective**: Describe the adaptation of an IP simulation course in an undergraduate setting and to report participant's experiences. **Ethics:** Project was approved by educational leadership. Participation was part of course and students were informed the data would be evaluated and published. ***Participants did not get a choice to opt out of participation. No consent obtained. Participants should have been given choice to have their data included or excluded in the study.***	**Sample:** n=310 ***Poor description of sample, in reality only a maximum of 165 students participated, but they determined n equaled the pre-simulation + post-simulation, despite some participants could have filled out both surveys.*** Students included medical, nursing, and nursing anesthesia. Participation was mandatory component of education. **Setting:** Simulation setting at University of Oslo. Performed 4 simulations. Debriefing after every simulation. ***Single setting limits generalizability.***	**Methodology:** A one day HF-IPE course focusing on an emergency room setting. Adapted the Better and Systematic Team Training. Teams consisted medical, nursing, nursing anesthesia. ***Poor team design, no randomization, not equal team structures. Difficult to compare between teams.*** Data collected using questionnaires before and after simulations. **Independent variable:** HF-IPE. **Dependent variable:** Self-reported experiences. **Analysis:** Systematic text condensation to assess thematic analysis of data. Descriptive statistics analyzed quantitative questionnaires. Pairwise comparisons using Dunn's procedure with Bonferroni correction for multiple comparisons. ***Proper statistical methods chosen.*** Statistical significance was set at p value < 0.05.	145 students completed survey pre-simulation. With 165 students responded to questionnaire post-simulation. Qualitative results: 1) Students believed simulations created an emotional activation of their engagement. 2) Simulations often lead to learning surrounding leadership roles. 3) Students felt they gained insights into teamwork and the value of communication. Quantitative results: 1) Medical students found facilitator feedback less helpful than nursing. 2) No difference between students when assessing communication. 3) Nursing anesthesia found role-tagged vests more useful than medical students 4) Medical students found more benefits surrounding leadership when compared to nursing students. **Conclusion:** The modified Student-BEST course was perceived as beneficial for students from various disciplines.	**Strengths:** Large sample size. Adapted a previously validated program. Simulation design and debriefing after each session. **Limitations:** Mandatory participation of participants. No consent. Self-assessments may over-estimate or over-estimate. Questionnaire did not undergo formal validation process. Poor team design, not equal team structures. **Critique:** The study does provide some useful information surrounding how students perceive benefits from HF-IPE. But limitations and poor design lead the results to be questioned and limit generalizability of results. More objective measures would be beneficial to determine if improvements in teamwork and communication come from participating in the designed student-BEST program.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
King et al., (2014) **Type of study**: Quasi-experimental, repeated measures design. **Objective**: Develop, deliver, and assess HF-IPE versus LF-IPE for pre-licensure students from different disciplines. **Ethics:** Consent was obtained prior to simulation setting. ***No reference to if ethical approval was obtained.***	**Sample:** n=78. Convenience sample of participants who volunteered to participate. Most were senior level students. ***Convenience sample may limit generalization of results as certain subjects are not part of sample.*** Students self-selected what simulation they wanted to participate it. ***Self-selection is a poor method of dividing sample.*** HF simulation used medicine, nursing, LPN, and respiratory therapy. LF simulation used respiratory therapy, nursing, OT, PT, recreational therapy, therapy assistant, health care aide and pharmacy. ***Involved more than two disciplines, more reflective of real-world situations.*** **Setting:** Students recruited from four different post-secondary institutions in Edmonton, Canada. ***Single setting limits generalizability.***	**Methodology:** Participants completed self-reported questionnaire pre and post simulation. Participants either took part in HF simulation or a LF simulation. Debriefing took place after each simulation focusing on communication and teamwork behaviour. **Independent Variable:** HF-IPE scenario and LF-IPE scenario. **Dependent Variable:** Self-reported changes in communication and teamwork behaviours. **Instrument:** University of the West of England Interprofessional Questionnaire. ***Previous reliability of tool established and team reassessed this tool to determine that only one subscale had internal consistency within their acceptable range.*** **Analysis:** Cronbach's alpha values for each subscale and full instrument. Paired t-test and repeated measures ANOVA also completed. Statistical significance was set at p value < 0.01. ***Proper statistical test chosen to analyze data.***	Regardless of simulation, participants perceived skills improved. Statistical significance was only achieved on two measured items: 1) I prefer to stay quiet when other people in a group express opinions that I don't agree with ($p=0.003$) 2) I am able to become quickly involved in new teams and groups ($p=0.002$). Total score also saw a statistically significant improvement ($p=0.004$). **Conclusion:** HF-IPE should be expanded beyond just nursing and medicine. The key is to create a simulation environment (HF or LF) that is reflective of a real-world situation. Teams should be devised based on relevancy to real-world practice, and not basing groups based upon academic credentials.	**Strengths:** Using a variety of participants from numerous disciplines makes the results likely to be more generalizable. Thorough analysis of the data while also using the subscale of the instrument that they decided meet their requirements for reliability and validity. **Limitations:** Small sample size did now allow for three way analysis of factors. No randomization of participants between test groups. Inconsistent durations between simulations (HF was 1 hour, LF was 3 hours). Self-assessments may over-estimate or under-estimate skill improvements. **Critique:** The study provided great insight in IPE simulations that reflect real-life simulations. While the sample size was small for three-way-analysis and self-reporting does not provide concrete evidence towards object improvements, the results are promising.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Dillon, Noble, & Kaplan, (2009) **Type of Study**: Pre-test, post-test design. **Objective**: Initiate an IP collaborative process between nursing and medical students and determine if HF-IPE can be perceived as useful by the participants. **Ethics**: Approval was obtained from review board. No mention of informed consent. ***While volunteering consent may be implied, informed consent should have still been obtained.***	**Sample:** n=82 senior level nursing and medical students pre-test. n=40 completed post-test questionnaire. ***Low retention rate. Blamed on scheduling conflicts.*** Convenience sample. ***This convenience sample may limit generalization of results as certain subjects are not part of the sample.*** **Setting:** A single simulation took place. Participants were recruited from a single educational setting. ***Single setting limits generalizability.***	**Methodology:** A mock-code HF simulation was developed for an IPE exercise for nursing and medical students. Measure students' perceptions of HF-IPE. Also asked open-ended questions. Debriefed after the simulation. **Independent Variable:** HF-IPE. **Dependent Variable:** Self-reported attitude's and beliefs related to IPE. **Instrument:** Jefferson Scale of Attitudes Toward Physician-Nurse Collaboration. ***Instrument has good reliability that was calculated by researchers.*** Also asked opened-ended questions. **Analysis:** ANOVA to detect differences between nursing and medical student pre and post test scores. Statistical significance was set at p value < 0.05. Content analysis of open-ended questions. ***Proper statistical test chosen to analyze data.***	Nursing students exhibited higher pre-test scores, but medical students exhibited higher post-test scores. Medical students had statistically significant gains in the areas of collaboration (p=0.013) and nursing autonomy (p=0.025). Themes emerged from the qualitative data: 1) Medical students had mixed feelings regarding a nurse prior to simulation, but felt their role was necessary post-simulation. 2) Nurses believed the physician would have final say on decision making in pre-test, but believed the teamwork was more important after the simulation. 3) Both disciplines felt excited entering the HF-IPE, and felt it was a positive experience post HF-IPE. **Conclusion:** Findings seem to indicate that HF-IPE can be a useful educational strategy within healthcare curriculums. These experiences can ultimately lead to more positive IP experiences and can potentially lead to improved patient outcomes.	**Strengths:** Qualitative and quantitative data collected with proper statistical analysis used. Simple research design that could be easily replicated. **Limitations:** Convenience sample and small sample size from a single setting limits generalizability of results. Poor retention between pre and post-test (lost 51% of participants), making the results difficult to generalize to target population. Qualitative data contradicted quantitative data at times, and researchers offered no explanation as to why. **Critique:** Overall, this paper provides limited reliable evidence into the role HF-IPE may play in teamwork and communication behaviours. This is due to the limitations outlined and conflicting evidence that was not properly explained within the article.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Tofil et al., (2014) **Type of Study**: Repeated measures, pre-test, post-test design. **Objective**: To determine if simulation training would improve nursing and medical students' knowledge, communication skills, and understanding of each other's professional role. **Ethics:** Institutional review board approved this study. No reference to informed consent. ***Informed consent should have been obtained.***	**Sample:** Senior level nursing and medical students. n=108 participated. n=100 completed pre and post-test. ***Fair sample size. Good retention of participants. But no mention of how they were recruited.*** Each team had 3 nursing students and 5-6 medical students. ***Consistent structure of teams, allows for comparison between groups.*** **Setting:** University of Alabama setting. Data collected From July 2011 to April 2012. ***Ten month span of data collection. Strength of the study.*** ***Single setting limits generalizability.***	**Methodology:** Medical and nursing students participated in 4 HF simulations over an 8 week span. Debriefed after each simulation. Pre-tests and post-tests was completed. Open-ended questionnaire also completed. **Independent Variable:** HF-IPE. **Dependent Variable:** Self-reported attitude's and beliefs related to IPE. **Instrument:** Case-specific knowledge questionnaire, self-efficacy questionnaire. ***Non-validated data collection instruments.*** **Analysis:** Self-efficacy scale was examined using Cronbach's alpha. Two-tailed t-tests to determine differences between pre and post test results. Statistical significance was set at *p* value < 0.05. Content analysis of open-ended questions. ***Proper statistical test chosen to analyze data.***	Both medical and nursing students showed significant improvements in self-efficacy scores (p<0.0001). Students from both disciplines felt this activity was applicable to their field and beneficial to their educational experience. Both disciplines felt the exercise increased medical knowledge, improved a sense of teamwork, and improved sense of communication. **Conclusion:** HF-IPE for nursing and medical students can potentially increase communication self-efficacy as well as improve attitudes towards team roles.	**Strengths:** Moderate sample size. Retention rate of participants. Length of data collection. Consistent team structure. More than one simulation experience allowed for teamwork to build over time. **Limitations:** Difficult to make inferences regarding information learned due to the two month span between pre-test and post-test. Self-reported findings may over-estimate or under-estimate skill improvements. Non-validated instruments. **Critique:** Good study design, although used non-validated data collection instruments. More information should also have been included surrounding participant recruitment and informed consent. A similar designed study with validated tools could provide valuable information surrounding HF-IPE.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Paige et al., (2014) **Type of Study**: Quasi experimental, pre/post intervention comparison design. **Objective**: Investigate the immediate impact of an HF-IPE on attitudes and behaviours. **Ethics:** Institutional Review Board granted approval for this study. ***No mention of informed consent. It should be discussed within the paper.***	**Sample:** n=66 consisting of medical, nursing and nurse anesthesia. ***No information regarding how it recruited participants or if they were random assigned to teams.*** Each team had 2 medical, 2 nursing, and 2 nurse anesthesia students. Except a few teams had extra medical students. ***Consistent structure of teams, allows for comparison between groups.*** **Setting:** Academic urban health sciences centre. ***Single setting limits generalizability.*** Each session was 2 hours in length.	**Methodology:** 10 HF-IPE sessions took place within 2 standardized scenarios that focused on an operating room setting. Each session was 2 hours in length. Debriefing took place after each simulation. Pre-tests and post-tests were completed. Trained observers also assessed team-based performances in each simulation. **Independent Variable:** HF-IPE. **Dependent Variable:** Attitudes and behaviours surrounding HF-IPE. Overserved team-based performances. **Instruments:** Specifically designed questionnaire asking open-ended questions. ***Instruments validity not discussed.*** Operating Room Teamwork Assessment Scales (ORTAS). **Analysis:** Paired t-tests and Bonferroni adjustments were completed to analyze data from questionnaires. Qualitative analysis used to identify themes from data. One-way ANOVA used to determine differences between mean calculations of observed scores between simulations. Paired sample t-test used to compare differences between observer and participants ratings of behaviour. No reference was made to what *p* value was considered statistically significant. ***Good statistical analysis covering many different facets of data analysis.***	Statistically significant ($p<0.0001$) increase in pre to post scores for 11 of 15 self-efficacy measurements. Statistically significant ($p<0.0001$) gains in mean observer-rated performance scores for all 3 subscales of ORTAS. Statistically significant ($p<0.0001$) gains noted within each role of the IP team that was evaluated by observers. Themes that emerged from qualitative analysis of data: 1) Enhanced communication, 2) Positive impact from debriefing and 3) Realism of simulation. Some individuals did feel they were not prepare for the HF-IPE, some felt they needed more time, and some believe repeated exposures would have been beneficial. **Conclusion:** A single session HF-IPE focused on an OR setting should be considered both feasible and an effective method of improving teamwork behaviours.	**Strengths:** Consistent team structures for the most part. Standardized simulations. Thorough analysis of the data. Assessed both qualitative and quantitative data, as well as focusing on self-reported and observer-collected data. **Limitations:** Small sample size may limit generalizability of results. Some scenarios had an excess of medial students who only watched one simulation, then switched in the second simulation which does not truly demonstrate changes within simulations. No validity regarding the attitudes and behaviours questionnaire. **Critique:** I believe the strengths of this study outweigh its limitations. This study collected a plethora of data, with a fairly consistent team structure, and had a thorough assessment of the data. Future research should replicate this design study but focus on instruments with established validity.

Article/ Design	Sample/ Settings	Methodology/ Analysis	Results/ Conclusion	Strengths/ Limitations/ Critique
Stewart, Kennedy, & Cuene-Grandidier, (2010) **Type of Study**: Post-intervention study design. **Objective**: Develop, implement and evaluate an HF-IPE program focused on paediatric simulations. **Ethics:** Ethical approval granted from ethics committee. Written consent obtained during student orientation.	**Sample:** Senior level medical and nursing students participated in the HF-IPE program. ***No description regarding how it recruited participants.*** n=95. 100% of participants completed survey. Students were allocated into small IP teams of 3-4 students. ***No description of team structures.*** **Setting:** Simulation room at Queen's University in Belfast. ***Single setting limits generalizability.***	**Methodology:** 6 HF scenarios were created. Groups got to work through 1 scenario while the other teams observed. Debriefing after each simulation. Including both participating and observing participants. Sessions were 20 minutes in length. Questionnaire given to all participants following the HF-IPE program. Included Likert-scale questions and open-ended questions. **Independent Variable:** HF-IPE. **Dependent Variable:** attitudes, behaviour, and experiences related to HF-IPE program. **Instrument:** Questionnaire based on previously published questionnaire. **Analysis:** Cron-Bach's alpha > 0.70 used to determine reliability of questionnaire. ANOVA and student's t-test used to assess quantitative data. Thematic-content analysis for qualitative data. ***Appropriate statistical analysis tests chosen while also determining reliability of questions in developed questionnaire while excluding questions that did not meet the determined Cron-Bach alpha score.*** ***Non-validated data collection instrument.***	No statistically significant differences between medical and nursing students' attitudes surrounding knowledge, communication, teamwork, professional identity, role awareness, and attitudes towards HF-IPE. Qualitative themes that emerged focused on: 1) HF-IPE was considered a better way of learning, 2) IPE provided opportunities to learn from other professions and 3) Increased role awareness was achieved from HF-IPE program. **Conclusion:** HF-IPE can be effective within medical and nursing curriculums. Students evaluated these experiences as positive and this positivity was also reflected in the quantitative data collected.	**Strengths:** Collected qualitative and quantitative data. Results were congruent among different data collected. Good statistical analysis of data while also determining reliability of own instrument. Good retention rate to complete questionnaire. **Limitations:** Lack of information surrounding participant recruitment and team structure. Questionnaire does not have validity. Teams took part in different simulations so experiences may be different. Only self-reported data, which can under-report or over-report findings. **Critique:** This study collected a plethora of data, and had a thorough assessment of the data. The results did seem to indicate that HF-IPE is perceived as beneficial but future research should also focus on observed performance.

Appendix C

Report on the Data Analysis Plan

The overall goal of this practicum project was to create a data analysis plan for the quantitative data obtained from the Communication and Teamwork Skills (CATS) assessment tool (Frankel, Gardner, Maynard, & Kelly, 2007) being used in MacDonald et al.'s (2016) research study "*Measuring the Effectiveness of High Fidelity Simulation in Interprofessional Education to Foster Teamwork Among Undergraduate Nursing, Medicine and Pharmacy Student*s". This data analysis plan will be used to guide the evaluation of the communication and teamwork behaviours observed while nursing, medicine, and pharmacy students practice within an interprofessional team during a high fidelity simulation (HF-IPE) and a low fidelity simulation (LF-IPE). Specifically, the practicum project's focus was to contribute to the data analysis phase of nursing research, as evident by the creation of the data analysis plan for the CATS assessment tool and the proof of concept. The purpose of this report is to outline the evidence used in the creation of the data analysis plan including a brief discussion of current research using CATS as an evaluation instrument, selecting an appropriate statistical analysis method, and creating the SPSS codebooks for data analysis. Proof of concept was implemented using a fictitious data set to test the SPSS codebooks, including description of the data analysis and visual representation of these data. The plan presented in this report is for the analysis of a fictitious data set that is normally distributed, the groups are equal, and participants are randomly assigned to teams.

Communication and Teamwork Skills Assessment Tool

Frankel et al. (2007) created the CATS assessment tool as an instrument to measure communication and teamwork skills of healthcare professionals in the real world and in simulated settings. The CATS assessment tool was designed to assess teamwork behaviors in a broad range of healthcare professionals, including nurses, physicians, social workers, and respiratory therapists (Aliner et al., 2014; Garbee et al., 2012; Garbee et al., 2013; Hughes et al., 2014; Joshi, Hernandez, Martinez, AbdelFattah, & Gardner, 2017; Passauer-Baierl, Baschnegger, Bruns, & Weigl, 2014; Smithburger, Kane-Gill, Kloet, Lohr, & Seybert, 2013).

The CATS assessment tool focuses on directly observing teamwork behaviour while quantitatively gathering data on the observed behaviours. Frankel et al. wanted to develop a quantitative assessment tool that focused on how often and how well particular teamwork behaviours were performed, while also having an opportunity to examine teamwork as a whole. The CATS assessment tool investigates four primary domains of team behaviour including: situational awareness, coordination, communication, and cooperation. Within these four domains, there are 21 behaviour markers that are assessed by a trained observer, including three behaviour markers that are scored if a crisis situation arises. Specific behaviour marker scores need to be added together to determine each respective domain score. For example, the coordination domain is comprised of the following behaviour markers: briefing, verbal plan, verbalize expected outcomes, debriefing, and establish event manager. Behavior markers are scored on the basis of how often an event occurs and the quality of the team’s communication and teamwork behaviours.

Each time a behaviour is observed it produces a raw data score as either “*Good*” = 1 point; “*Variable in Quality*” = 0.5 points, or “*Expected but Not Observed*” = 0 points under the appropriate behaviour marker. The raw data under each behaviour marker is subsequently used to determine raw scores for each of the four domains, and as an overall score. To determine the raw data within each domain, this requires the addition of the raw scores under the corresponding behaviour markers. Likewise, to determine the raw data of the overall score, this requires the addition of the raw scores for all behaviour markers. The raw data collected using the CATS assessment tool is initially calculated into non-weighted total scores. The non-weighted total scores need to be further calculated into a weighted total score. The weighted total score out of 100 is calculated for each individual behaviour marker, each domain, and as an overall score. The weighted total scores can then be used to compare team performance either between teams, or pre and post an intervention, or across two different testing conditions such as HF-IPE and LF-IPE. The data collected using this tool is considered ratio level data.

SPSS Codebooks

Two codebooks were created in SPSS with the first codebook being used to input the raw data and compute the non-weighted total scores along with the weighted total scores, and the second codebook being used to separate the weighted total scores for all variables to allow for data analysis. The creation of two codebooks makes separation of data and analysis an easier process, however, one had to be cognizant of manual transcription errors that could take place when transferring data between codebooks, or when manual addition of the raw data was necessary.

Due to the manual addition and transferring of information, these steps of the process could result in transcription errors. To limit human addition error, a voice command program was used whereby it would automatically add the numbers together as they were read aloud. These numbers were double-checked by manual addition. This method was used to calculate the raw data within each domain and as an overall score. To prevent errors from happening during the transferring of information from the first codebook to the second codebook, both codebooks were opened on the same computer monitor and the cut and paste function was used to manually transfer the data from the first codebook to the second codebook. The cut and paste method prevented manual transcription errors, and having both windows visible at the same time allowed for an easy visualization that the data was being transferred into the appropriate place. These approaches to transferring data helped to prevent transcription errors.

Proof of Concept

For the purpose of testing the SPSS codebooks, data analysis process, and visual representation of data, a fictitious data set was created for seven HF-IPE and seven LF-IPE scenarios. Analysis of this fictitious data set would be used to ensure the SPSS codebooks worked correctly and could produce the desired results if a real collected data set were to be inputted. Since the data set entered was fictitious, there will be no discussion of the findings related to the literature, but the focus will be on describing the statistical methods and visual representation of the fictitious data set and the subsequent fictitious results.

Fictitious Data Generation and Input

Fictitious data for seven teams were created with all seven teams participating in a HF-IPE scenario and in a LF-IPE scenario. Thus, 14 scores were created and the sample size for this proof of concept was $n = 7$. All 14 of the CATS scores included the crisis situation behaviour markers, to ensure that all of the behaviour markers were entered and analyzed. Once all the raw data was inputted, there were 26 different raw data scores for each scenario including: 21 behaviour markers, four domain scores, and one overall score

Using the ‘compute variable’ function with SPSS, the non-weighted total scores and weighted scores were calculated for the 26 different variables for all 14 scenarios. The non-weighted score was obtained adding the *total number of times* a behaviour was observed within each respective behaviour marker. These observed behaviours were scored for each behaviour marker within the following categories: Observed and Good = *GB*; Variable in Quality = *VQB*, and Expected but Not Observed = *NOB*. As stated previously, the behaviour marker raw data scores were used to calculate the raw data within four domains and an overall score. The total number of times the behaviour was observed was added together for a non-weighted total score coded as “A”. See Equation 1.

$$\text{Non-Weighted Total Score (A)} = GB + VQB + NOB \tag{1}$$

A weighted score was then computed for each of the 21 behaviour markers, the four domains, and as an overall score. As part of the process of calculating the weighted total score, the raw data under GB, VGB, and NOB for each variable had to be multiplied by 1.0, 0.5, and 0 respectively. These weighted total scores were the variable of interest because this value allowed for the statistical analysis of comparisons between the HF-IPE and LF-IPE scores. The weighted total score coded as “B” was determined for the 21 behaviour markers, the four domains, and as an overall score. See Equation 2.

$$\frac{(GB \times 1.0) + (VQB \times 0.5) + (NOB \times 0)}{A} \times 100 = \text{Weighted Total Score (B)} \tag{2}$$

The weighted total scores (B) were transferred into the second codebook under the respective HF-IPE and LF-IPE scenarios to allow for a comparison of teamwork behaviors across the different environments. Once all the data was successfully transferred to the second codebook, the data entry was completed and the statistical analysis occurred.

Analyzing the Fictitious Data

This proposed data analysis plan assumes that the data is normally distributed, the groups were equal, and participants were randomly assigned to different teams. In comparing the HF-IPE scores to the LF-IPE scores there were 26 separate analyses completed looking at each behavior marker or value of interest. The fictitious generated data was paired together to simulate seven different groups taking part in one HF-IPE and one LF-IPE. It was determined through consultations with a statistician and the use of a decision tree created by Simpson (2015) that the Paired t-test could be used to analyze the data and determine any statistically significant differences between the HF-IPE and the LF-IPE scores. The confidence intervals were set at 95% with the level of significance of $p < 0.05$. A *p* value set at this significance level would mean that the likelihood of the differences detected between the scores would emerge due to chance only 5% of the time (Knapp, 2016). A *p* value significance level set at less than 0.05 and confidence intervals set at 95% are considered the standard parameters used for many research studies (Polit & Beck, 2017).

For the purpose of this practicum project, the null hypothesis would postulate that HF-IPE and LF-IPE would produce the same scores when assessed using the CATS assessment tool. The alternate hypothesis would postulate that HF-IPE would produce a higher quality score using the CATS assessment tool when compared to the LF-IPE scores. The *p* value would enable either an acceptance or rejection of the null hypothesis and alternate hypothesis.

Fictitious Data Analysis Results and Interpretations

When looking at the weighted total scores, it was evident that the fictitious HF-IPE scenarios scored higher on the CATS assessment tool when compared to the LF-IPE scenarios. When analyzing the Paired t-test scores, 19 out of the 26 variables showed a statistically significant difference between HF-IPE scores and LF-IPE scores. The overall scores, and the four domain scores all had p values < 0.05, and demonstrated that the HF-IPE scores were significantly higher scores when compared to LF-IPE. Figure C1 outlines the range of scores – including the mean scores - during the HF-IPE scenarios and LF-IPE scenarios as they relate to the overall score and the four domain scores.

When analyzing the 21 behaviour markers using the Paired t-test, 14 showed a statistical significance with a p value < 0.05 (Table C1). The behaviour markers that had a p value > 0.05 included: request external resources as needed, verbally request team input, cross monitoring, verbal assertion, receptive to assertion and ideas, communicates with patient, and establish event manager. While these individual behaviour markers do not show a significant difference between HF-IPE and LF-IPE, the overall scores and four domain scores that encompass all the behaviour markers all showed a statistically significant difference. A full list of mean scores for all variables analyzed during the HF-IPE and LF-IPE scenarios, along with the differences between the mean scores, the standard deviations, and the p values calculated using the Paired t-test are presented in Table C1.

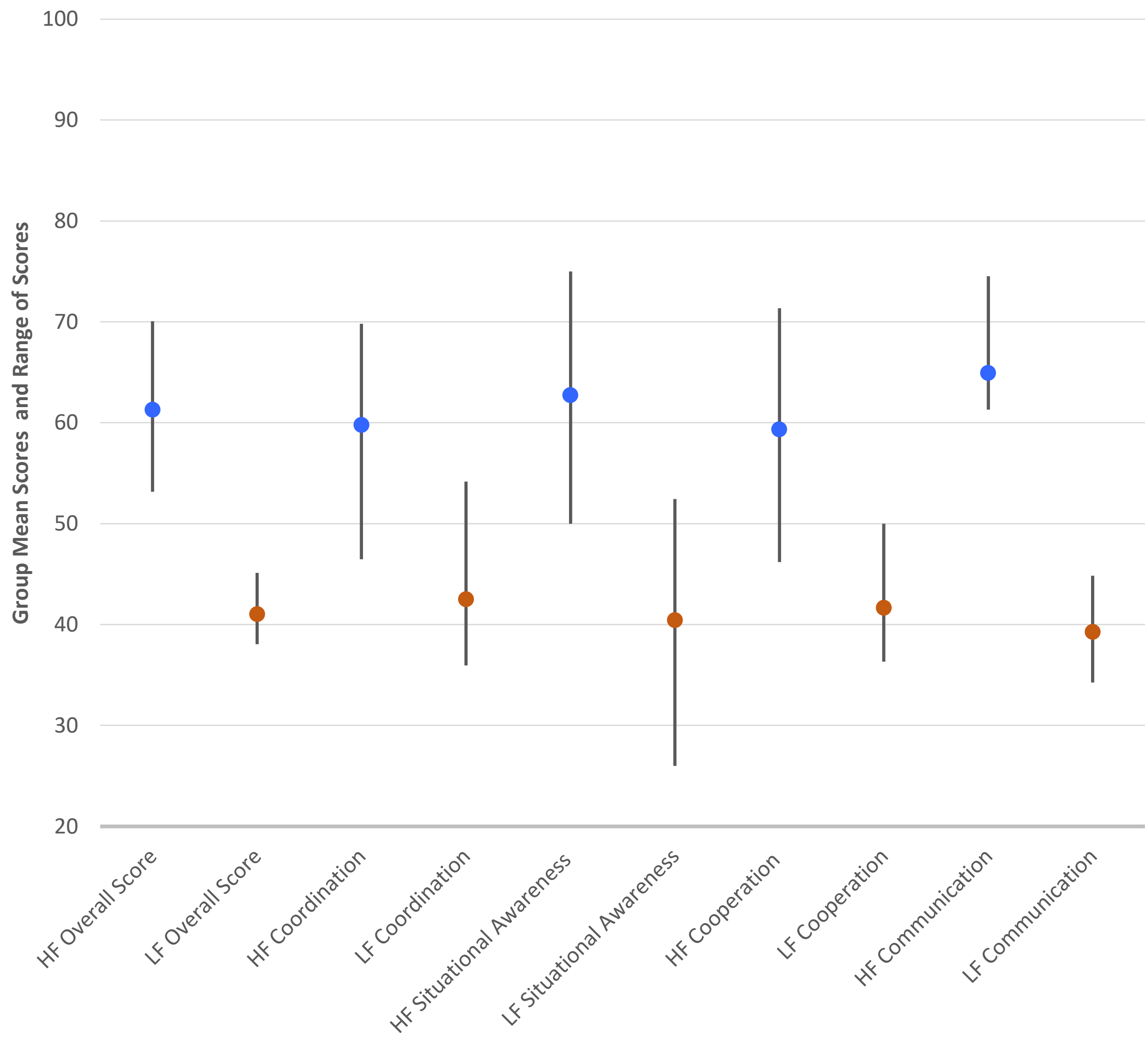


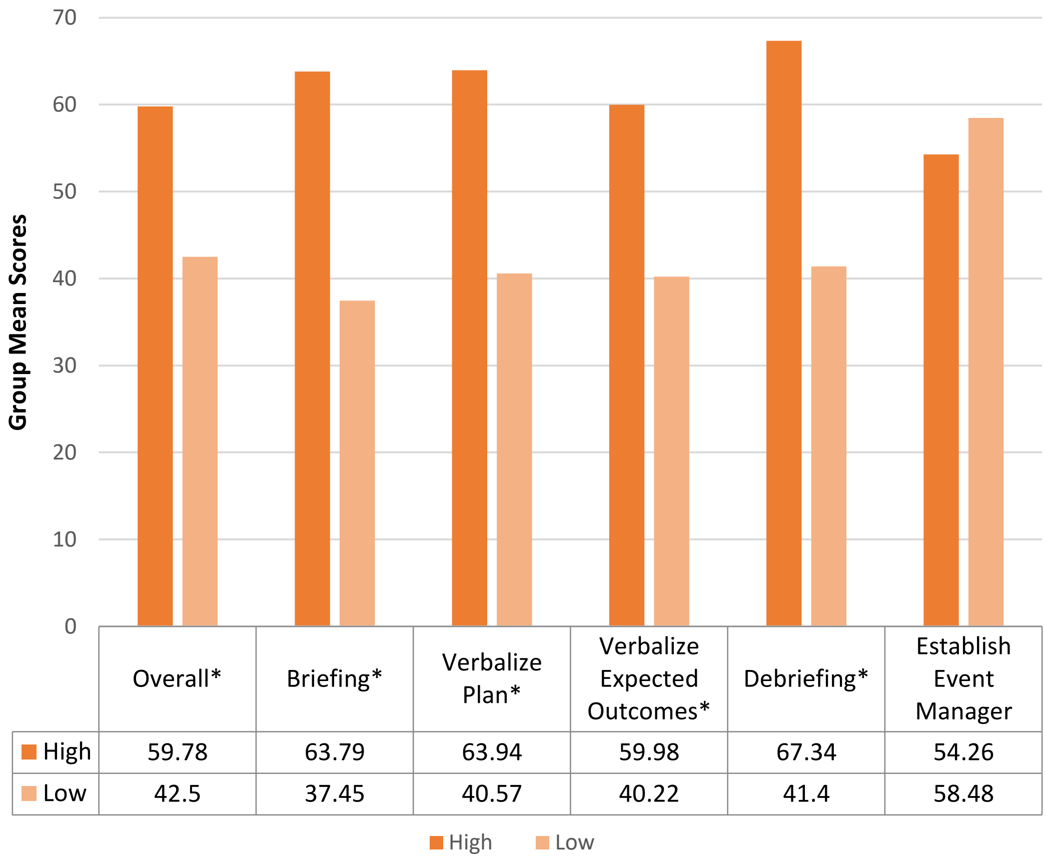
Figure C1. *Group mean scores and range of scores by domains.*

Table C1

Group Mean Scores Comparison between HF-IPE and LF-IPE

Variable	High Fidelity Mean± Standard Deviation	Low Fidelity Mean± Standard Deviation	Difference in Means	*p* value
Overall Group Score	61.27 ± 5.81	41.03 ± 2.25	20.24	.00028
Coordination Domain	59.78 ± 9.17	42.50 ± 6.53	17.28	.001
Situational Awareness Domain	62.73 ± 8.26	40.42 ±10.58	22.31	.002
Cooperation Domain	59.32 ± 9.48	41.66 ± 4.46	17.66	.007
Communication Domain	64.95 ± 5.12	39.28 ± 4.11	25.67	.00023
Briefing	63.79 ± 12.19	37.45 ± 18.30	26.34	.021
Verbalize Plan	63.94 ± 17.02	40.57 ± 8.49	23.37	.001
Verbalize Outcomes	59.98 ± 15.32	40.22 ± 12.06	19.76	.011
Debriefing	67.34 ± 16.39	41.40 ± 9.13	25.94	.008
Visually Scan Environment	64.05 ± 10.01	41.95 ± 8.60	22.10	.008
Visually Adjust Plan as Changes Occur	61.00 ± 10.07	38.06 ± 15.01	22.94	.002
Request External Resources	55.86 ± 19.03	42.27 ± 10.72	13.59	.169
Ask for Help From Team	65.73 ± 17.37	37.75 ± 10.40	27.93	.005
Verbally Request Team Input	64.07 ± 17.27	44.48 ± 23.49	19.23	.181
Cross Monitoring	58.65 ± 17.30	44.93 ± 15.65	13.72	.136
Verbal Assertion	54.19 ± 15.72	46.65 ± 10.40	7.54	.271
Receptive to Assertion and Ideas	57.07 ± 15.15	41.43 ± 7.65	15.64	.075
Closed Loop	71.56 ± 10.23	42.84 ± 10.57	28.72	.008
SBAR	65.76 ± 11.29	38.34 ± 16.21	27.42	.025
Verbal Updates Think Aloud	75.36 ± 15.18	45.27 ± 8.84	30.09	.004
Uses Names	65.15 ± 10.70	39.39 ± 5.87	25.76	.00015
Communicates With Patient	60.62 ± 14.30	41.11 ± 12.99	19.51	.064
Appropriate Tone of Voice	63.68 ± 9.45	36.07 ± 10.71	27.61	.002
Establish Event Manager	54.26 ± 7.30	58.48 ± 12.03	-4.22	.539
Escalation of Concern	61.53 ± 11.97	37.46 ± 12.44	24.07	.009
Critical Language	55.79 ± 12.89	31.72 ± 17.63	23.98	.009

The overall coordination domain score and the respective behaviour marker group mean scores are presented in Figure C2. Four out of the five behaviour markers that influence the coordination domain score were significantly higher in the HF-IPE as compared to the LF-IPE.



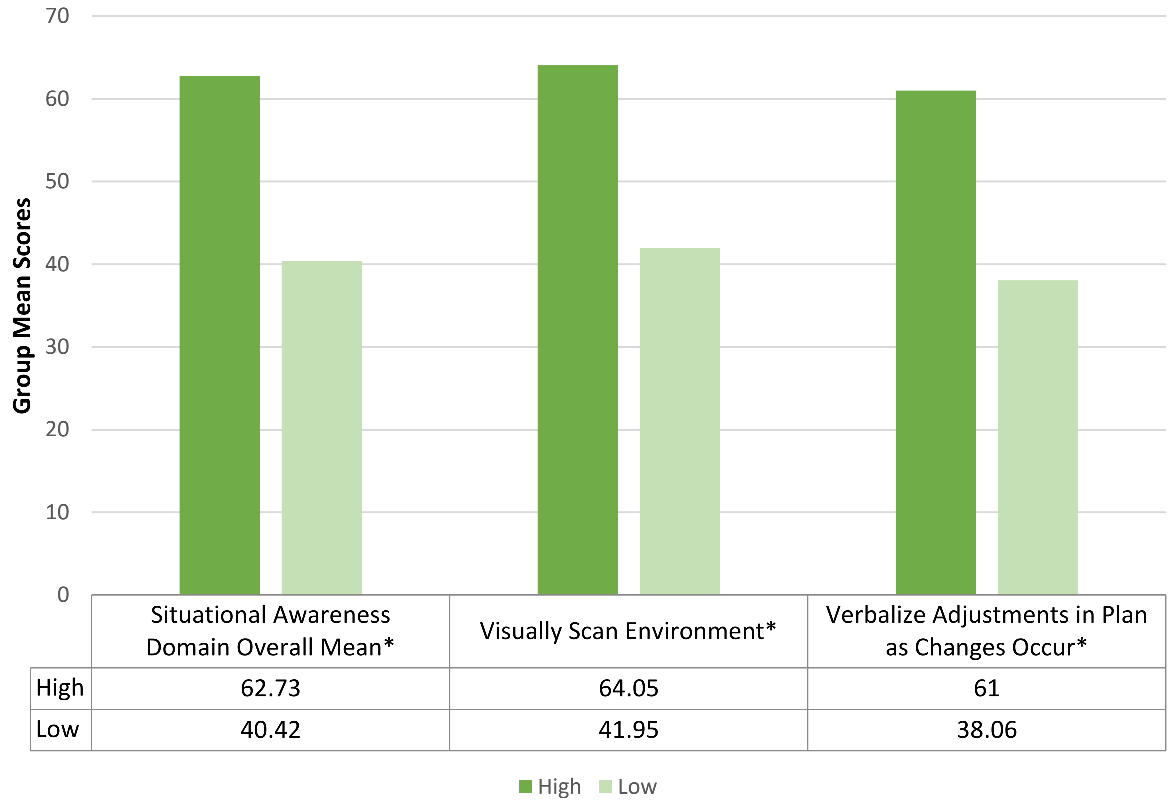
	Overall*	Briefing*	Verbalize Plan*	Verbalize Expected Outcomes*	Debriefing*	Establish Event Manager
High	59.78	63.79	63.94	59.98	67.34	54.26
Low	42.5	37.45	40.57	40.22	41.4	58.48

*p < .05.

Figure C2. Coordination domain, behaviour marker by group mean scores for high and low fidelity interprofessional simulation.

One negative mean score difference, which was the establish event manager score was not considered significant (p = 0.539). The small sample size of this fictitious data set (n = 7) could have had a significant influence on the standard deviation and thus affected the volatility of the data. Knapp (2016) believed that while the t-test can be completed on any sample size, for a t-test to be considered robust the sample size should be greater than 30 subjects. Future research using a larger sample size could produce results that could be considered more robust.

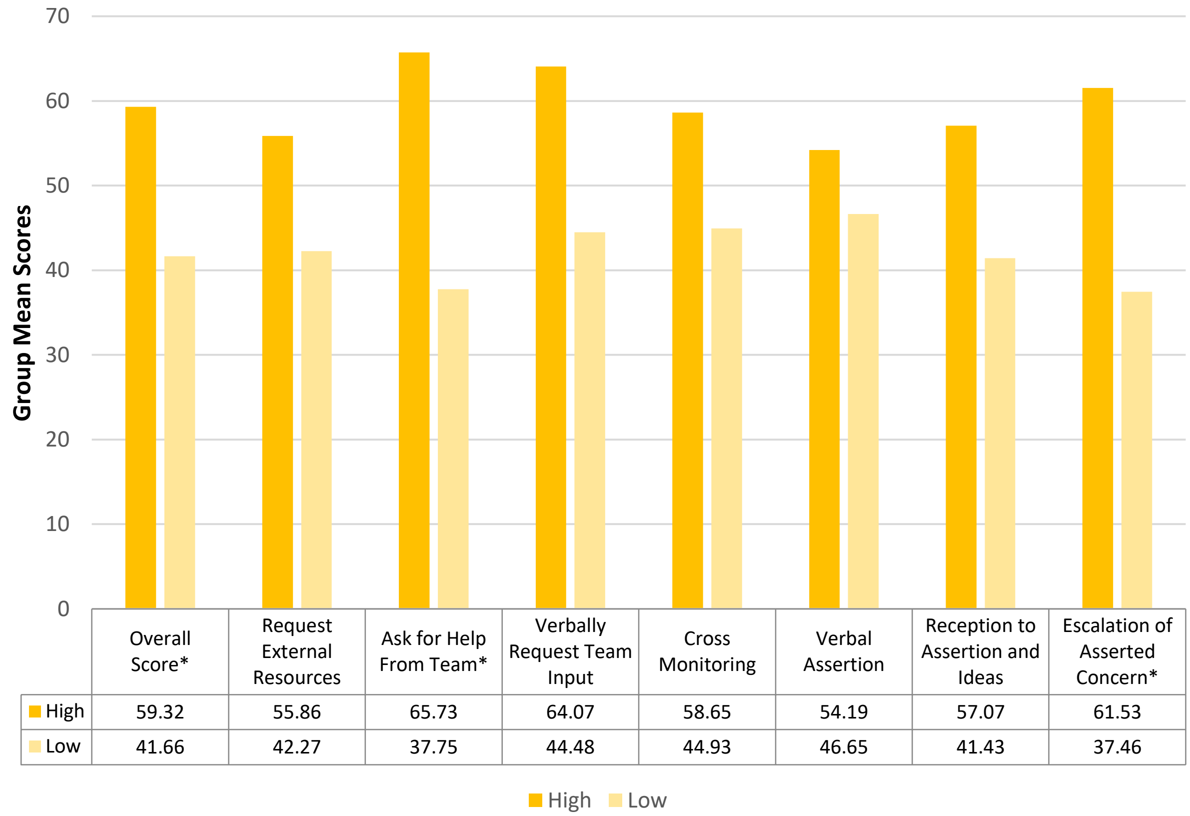
When looking at the situational awareness domain score, all of the behavior markers were significantly higher in the HF-IPE scenario as compared to the LF-IPE scenario (Figure C3).



*p < .05.

Figure C3. Situational awareness domain, behavior markers by group mean scores for high and low fidelity interprofessional simulation.

When looking at the cooperation domain score and the respective behaviour markers, only two out of the seven behaviour markers showed a statistically significant difference between the HF-IPE and LF-IPE scores (Figure C4). Despite this, the overall score for this domain showed a statistically significant difference with a $p = 0.007$. As stated previously, these numbers could be attributed to the small sample size used for this statistical analysis. If these results were from a collected data set, it could provide researchers with a good direction to focus subsequent research to further investigate if cooperation behaviours are displayed differently in HF-IPE as compared to LF-IPE.

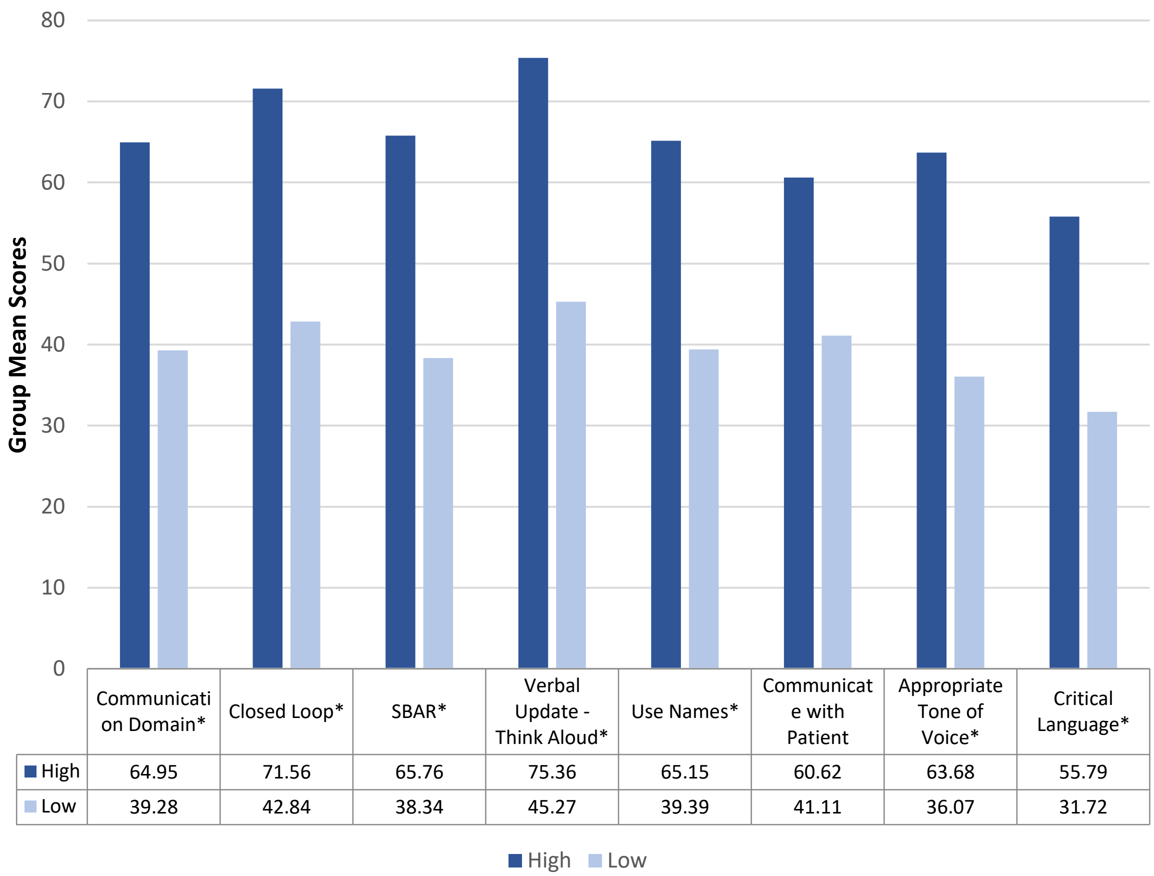


	Overall Score*	Request External Resources	Ask for Help From Team*	Verbally Request Team Input	Cross Monitoring	Verbal Assertion	Reception to Assertion and Ideas	Escalation of Asserted Concern*
High	59.32	55.86	65.73	64.07	58.65	54.19	57.07	61.53
Low	41.66	42.27	37.75	44.48	44.93	46.65	41.43	37.46

*$p < .05$.

Figure C4. Cooperation domain, behaviour markers by group mean scores for high and low fidelity interprofessional simulation.

When looking at the communication domain score and the respective behaviour markers, six out of seven behaviour markers showed a statistically significant difference between HF-IPE and LF-IPE scores (Figure C5).



	Communication Domain*	Closed Loop*	SBAR*	Verbal Update - Think Aloud*	Use Names*	Communicate with Patient	Appropriate Tone of Voice*	Critical Language*
High	64.95	71.56	65.76	75.36	65.15	60.62	63.68	55.79
Low	39.28	42.84	38.34	45.27	39.39	41.11	36.07	31.72

*$p < .05$.

Figure C5. Communication domain, behaviour markers by group mean scores for high and low fidelity interprofessional simulation.

Analysis of the fictitious data would conclude that the null hypothesis is rejected and the alternate hypothesis is accepted. The null hypothesis postulated that HF-IPE and LF-IPE would produce the same communication and teamwork scores on the CATS assessment tool. Rejecting the null hypothesis means that HF-IPE and LF-IPE did not produce the same results on the CATS assessment tool. The alternate hypothesis postulated that HF-IPE would produce higher group means as compared to the LF-IPE scores on the CATS assessment tool. Accepting the alternate hypothesis means that participation in HF-IPE produced higher group means using the

CATS assessment tool when compared to LF-IPE. If the "real" collected data set produces similar results to this fictitious data set, it would be clear that participation in the HF-IPE fosters higher quality and more communication and teamwork behaviors as compared to participation in the LF-IPE.

Conclusion

It is clear that the data analysis plan developed for this practicum project can be used to analyze and present the quantitative data collected using the CATS assessment tool. This data analysis report and proof of concept exercise demonstrated that the data analysis plan including the SPSS codebooks and statistical methodology chosen to analyze the quantitative data from the CATS assessment tool was appropriate and can produce the desired results.

References

Aliner, G., Harwood, C., Harwood, P., Montague, S., Huish, E., Ruparelia, K., . . . Antuofermo, M. (2014). Immersive clinical simulation in undergraduate health care interprofessional education: Knowledge and perceptions. *Clinical Simulation in Nursing, 10*(4), E205-E216. doi:10.1016/j.ecns.2013.12.006

Frankel, A., Gardner, R., Maynard, L., & Kelly, A. (2007). Using the Communication and Teamwork Skills (CATS) assessment to measure health care team performance. *Joint Commission Journal on Quality and Patient Safety*, *33*(9), 549-558. doi:10.1016/S1553-7250(07)33059-6

Garbee, D., Paige, J., Bonanno, L., Rusnak, V., Barrier, K., Kozmenko, L., . . . Nelson, T. (2012). Effectiveness of teamwork and communication education using an interprofessional high-fidelity human patient simulation critical care code. *Journal of Nursing Education and Practice, 3*(3), 1-12. doi:10.5430/jnep.v3n3p1

Garbee, D., Paige, J., Barrier, K, Kozmenko, V., Kozmenko, L., Zamjahn, J., . . . Cefalu, J. (2013). Interprofessional teamwork among students in simulated codes: A quasi-experimental study. *Nursing Education Perspective*, *34*(5), 339-344. doi:10.5480/1536-5026-34.5.339

Hughes, K. M., Benenson, R. S., Krichten, A. E., Clancy, K. D., Ryan, J. P., & Hammond, C. (2014). A crew resource management program tailored to trauma resuscitation improves team behavior and communication. *Journal of the American College of Surgeons*, *219*(3), 545-551. doi:10.1016/j.jamcollsurg.2014.03.049

Joshi, K., Hernandez, J., Martinez, J., AbdelFattah, K., & Gardner, A. K. (2017). Should they stay or should they go now? Exploring the impact of team familiarity on interprofessional team training outcomes. *The American Journal of Surgery*. 1-7. doi:10.1016/j.amjsurg.2017.08.048

Knapp, H. (2016). *Practical Statistics for Nursing Using SPSS.* Los Angeles, CA: SAGE Publications.

MacDonald, S., Manual, A., Dubrowski, A., Bandrauk, N., Law, R., & Curran, V. (2016). *Measuring the effectiveness of high fidelity simulation in interprofessional education to foster teamwork among undergraduate nursing, medicine and pharmacy students.* Teaching and Learning Proposal, Memorial University of Newfoundland.

Passauer-Baierl, S., Baschnegger, H., Bruns, C., & Weigl, M. (2014). Interdisciplinary teamwork in the OR: Identification and measurement of teamwork in the operating room. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen, 108*(5), 293-298. doi:10.1016/j.zefq.2013.05.007

Simpson, S. H. (2015). Creating a data analysis plan: What to consider when choosing statistics for a study. *The Canadian Journal of Hospital Pharmacy*, *68*(4), 311-317. Retrieved from https://www.cjhp-online.ca/

Smithburger, P., Kane-Gill, S., Kloet, M., Lohr, B., & Seybert, A. (2013). Advancing interprofessional education through the use of high fidelity human patient simulators. *Pharmacy Practice, 11*(2), 61-65. Retrieved from https://www.pharmacypractice.org/

Polit, D. F., & Beck, C. T. (2017). *Nursing research: Generating and assessing evidence for nursing practice* (10th ed.). Philadelphia, PA: Walters Kluwer/Lippincott Williams & Wilkins.

Appendix D

SPSS Codebooks

First Codebook: Used to Calculate Weighted Scores

In order to create a codebook that can statistically analyze the differences between HF-IPE and LF-IPE, an initial codebook had to be created to compute the raw data into a weighted total score. To calculate the weighted total scores, the raw data will have to be entered into the first codebook. Using the "compute variable" function within SPSS, the raw data would be calculated into non-weighted scores and weighted scores.

Name Column

Within "Variable View" the name column was used to delineate which variable to input. A list of the names of the variables used within the first codebook, along with the labels used and values assigned are presented in Table D1. The **"Group"** variable referred to the different interprofessional teams that took part, along with if they took part in HF-IPE or LF-IPE. These differentiations were labeled within the Values column (Table D1). The **"Behaviour"** variable referred to which behavior marker was being assessed. Labels within the Values column were used to identify the 21 behaviour markers, the four domains, and an overall score (Table D1). The **"GB"** variable, **"VGB"** variable, and **"NOB"** variable are the three levels of measurements noted for each behaviour marker within the CATS assessment tool. These three variables are where the raw data scores would be inputted into the codebook. **"A"** variable and **"B"** variable are calculated variables of interest within this first codebook. The raw scores for **"GB", "VQB"**, and **"NOB"** will be used to calculate the **"A"** scores, which is the non-weighted score. Once the **"A"** scores are determined, the **"B"** scores – known as weighted score – will be calculated for each value identified within the **"Behaviour"** variable. The reason these variables were named "**A"** and "**B**" was because it allowed for an easier process when using the "compute variable" function.

Table D1

Weight Totals Codebook Names, Labels, Values, and Measure

Name	Label	Values
Group		1.0 = Group 1 HF-IPE 2.0 = Group 1 LF-IPE 3.0 = Group 2 HF-IPE 4.0 = Group 2 LF-IPE *This trend will continue for all groups
Behaviour		1.0 = Briefing 2.0 = Verbalize Plan 3.0 = Verbalize Expected Outcomes 4.0 = Debriefing 5.0 = Visually Scan Environment 6.0 = Verbalize Adjustments in plan as changes occur 7.0 = Request external resources if needed 8.0 = Ask for help from team as needed 9.0 = Verbally request team input 10.0 = Cross Monitoring 11.0 = Verbal Assertion 12.0 = Receptive to assertion and ideas 13.0 = Closed Loop 14.0 = SBAR 15.0 = Verbal updates – think aloud 16.0 = Use Names 17.0 = Communicate with Patient 18.0 = Appropriate tone of voice

		19.0 = Establish event manager 20.0 = Escalation of asserted concerns 21.0 = Critical Language 22.0 = Coordination Domain 23.0 = Situational Awareness Domain 24.0 = Cooperation Domain 25.0 = Communication Domain 26.0 = Overall Score
GB	Good Behaviour	
VQB	Variable Quality Behaviour	
NOB	Not Observed but Expected Behaviour	
A	Non-Weighted Score	
B	Weighted Score	

Compute Variables

Below are the two equations used to determine the non-weighted scores and weighted scores under the "Compute Variable" function:

[GB+ VQB + NOB = A]

[((GB + (VQB * 0.5) + (NOB * 0)) / A) * 100 = B]

Type, Width, and Decimal Columns

All variables were set to the numeric type. The width of each value was consistent at eight and the decimal value was placed at two. Since the significance of the *p* value was set at 0.05 having the decimal value at two was sufficient given the context of this design.

Label Column

The variables **"GB"**, **"VQB"**, and **"NOB"** were respectively labeled as "Good Behaviour", "Variable Quality Behaviour", and "Not Observed but Expected Behaviour". Variables **"A"** and **"B"** were labeled as "non-weighted score" and "weighted score" respectively.

Missing Column

The value "88" under the "Discrete Missing Values", was assigned to represent data that was missing. The number "88" was chosen because it is unlikely that the raw data would produce such a unique number under any of the behaviour markers. "99" was also assigned under the "Discrete Missing Values" to represent data not collected because it might not be applicable given the context of the simulation. For example, three behaviour markers within the CATS assessment tool are only assessed if a crisis situation arises. By assigning these numbers for missing data it will ensure this information will be excluded from the data analysis process if applicable.

Measure Column

The data to be entered under the **"Group"** and **"Behaviour"** variables was considered nominal data. All other variables were considered scale, due to the numeric nature of the data to be coded.

Second Codebook: Used to Analyze HF-IPE versus LF-IPE

Once the weighted scores for each measurement was calculated using the first codebook, a second codebook needed to be created to organize the data into HF-IPE scores and LF-IPE scores. Once organized into these two distinguishable groups, the data could be analyzed to determine if there were any statistically significant differences between the means. The weighted scores were calculated within the first codebook and then manually transferred to their respective variables within the second codebook. It is important to note that the manual transfer of data could be a limitation of this data analysis plan, due to the risk of human transcription error. This could be addressed if one person read out loud the weighted scores from the first codebook and a second person cross-reference the scores with the second codebook, to ensure they were transferred correctly.

Name Column

For the second codebook, 52 variables were identified, 26 variables for HF-IPE and 26 variables for LF-IPE. These 26 variables include one overall group mean score, 21 behaviour marker means, and four domain mean scores. Naming of each variable follows the same pattern of **"HF_OS"** or **HF_CoorD"** with proper names being applied in the Labels column. Abbreviations of names were used to keep the names within this column short, which will hopefully allow for a more visually pleasing representation of the data when transferred to bars and graphs. Please refer to Table D2 for a full list of the name of each variable and their respective labels. It is important to note that no information was placed in the Values column in the second codebook.

Table D2

Names and Labels for Second Codebook

Name	Labels
HF_OS	HF IPE Overall Score
LF_OS	LF IPE Overall Score
HF_CoorD	HF IPE Coordination Domain Score
HF_SAD	HF IPE Situation Awareness Domain Score
HF_CoopD	HF IPE Cooperation Domain Score
HF_CommD	HF IPE Communication Domain Score
LF_CoorD	LF IPE Coordination Domain Score
LF_SAD	LF IPE Situation Awareness Domain Score
LF_CoopD	LF IPE Cooperation Domain Score
LF_CommD	LF IPE Communication Domain Score
HF_B	HF IPE Briefing Score
HF_VP	HF IPE Verbalize Plan Score
HF_VEO	HF IPE Verbalize Expected Outcomes Score
HF_DB	HF IPE Debriefing Score
HF_VSE	HF IPE Visually Scan Environment Score
HF_VAP	HF IPE Visually Adjustment in Plan as Changes Occur Score
HF_RER	HF IPE Request External Resources as Needed Score

HF_AFH	HF IPE Ask for Help From Team as Needed Score
HF_VRI	HF IPE Verbally Request Team Input Score
HF_CM	HF IPE Cross Monitoring Score
HF_VA	HF IPE Verbal Assertion Score
HF_RTA	HF IPE Receptive To Assertion and Ideas Score
HF_CL	HF IPE Closed Loop Score
HF_SBAR	HF IPE SBAR Score
HF_VUTA	HF IPE Verbal Updates Thinks Aloud Score
HF_UN	HF IPE Uses Names Score
HF_CWP	HF IPE Communicates with Patient Score
HF_ATOV	HF IPE Appropriate Tone of Voice Score
HF_EEM	HF IPE Establish Event Manager Score
HF_EAC	HF IPE Escalation of Asserted Concern Score
HF_CL	HF IPE Critical Language Score
LF_B	LF IPE Briefing Score
LF_VP	LF IPE Verbalize Plan Score
LF_VEO	LF IPE Verbalize Expected Outcomes Score
LF_DB	LF IPE Debriefing Score
LF_VSE	LF IPE Visually Scan Environment Score
LF_VAP	LF IPE Visually Adjustment in Plan as Changes Occur Score
LF_RER	LF IPE Request External Resources as Needed Score
LF_AFH	LF IPE Ask for Help From Team as Needed Score
LF_VRI	LF IPE Verbally Request Team Input Score

LF_CM	LF IPE Cross Monitoring Score
LF_VA	LF IPE Verbal Assertion Score
LF_RTA	LF IPE Receptive To Assertion and Ideas Score
LF_CL	LF IPE Closed Loop Score
LF_SBAR	LF IPE SBAR Score
LF_VUTA	LF IPE Verbal Updates Thinks Aloud Score
LF_UN	LF IPE Uses Names Score
LF_CWP	LF IPE Communicates with Patient Score
LF_ATOV	LF IPE Appropriate Tone of Voice Score
LF_EEM	LF IPE Establish Event Manager Score
LF_EAC	LF IPE Escalation of Asserted Concern Score
LF_CL	LF IPE Critical Language Score

Type, Width, and Decimal Columns

All variables are considered numeric due to the nature of the data. The width is set to eight and the decimal is set to two to maintain consistency among the different codebooks.

Missing Column

"88" and "99" were used again in the second codebook to delineate between data that is missing or variables that are not applicable given the context of the situation.

Measure Column

All variables are considered scale due to the numeric nature of the data to be coded and analyzed.

Appendix E

Decision Tree to Determine Statistical Analysis