A DATA ANALYSIS PLAN FOR THE INTERPROFESSIONAL TEAMWORK QUESTIONNAIRE TO MEASURE THE IMPACT OF HIGH FIDELITY INTERPROFESSIONAL EDUCATION

by ©Amy Warren

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

April 2018

St. John's Newfoundland and Labrador

Abstract

High-fidelity interprofessional education can help to create realistic patient scenarios for active student engagement however there is a need to measure whether it has an impact on interprofessional teamwork. The purpose of this practicum was to demonstrate advanced nursing competencies by developing a data analysis plan for the Interprofessional Teamwork Questionnaire. That questionnaire is designed to measure nursing, medicine, and pharmacy students' changes in knowledge and attitudes towards high fidelity interprofessional education. A literature review was conducted, and consultations were held with researchers and a statistician to inform the development of the data analysis plan. The data analysis plan includes methods for analyzing, summarizing, interpreting and displaying the quantitative data from the questionnaire. The data analysis plan considers measures of central tendency, consideration of normal distribution, measures of variability, and measures of symmetry to determine whether parametric or non-parametric measures are appropriate in data analysis. Two non-parametric measures that are appropriate for analysis of the quantitative data collected from the Interprofessional Teamwork Questionnaire include the Wilcoxon Matched-Pairs Signed Rank Test, and Friedman's two-way analysis of variance.

Key Words: data analysis plan, quantitative research, parametric, nonparametric, advanced nursing competencies, Interprofessional Teamwork Questionnaire

Acknowledgements

I would first like to express my profound gratitude to my practicum supervisor Professor Sandra MacDonald of the Nursing Faculty at Memorial University. She has been the compass guiding me through uncharted waters with all the patience and grace of a true mentor. She has been a role model for all that I someday hope to be, and a consistent positive force through this practicum.

I would also like to thank Joanne Smith-Young, the Research Coordinator for the Nursing Faculty at Memorial University. She has provided tremendous support in clarifying aspects of the data analysis plan with all the passion of someone who truly enjoys helping others. Her amicable nature made working with both the statistician and I, a positive experience.

Finally, I would also like to acknowledge my family for their unfailing support and continuous encouragement through the many tribulations faced in reaching this educational milestone. They have been my shelter from life's storms and the gasoline in my tank when I felt as though I was running on empty. I would not be who I am without them, and for that, I am eternally grateful.

Table of Contents

List of Tables

List of Figures

Introduction

Interprofessional education (IPE) is a collaborative approach to teaching and learning that fosters teamwork among students in health-related fields such as nursing, medicine and pharmacy (de Voest, Raguckas, Bambini, & Beel-Bates, 2013; Garbee, et al., 2013; Gough, Hellaby, Jones, & MacKinnon, 2012; Krueger, Ernstmeyer, & Kirking, 2017; Rossler, & Kimble, 2016). IPE encourages students to use their varied educational backgrounds to learn together as teams during their education programs. High-fidelity human patient simulation (HF) is one particularly useful teaching and learning approach for IPE that can help to create realistic patient scenarios for active student engagement (Kardong-Edgren, Adamson, & Fitzgerald, 2010; Paige, et al., 2014; Stewart, Kennedy, & Cuene-Grandidier, 2010). One of the challenges when measuring the impact of high fidelity interprofessional simulation education (HF-IPE) is the need for reliable and valid instruments and a clear data analysis plan for those instruments (Gough et al., 2012; Kardong-Edgren et al., 2010).

The Interprofessional Teamwork Questionnaire (ITQ), is the focus of this practicum and it consists of a 5 point Likert scale with an ordered set of discrete terms or statements from which participants are asked to choose the response that best describes their state or experience with the simulation. The ITQ extracts nominal and ordinal level data by use of a Likert scale, which includes seven statements that are rated by the participant on a five-point scale from “Strongly agree” to “Strongly disagree.” Those statements focus on areas of individual functionality in an interprofessional team, the significance of interprofessional collaboration in the simulation, comprehension of the role of each

profession involved, individual communication confidence in an interprofessional team, confidence in collaboration for care planning and satisfaction with the learning experience.

The data analysis plan for the ITQ focuses on the statistics relevant to a pretest, posttest, repeated measures research design. It is appropriate to consider both parametric and non-parametric measures when analyzing Likert scale data, but the majority of research studies using Likert scales to evaluate interprofessional teamwork used parametric methods of data analysis (Curran, Mugford, Law, & MacDonald, 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Liaw, Zhou, Lau, Siau & Chan, 2014; Lin et al., 2013; Paige et al., 2014; Reising, Carr, Shea, & King, 2011; Rossler & Kimble, 2016; Sigalet, Donnon, & Grant, 2012; Stewart et al., 2010; Wellmon, Lefebvre, & Ferry, 2017). The detailed data analysis plan for the ITQ includes a discussion of sample size, descriptive statistics, levels of significance, and recommended non-parametric and parametric measures.

Purpose of Project

The goal of this research practicum project was to develop advanced nursing research skills through participating in the data analysis phase of a research study. The data analysis plan developed for this practicum will be used by Dr. Sandra MacDonald in her research study titled "Measuring the Effectiveness of High Fidelity Simulation in Interprofessional Education to Foster Teamwork Among Undergraduate Nursing, Medicine and Pharmacy Students" to analyze the quantitative data collected from the Interprofessional Teamwork Questionnaire (ITQ). The ITQ is being used in that study as the pretest, posttest instrument to measure the impact of participation in HF-IPE on knowledge and attitudes towards teamwork.

The objectives that guided the achievement of the practicum goal included:

1. Demonstrate advanced nursing practice competencies through research, leadership, clinical, and collaborative activities.
2. Analyze and synthesize existing literature to provide evidence for conclusions that will inform the development of the data analysis plan.
3. Apply evidence to create and implement a plan for quantitative data analysis.
4. Apply knowledge of data analysis, interpretation of results and nursing research methods when analyzing quantitative data.
5. Disseminate research findings by participating in knowledge – transfer techniques.

Methods

Three main methods were used to achieve the objectives of this practicum including an integrative literature review, consultations with nurse researchers and statisticians, and the development of a data analysis plan including an SPSS data analysis program for quantitative data analysis of the ITQ. A summary of the results from those methods will be discussed in this report. The complete literature review is presented in Appendix A and the report on the consultations in Appendix C. The proposed data analysis plan was developed based on the literature review and the consultations.

Summary of Literature Review

Search Methods

MUN University Libraries search service, CINHAL, PubMed, and Google Scholar databases were utilized in the identification of relevant articles for this literature review. The following search terms were used: Likert scale, data analysis plan, interprofessional teamwork questionnaire, quantitative questionnaire, simulation measurement,

interprofessional, teamwork, high fidelity simulation, simulated scenarios, interprofessional education, undergraduate medicine, pharmacy, and nursing students. This search generated approximately 30 pertinent abstracts and included searching the references of relevant articles.

The criteria used to screen the abstracts included: (a) the article related to high fidelity simulation and undergraduate IPE, (b) the article was a research study, (c) the article included quantitative evaluation measures and data analysis plans with Likert scales, and (d) the article included nursing students and at least one other health-related student group. For the purposes of this literature review, Zou, Carlsson, and Quinn's (2010) definition of the Likert scale as an ordered set of discrete terms or statements from which participants are asked to choose the response that best describes their state or experience, was used for article selection.

Likert Scale Data Analysis Plans

Data analysis plans help to identify the specific methods to be used when analyzing data collected in a research study. A plan is used to organize and guide statistical analysis to ensure aspects like sample sizes, methods of data collection, and instruments are appropriate for reaching the goal of the study. The process of creating a data analysis plan involves defining variables, especially in terms of designating independent and dependent variables. It also involves specifying the levels of these variables: nominal, ordinal, interval, and ratio (Simpson, 2015). This is particularly important for creating an effective SPSS file for data input. Additionally, data analysis plans involve outlining the descriptive statistics, which is particularly important for determining whether parametric, nonparametric, or a mixture of both are appropriate in data analysis (Simpson, 2015).

Furthermore, data analysis plans involve consideration of the research questions, the design of the study, the level of measurement, the level of significance, and the most common types of tests used for the types of data identified. This is particularly important if inferential statistics with hypothesis testing is to be performed (Simpson, 2015).

Likert Scale research instruments with accompanying data analysis plans were explored. Of the 12 studies analyzed, 9 data analysis plans utilized parametric measures (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Sigalet et al., 2012; Stewart et al., 2010); three utilized nonparametric measures (Liaw et al., 2014; Rossler & Kimble, 2016; Wellmon et al., 2017) and six utilized ANOVA (Curran et al., 2005; Krueger et al., 2017; Lin et al., 2013; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017). Six utilized paired *t*-tests (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Paige et al., 2014; Sigalet et al., 2012; Wellmon et al., 2017).

Although Kardong-Edgren et al., (2010) reported there was a paucity of reliability and validity data in their review of evaluation instruments for high fidelity simulation, the majority of these studies did address some aspect of validity and reliability of the instruments (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Sigalet et al., 2012), four studies included only the reliability of the instruments (Paige et al., 2014; Rossler & Kimble, 2016; Stewart et al., 2010; Wellmon et al., 2017), and one study discussed only the validity of the instruments (de Voest et al., 2013). Of the six studies that referred to the validity of the measurement instruments, three reported the validity of the existing measures (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014), two calculated the validity of the purpose designed instruments (Lin et al., 2013;

Sigalet et al., 2012), and one referred to the content validity analysis from a panel of experts (de Voest et al., 2013). Only seven of the research studies established both internal consistency and reliability of the instrument using Cronbach's α (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010).

Of the 12 studies included in this review, only one (Stewart et al., 2010) included a confidence interval (CI) among the provided data. However, that study did not state a level of statistical significance. All other studies stated a level of significance of 0.05, or declared data significant that had a *p*-value of 0.05 or below. Ten of the studies chose a significance level of 0.05 (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Sigalet et al., 2012; Wellmon et al., 2017). One study reported confidence intervals, with that confidence interval being set to 95% (Stewart et al., 2010). This evidence has been applied to the development of the quantitative data analysis plan for the ITQ, which includes the creation of the SPSS data analysis program for the ITQ.

Data analysis plans for Likert scales included the statistical measures of Tukey's honestly significant difference (HSD) (Lin et al., 2013; Wellmon et al., 2017); Cohen's *d* and component analysis with varimax rotation (Sigalet et al., 2012); and analysis of covariance (ACOVA) (Liaw et al., 2014). Three articles utilized nonparametric methods in the data analysis plan including the Mann-Whitney *U* test (Rossler & Kimble, 2016; Wellmon et al., 2017), chi-square test (Liaw et al., 2014; Rossler & Kimble, 2016), as well as the Wilcoxon signed rank and a Kruskal-Wallis test (Rossler & Kimble, 2016). These studies show there are several appropriate statistical tests that could be used in the data

analysis plan for the ITQ data including, but not limited to; Cronbach's α, t-test, paired t-tests, ANOVA, HSD, Cohen's *d*, ACOVA, Mann-Whitney *U* test, Wilcoxon signed rank, and a Kruskal-Wallis test.

Summary of Data Analysis Plan

The data analysis plan developed for this practicum focuses on analysis of ordinal data from the ITQ. An SPSS data analysis file was created to calculate the descriptive statistics including means, mode, median, frequencies, and normality assessment. Normalcy of the ITQ data can be determined using Pearson's Coefficient to help determine whether the data is normally distributed and whether it is possible to proceed with parametric or non-parametric measures. If the data from the ITQ proves to be too skewed for parametric measures, the nonparametric equivalents of the Friedman Matched Samples Test for the repeated measures analysis of variance and the Wilcoxon Matched-Pairs Signed Rank Test could be performed. A Bonferroni correction could also be performed (Paige et al., 2014). Two of the limitations noted in the proposed data analysis plan is that it does not address the reliability and validity of the ITQ, and the sample size that will be needed to determine statistical significance. Further consultation with a statistician is recommended to determine appropriate methods for calculating the required sample size for the study as well as calculating the reliability and validity of the instrument.

Summary of Consultations

Meetings were conducted via teleconferences, emails and telephone conversations with a researcher and a statistician. The purpose of those consultation meetings was to obtain feedback on the SPSS data analysis file for the ITQ. Based on the recommendations from the Nurse Researcher, the SPSS file was reorganized to cluster the pre and posttest

values for each item, as well as designate 0 for “not applicable” responses and 99 for missing data. Initially the recommendations from the statistician focused on the descriptive statistics for analyzing the Likert scale data. However, after further review of the literature and discussion with the statistician, it was determined that both descriptive and inferential statistics could potentially be applied to the analysis of the data from the ITQ.

Consultations with the statistician helped to determine that the ITQ Likert scale data could be considered ordinal or nominal and the Wilcoxon Signed-Rank test was determined to be appropriate for the data analysis. The statistician also provided valuable input into how to analyze the Likert scale data and that discussion essentially formed the basis of the selection of the statistical tests that were recommended in the data analysis plan. This was the most valuable output received from the consultation meetings and greatly clarified which tests would likely be most appropriate. Both the Nurse Researcher and the Statistician reviewed the data analysis plan and SPSS file for the ITQ and agreed there were no concerns and the data analysis plan could be used to analyze data from the ITQ.

Advanced Practice Competencies

This practicum project helped to develop advanced practice competencies in the areas of clinical practice, research, consultation and collaboration, and leadership (Canadian Nurses Association, 2008). The following is a discussion of examples of activities that demonstrate each competency.

Clinical Competency

Advanced clinical competencies were demonstrated during this practicum project by identifying and assessing trends in nursing research, specifically related to developing data analysis plans for Likert scales and integrating HF-IPE into undergraduate education.

Through analysis of the evidence collected for this project, it is clear that there is a lack of research on the impact of HF-IPE as a teaching learning approach to foster teamwork with undergraduate nursing, medicine and pharmacy students. It is also clear that there is a lack of valid and reliable psychometric instruments to evaluate the impact of HF-IPE on student's knowledge of teamwork and attitudes towards HF-IPE. Completing this practicum has contributed to my understanding of a trend in research and education that could delay adoption of HF-IPE into health education programs.

This practicum has also involved exploring the use of data from multiple sources, often in ambiguous and complex situations. There was a degree of ambiguity as to whether Likert scale data was nominal or ordinal in nature, requiring the comparison of several different articles, meetings with the researchers, and the input of a statistician to clarify the quandary. The same sort of complexity and ambiguity existed regarding the choice of statistical tests. Even though parametric measures dominated the landscape of the literature review, very few actually reported any assessment of normality before progressing to using parametric measures. Furthermore, these situations provided an opportunity to demonstrate decision-making in complex clinical situations. Though the choice of tests and data types may not affect direct clinical care, it is an important decision for clinical research.

Through the discussion of the mock results, the clinical competency regarding the explanation of client responses was also partially obtained during this practicum. In the discussion section of this paper some of the mock phenomena are explored. Furthermore, potential explanations for these phenomena are provided from the literature. For example, since mock nursing students had such low pretest scores, the literature was analyzed to determine whether any of the studies from the literature review shared a similar trend. One

research study by Sigalet et al., (2012) connected low pretest scores in nursing students to a lack of exposure to interprofessional collaboration concepts. Even the mock results could be linked to the literature therefore proving the data analysis plan could effectively analyze the data from the ITQ.

Although the data analysis plan was based on fictitious data, it could contribute to the generation of new nursing knowledge for practice by providing clear guidelines for nurse researchers conducting program evaluation research. This practicum provided an opportunity to add to the growing number of research practicums by Master of Nursing students. It also added to the body of knowledge needed for nurse education when pursuing the evaluation of teaching and learning methodologies in undergraduate program.

Research

This practicum provided an opportunity to participate as a collaborator on the research team by creating a data analysis plan, analyzing data and compiling a report. Although the data analyzed for this project was not from real students, the creation of the SPSS data analysis file and the analysis of mock data were completed as if it were collected from the study. These are examples of the advanced research competencies that were performed in the area of data analysis. Furthermore, advanced research competencies were demonstrated by analyzing the trends noted above regarding delayed adoption of interprofessional education into health education programs.

Under the research competency "…critique, interpret, apply, and disseminate evidence-based findings" the integrated literature review demonstrated the advanced research competencies of critiquing the evidence. The analysis of the findings from the mock data of the ITQ demonstrated an ability to interpret statistical tests and whether

results are significant when given significance levels. The development of the data analysis plan demonstrated an awareness of the limitations of parametric and nonparametric statistical tests along with the assumptions required of each measure. Analyzing the mock data from the ITQ also demonstrated a beginning research competency in the ability to interpret descriptive and inferential statistics. Furthermore, the results of this practicum will be disseminated through the practicum presentation and will be applied in Dr. MacDonald's research study.

Consultation and Collaboration

The advanced practice competency of consultation and collaboration was demonstrated by initiating timely and appropriate consultations with nurse researchers and the statistician. This competency was demonstrated using teleconferences, email, and telephone meetings to engage in collaboration and consultations with appropriate experts in the field. This practicum also involved practicing collaboratively to build effective coalitions and demonstrating knowledge in communication. Interpersonal relations are extremely important for building effective partnerships. When meetings were canceled and rescheduled several times, it did provide a challenge to meeting this competency. However, communication was always appreciative and assertive of needs, leading to no unnecessary conflict. The relationship built with both the research coordinator and the practicum supervisor was collegial, also providing evidence for some attainment of this competency.

Applying group dynamics is also a relevant competency in this area. A student must always remember that they are only one of many interests competing for the limited time of experts in the organization. There are likely many other interests that are much more pressing than the problem to be presented by the student. Therefore, it is imperative that the

dynamics of the organization be considered and to approach those situations with patience. Like the other competency, this was best illustrated by the recognition of this dynamic when challenges, like cancelled meetings, arose.

Leadership

This competency was demonstrated by taking the initiative to develop a research practicum when it was not a common topic for the MN practicums. This is the first MN practicum that has focused on the development a data analysis plan, which makes this is an example of the leadership competency expected in advanced nursing practice. Leadership was also demonstrated though identifying problems and initiating change. A problem occurred when the statistician made the suggestion that only descriptive statistics could be utilized for the data due to the proposed nominal nature of Likert items. However, through a more thorough review of the literature and the support of other experts like the research coordinator, nonparametric measures were still deemed possible. Therefore, the initiated change was the compromise of completing descriptive statistics as suggested by the statistician but also completing nonparametric statistics.

Proof of Concept for the Data Analysis Plan

The proposed data analysis plan for the ITQ was tested using "mock" data representing 12 subjects in a repeated measures pretest posttest research design. The mock data analysis includes descriptive statistics, inferential statistics, analysis of variance and visual presentation of the findings. Although the data presented here is fictitious, it serves as proof of concept that these statistical tests can be applied successfully to ITQ data. The limitations of this "mock" data analysis include the lack of established reliability and validity of the ITQ, the lack of a representative sample and the small "mock" sample size.

Following is a discussion of the mock data analysis with implications for the ITQ data analysis plan developed for this practicum.

When normality on the mock data was assessed using measures of symmetry, it was found that the data for all items were significantly skewed. For example, the skewness value calculated for item eight, regarding whether the objectives were clear, from the post LFS ITQ was 3.464. The kurtosis value for same item was 12.00. The standard error for the skewness and kurtosis of this item was 0.637 and 1.232 respectively. This would indicate a very large skewness and kurtosis. This is not surprising given the particularly small sample size utilized for data analysis (Munro, 2005). Therefore, the inferential statistics performed included the nonparametric measure of the Wilcoxon Matched Pairs Signed Rank Tests to address the small sample size and significantly skewed distributions. Even though the mock sample size was small, significant results were calculated which is particularly intriguing given the fact that sample sizes should be large enough to detect significant differences (Munro, 2005). The following is a brief discussion of the mock data findings and a discussion of those mock findings.

Percentages of Students' Responses

One of the methods of data analysis for Likert Scales includes percentages, which are often perceived as more meaningful and easier to understand than other methods (Munro, 2005). When comparing the combined total groups' percentages of mock student responses, several trends emerged (Table 1 and Table 2). Prior to participating in simulation, 66.6% reported their ability to function effectively as a team member as "Neutral" (33.3%) to "Agree" (33.3%). After LFS, 58% reported this as "Neutral."

Table 1

Percentage Breakdown of ITQ Responses By Profession

Statement	Nursing					Medicine					Pharmacy				
	SD	D	N	A	SA	SD	D	N	A	SA	SD	D	N	A	SA
Pretest 1		50	25	25			25	25	25	25			50	50	
Post-LFS 1		50	25	25				50	50				100		
Post-HFS				50	50				75	25			25	25	50
Pretest 2			25	75				25	75				50	50	
Post-LFS 2		50	50				25	50	25			50	50		
Post-HFS			25	25	50				50	50				50	50
Pretest 3		25	75				50	25	25			50	50		
Post-LFS 3		50	50				25	50	25			25	50	25	
Post-HFS 3			25	50	25			25	50	25			50	50	
Pretest 4		50	50				50	25	25			25	50	25	
Post-LFS 4		25	50	25			50	50				50	25	25	
Post-HFS 4			25	75				25	50	25			25	75	
Pretest 5		25	50	25			50	25	25			50	50		
Post-LFS 5		25	75					50	50			25	50	25	
Post-HFS 5			25	50	25				100					50	50
Pretest 6		25	50	25			25	25	25	25		25	50	25	
Post-LFS		25	50	25			25	50	25			25	50	25	
Post-HFS 6			50	25	25			25	50	25				75	25
Pretest 7		25	50	25			25	50	25			50	50		
Post-LFS		50	50				25	50	25			25	50	25	
Post-HFS 7				75	25			25	50	25				50	50
Post-LFS 8				100					75	25				100	
Post-HFS 8				75	25				25	75				75	25
Post-LFS 9				100					100					100	
Post-HFS 9					100				100					25	75
Post-LFS 10				100					100					100	
Post-HFS 10				75	25			25	25	50			25	75	
Post-LFS 11			25	75					75	25			25	50	25
Post-HFS 11				100						100				50	50
Post-HFS 12				100					25	75				100	
Post-LFS 13 -				75	25			25	75					75	25
Post-HFS 13				100					50	50			25	50	25
Post-LFS 14				100					100					100	
Post-HFS14					100					100					100
Post-LFS 15				100					75	25				100	
Post-HFS 15				25	75				75	25				50	50

Conversely, after HFS, 91.7% reported the same item as “Agree” (50%) to “Strongly Agree” (41.7%). This difference could indicate that the mock students perceived they could function more efficiently as a team after HFS as compared to LFS.

Table 2

Breakdown of Percentage of Total Group Responses

	PreTest					Post-LFS					Post-HFS				
	SD	D	N	A	SA	SD	D	N	A	SA	SD	D	N	A	SA
Statement	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%
Function as a member of the team		25	33.3	33.3	8.3		16.7	58.3	25				8.3	50	41.7
Understand the importance of teamwork			33.3	66.7			41.7	50	8.3				8.3	41.7	50
Understand the role of nursing		41.7	50	8.3			33.3	50	16.7				33.3	50	16.7
Understand the role of medicine		41.7	41.7	16.7			41.7	41.7	16.7				25	66.7	8.3
Understand the role of pharmacy		41.7	41.7	16.7			16.7	58.3	25				8.3	66.7	25
Confidently communicate		25	41.7	25	8.3		25	50	25				25	50	25
Confidently collaborate		33.3	50	16.7			33.3	50	16.7				8.3	58.3	33.3
Learning objectives clear.	-	-	-	-	-				91.7	8.3				58.3	41.7
Workload was fair.	-	-	-	-	-				100					41.7	58.3
Experience was organized	-	-	-	-	-				100				16.7	58.3	25
Pre-Briefing was useful	-	-	-	-	-			16.7	66.7	16.7				50	50
Orientation was useful	-	-	-	-	-	-	-	-	-	-				75	25
Debriefing was useful	-	-	-	-	-			8.3	75	16.7			8.3	66.7	25
I would recommend this experience to others.	-	-	-	-	-				100						100
Overall a meaningful experience.	-	-	-	-	-				100						100

Before participating in simulation, these mock participants appeared to understand the importance of interprofessional teamwork with 66.7% of the total group choosing “Agree” and 33.3% choosing “Neutral.” Comparatively, post LFS these values dropped with 41.7% choosing “disagree” and 50% now choosing “Neutral.” On the other hand, after HFS these values shifted upward with 50% choosing “Strongly Agree” and 41.7%

choosing “Agree.” This difference could indicate the possibility that LFS may have negatively impacted the students understanding of the importance of interprofessional teamwork.

The pretest results of these mock participants indicated that they reported a relatively low understanding of the role of nursing on the interprofessional team with 50% responding with “Neutral” to this item and 41.7% responding with “Disagree.” These values showed very little change post participation in LFS. However, post HFS, 50% responded with “Agree” and only 33.3% responded with “Neutral” showing a positive improvement in this ITQ item. A similar trend held true for understanding the role of medicine and pharmacy on the interprofessional team.

With regards to the mock respondents’ confidence in communicating effectively with the interprofessional team, pretest results show relatively nonaligned confidence levels with 41.7% choosing “Neutral.” Compared to HFS, LFS showed relatively little change in this item with 50% choosing “Neutral.” However, post HFS 50% of mock respondent answered with “Agree,” indicating a positive improvement in this domain. A similar trend held true for their confidence in collaborating effectively with the interprofessional team with post LFS scores showing literally no change from pretest scores and a clear improvement post HFS.

Though the mock respondents did show clear differences in the above domains, their overall evaluation of the simulation experience was similar regardless of whether it was HFS or LFS. Most of the mock respondents chose “Agree” for all items regarding the objectives, organization, pre-briefing, orientation to HFS, and debriefing. However, there was some positive differences regarding the workload, the meaningfulness of the

experience, and whether they would recommend the experience to others. Regarding the workload, post LFS 100% of students agreed that the workload was fair. However, post HFS 58.3% of students chose “Strongly Agree” and 41.7% chose “Agree.” This indicates a greater perception of fairness regarding HFS versus LFS.

Even though 100% of the students chose “Agree” to whether they would recommend the LFS to others, post HFS 100% chose “Strongly Agree” indicating the potential for higher levels of satisfaction with HFS compared to LFS. A similar, but less intense, trend was seen regarding meaningfulness. Post LFS, 91.7% of students chose “Agree.” Conversely, post HFS students were split between “Agree” (50%) and “Strongly Agree” (50%).

There was a clear shift upward in percentages for HFS in these areas and often a shift downward or no change at all for LFS. Analysis of mock data showed that participation in HFS resulted in a higher level of agreement with the statement that they could function effectively as a member of a team, and collaboration and communication as compared to participation in LFS. The percentages, as well as the mean scores support the assertion that students were more confident in their ability to communicate with the interprofessional team after participating in HFS. A similar trend was true regarding their perception of their ability to collaborate effectively with the team.

Analysis of the breakdown of mean scores by profession revealed percentages for specific professions were different in several areas. The mock-nursing students rated their pretest level of functioning effectively as a team member lower than both medical and pharmacy students. These nursing students largely responded with “Disagree” to this item (50%), whereas 100% of pharmacy students rated themselves as “Neutral” (50%) to

"Agree" (50%). Medical students, on the other hand, had responses spread evenly across "Disagree" (25%), "Neutral" (25%), "Agree" (25%), and "Strongly Agree" (25%). Medical students also rated the same item as higher after LFS than their nursing and pharmacy counterparts with 50% of them rating it as "Neutral" and 50% rating it as "Agree." Conversely, nursing students showed no difference and 100% of pharmacy students rated the item as "Neutral" post LFS indicating a decrease from pretest values.

Though all mock student groups showed a positive shift upward in their ratings post HFS, nursing and pharmacy students were more likely than medical students to choose "Strongly Agree." (Figure 1) With the statement that they performed effectively as a member of the team for both, nursing and pharmacy students, 50% of them chose "Strongly Agree" post HFS whereas only 25% of medical students chose the same item. This is particularly interesting for the nursing students who disagreed with the statement and rated LFS comparatively poorly in their post LFS evaluation. This indicates the possibility that nursing students in particular may perceive themselves as functioning more effectively as a team member of the interprofessional team after completing HFS and may gain more from HFS.

When it comes to understanding the importance of interprofessional teamwork, both mock nursing and medical students had similar pretest results with 75% of them responding with "Agree" to this item (Figure 2). However, pharmacy students were more split with 50% of them responding with "Neutral" to this pretest item and 50% responding with "Agree." Though all student groups showed a negative shift in this item post LFS, nursing and pharmacy students showed a greater negative shift with 50% answering that

item with “Disagree.” None of the pharmacy and nursing students rated it as “Agree” or “Strongly Agree” post LFS, however 25% of medical students still rated it as “Agree.”

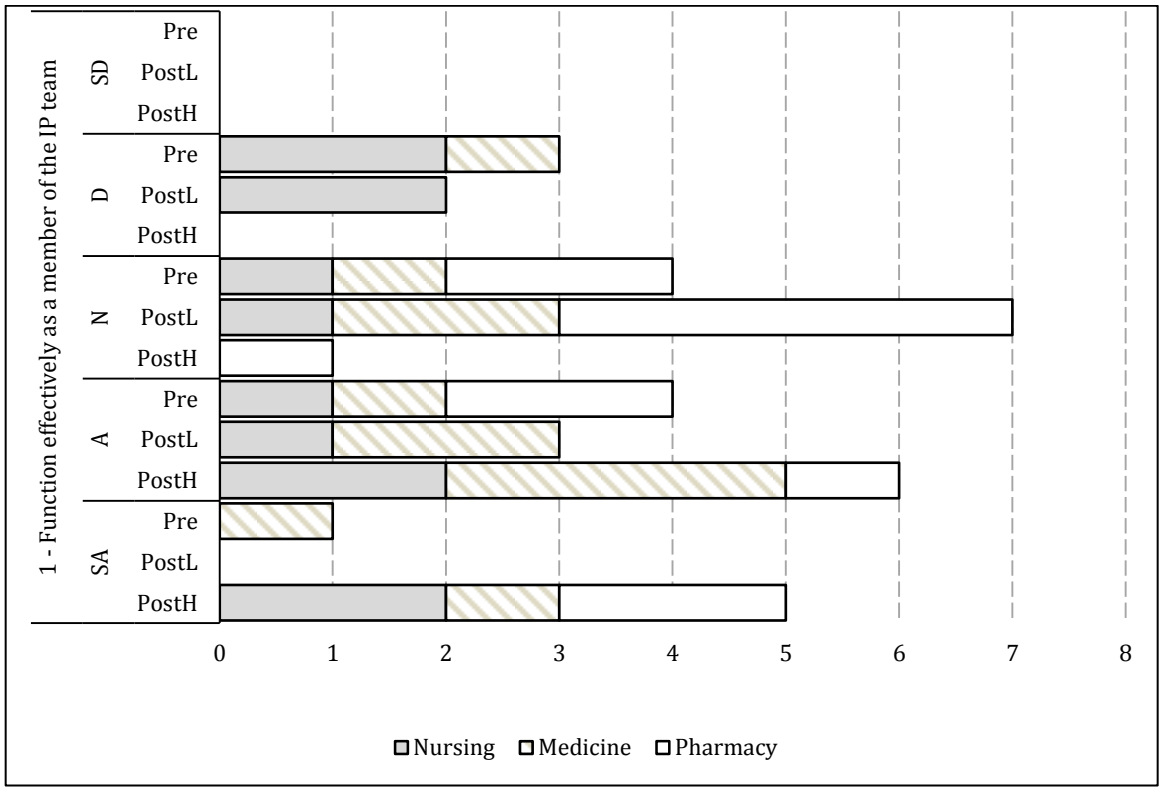

Figure 1. Function Effectively as a Member of the Team

After HFS, it appeared that the mock medical and pharmacy students rated their understanding of the importance of interprofessional teamwork as higher than the nursing students did. Even though 50% of all three student types chose “Strongly Agree” post HFS simulation, 25% of nursing students still remained “Neutral” on this item. This is unlike the medical and pharmacy students who chose only “Agree” (50%) or “Strongly Agree” (50%) for this item.

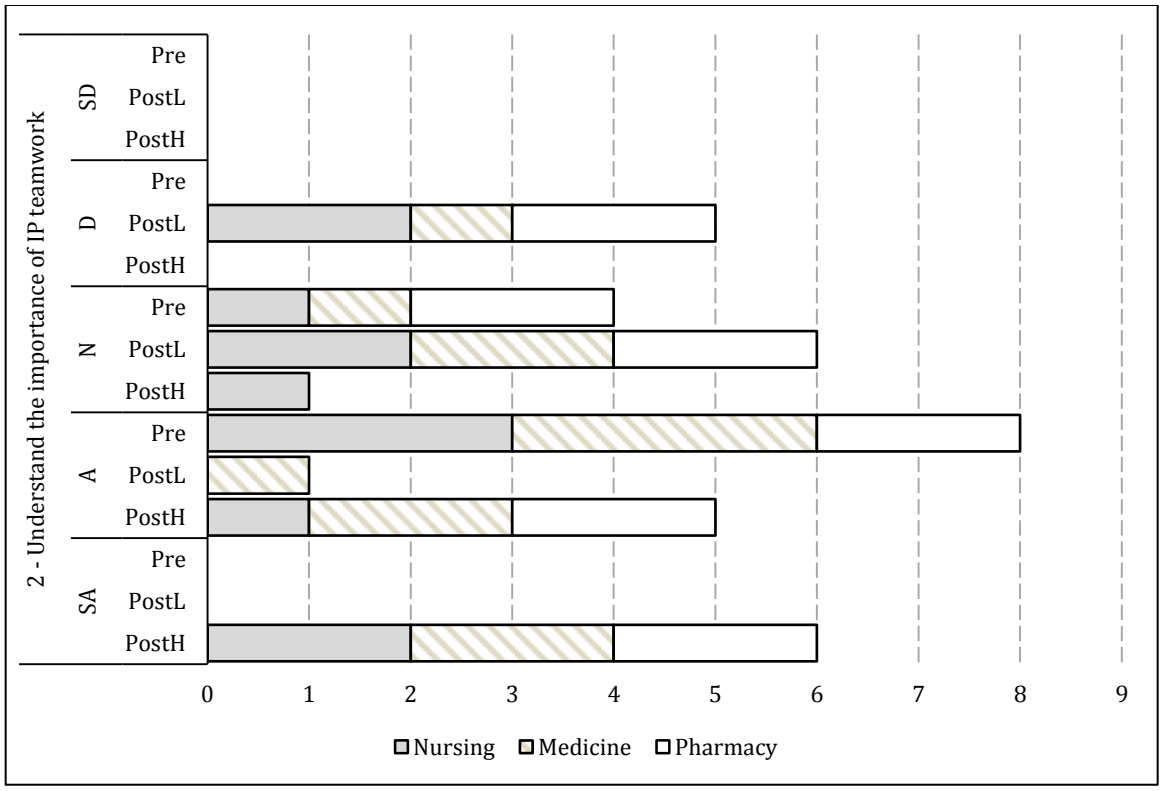

Figure 2. Importance of Interprofessional Teamwork

When it comes to understanding the role of the nurse on the interprofessional team, 25% of the mock medical students rated their understanding higher than nursing students themselves did in the pretest period (Figure 3). Nursing students rated this item largely as "Neutral" (75%) to "Disagree" (25%) whereas medical students were more spread with 50% responding with "Disagree", 25% responding with "Neutral," and 25% responding with "Agree." Pharmacy students were evenly split between "Neutral" (50%) and "Disagree." (50%). LFS appeared to have a more negative effect on nursing students in this ITQ item compared with the other two student groups. Post LFS, 50% of nursing chose "Disagree." An opposite shift occurred in medical students with 50% of them choosing "Neutral" and 25% choosing "Agree." Pharmacy students appeared the most optimistic with 50% choosing "Agree" post LFS.

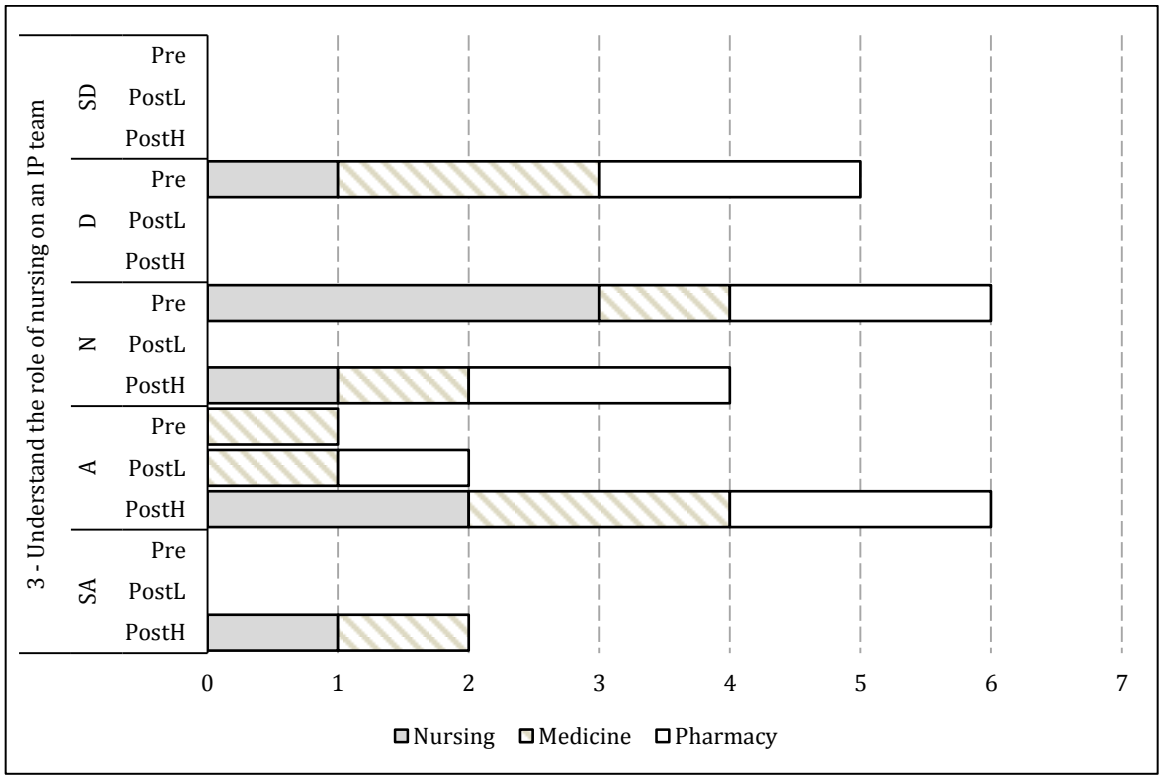

Figure 3. The Role of the Nurse

After HFS, both the mock medical and nursing students were more likely to rate their understanding of the role of nursing on an interprofessional team as "Strongly Agree" with 25% of these student groups choosing that item. However, none of the pharmacy students chose "Strongly Agree." Instead, 50% remained "Neutral" and 50% "Agreed." Only 25% of the medical and nursing students remained "Neutral" on this ITQ item comparatively post HFS.

When it comes to understanding the role of medicine on the interprofessional team, the mock nursing students rated their pretest understanding as lower than both medical and pharmacy student did (Figure 4). For the nursing students, their pretest results were spread evenly between "Disagree" (50%) and "Neutral" (50%). Comparatively, 50% of medical

students rated this item as “Disagree” with the remaining percentages spread evenly between “Neutral” (25%) and “Agree” (25%). Interestingly, pharmacy students were more nonaligned on the subject than medical students with 50% of them choosing “Neutral” for this item in the pretest period and the remaining percentages spread between “Disagree” (25%) and “Agree” (25%).

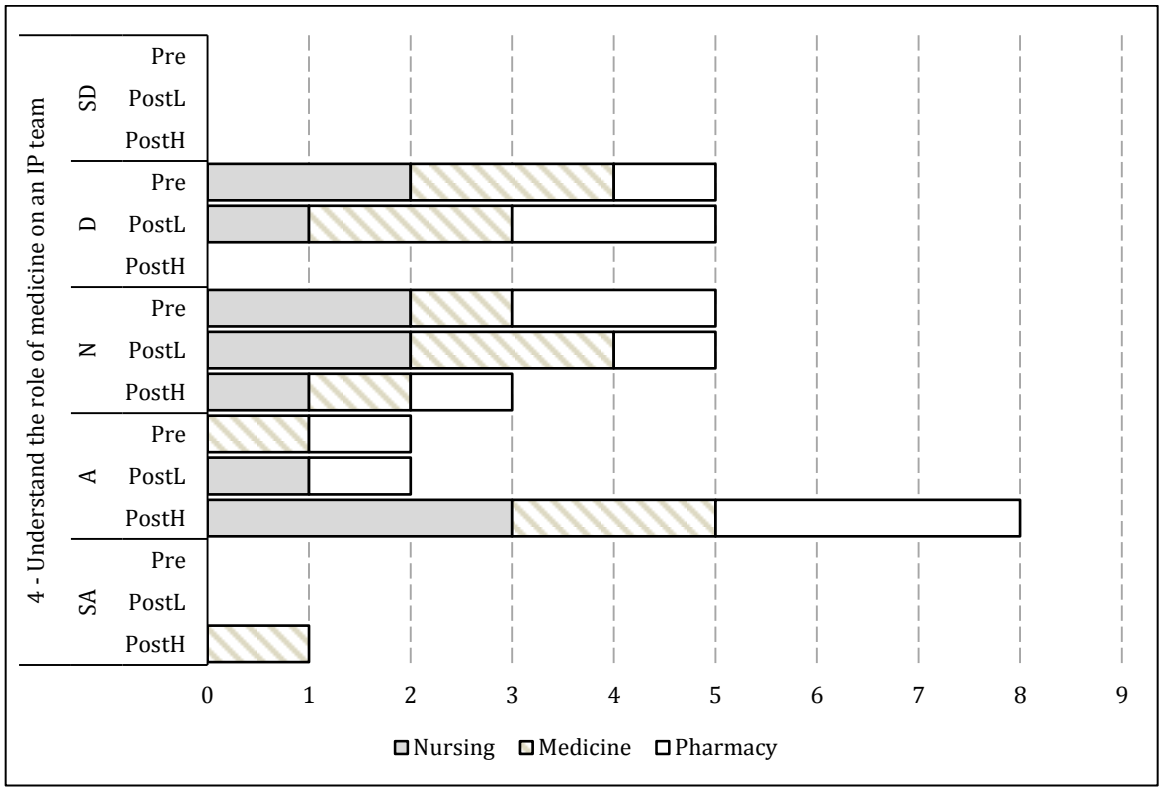

Figure 4. The Role of Medicine

After LFS, the mock-nursing students showed a small improvement in their rating of their understanding of the role of medicine. However, this improvement was not present for medical students and the pharmacy students. Post LFS, 25% of nursing students now rated this item as “Agree” where before they leaned toward neutral and disagreement. Post LFS, none of the medical students “Agreed” with this item with 50% choosing “Neutral”

and 50% choosing “Disagree.” The pharmacy students who were previously largely “Neutral” now leaned toward “Disagree” (50%) with this item. However, 25% of these pharmacy students still chose “Agree” with the statement, unlike the medical students.

Post HFS, both the mock nursing and pharmacy students largely chose “Agree” (75%) on this item regarding the role of medicine. Only 25% of nursing, pharmacy, and medical students remained neutral on these items after LFS. The medical students appeared more optimistic in their choices for this item with 50% choosing “Agree” and 25% choosing “Strongly Agree” whereas none of the other student types chose that option.

When it comes to understanding the role of pharmacy on the interprofessional team, the mock-nursing students rated this item the highest in the pretest period compared to medical and pharmacy students themselves (Figure 5). For the nursing students, this item was rated largely as “Neutral” (50%) to “Agree” (25%). Though 25% of medical students also chose “Agree,” 50% chose “Disagree” therefore dampening that result. Pharmacy students were the most conservative in their choice with responses spread evenly between “Disagree” (50%) to “Neutral” (50%).

Interestingly, LFS appeared to have a stronger negative impact on the mock-nursing students compared to the other student types. None of these students chose “Agree” or “Strongly Agree” post LFS, with 75% instead choosing “Neutral” and the remaining choosing “Disagree.” Unlike the nursing students, the pharmacy and medical students did show improvements in this area. For the medical students, responses were evenly spread between “Neutral” (50%) to “Agree” (50%) where in the pretest period only 25% chose “Agree” and half chose “Disagree” (50%). For the pharmacy students, 50% remained

"Neutral" but 25% now chose "Agree" and only 25% still chose "Disagree" therefore showing a small improvement in ratings.

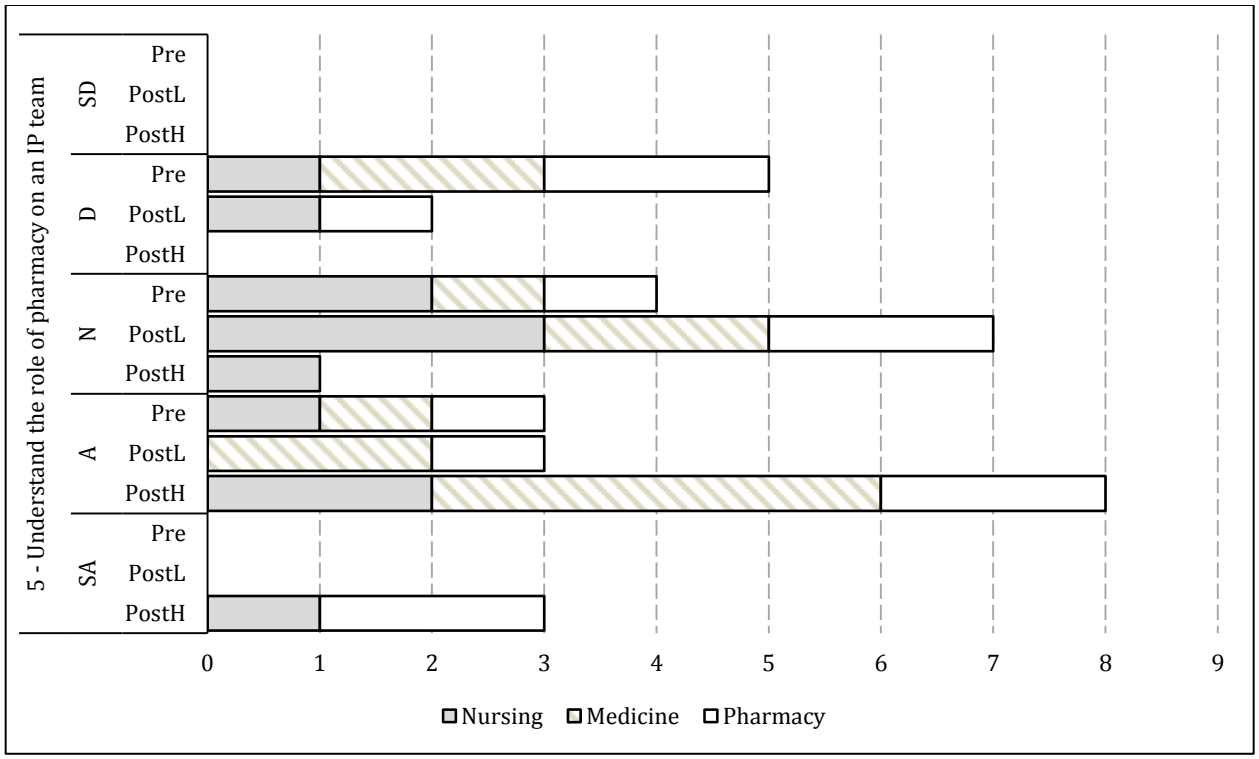

Figure 5. Role of Pharmacy

Remarkably, HFS did not appear to have the same negative impact LFS had on the mock-nursing students. Post HFS, 75% of nursing students chose "Agree" (50%) or "Strongly Agree" (25%). However, the pharmacy students appeared to gain the most benefit with previously conservative scores now split between "Agree" (50%) and "Strongly Agree" (50%). The medical students unanimously chose "Agree" (100%) post HFS, therefore also showing improvement.

In relation to the student groups' ratings of their confidence in communicating effectively with the interprofessional team, both the mock nursing students and pharmacy students had identical pretest scores (Figure 6). For these students, 50% chose "Neutral"

and the remaining 50% was split between “Disagree” and “Agree.” Conversely, the medical students were spread evenly across the board from “Disagree” (25%) to “Strongly Agree” (25%). Post LFS, the scores for nursing students and pharmacy students showed no change, remaining identical to their pretest scores. However, post LFS none of the medical students chose “Strongly Agree” and 50% chose “Neutral” thus appearing to show a more conservative stance post LFS.

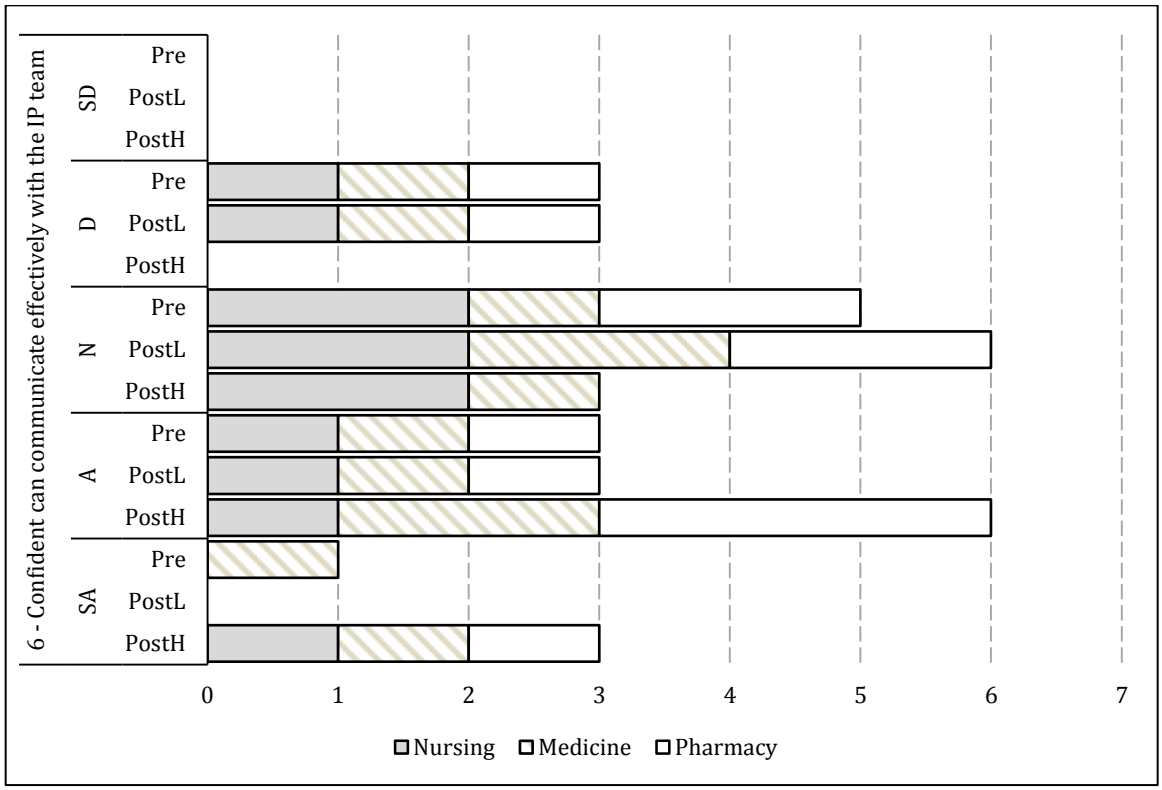

Figure 6. Confidence in Communicating

Though all student groups saw improvements in their scores, the mock pharmacy students appeared to show the greatest improvement. Post HFS, 75% of pharmacy students chose “Agree” and 25% chose “Strongly Agree” leaving no neutral or negative ratings. Medical students also showed improvement with 50% of them choosing “Agree,” 25%

choosing “Strongly Agree” and only 25% remaining “Neutral.” On the other hand, 50% of nursing students rated this item as “Neutral” and the remaining scores were spread evenly between “Agree” (25%) and “Strongly Agree” (25%).

When it comes to the student groups’ ratings of their confidence in their ability to collaborate effectively with the interprofessional team, the pretest scores of both the mock nursing and the medical students were identical (See Figure 7). For these student groups, 50% chose “Neutral” and the remaining 50% were spread evenly between “Disagree” (25%) and “Agree” (25%). The pharmacy students rated themselves less optimistically with 50% choosing “Disagree” and 50% remaining “Neutral” on the subject.

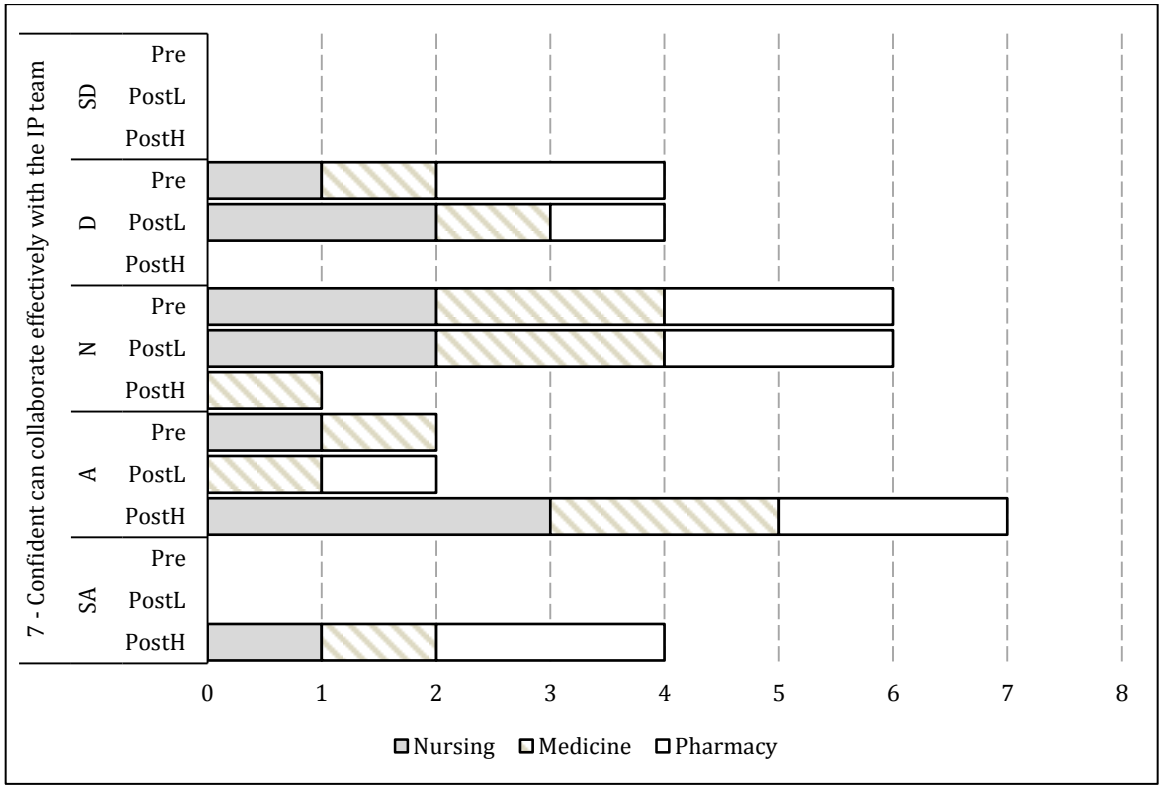

Figure 7. Confidence in Collaborating

Like the previous ITQ item, the mock nursing students' data showed a decline in scores post LFS. The ratings for this item from nursing students were spread evenly post LFS between "Disagree" (50%) and "Neutral" (50%). Remarkably, the rating of the medical students remained unchanged from the pretest ratings. Out of the three student groups, pharmacy appeared to improve the most out of LFS for this item, though the improvement was relatively small. Post LFS 25% of pharmacy students chose "Agree," 50% remained "Neutral," and 25% chose "Disagree."

Though all student groups showed improvement in their scores post HFS, the mock pharmacy students appeared to gain the greatest benefit. For this student group the post HFS scores were spread evenly between, "Agree" (50%) and "Strongly Agree" (50%). Nursing students also gained a strong benefit with 75% choosing "Agree" post HFS and 25% choosing "Strongly Agree." The medical students were a little more conservative as 25% of this student group remained "Neutral." However, they still did show improvement in scores with 50% choosing "Agree" and 25% choosing "Strongly Agree" post HFS.

Satisfaction scores post HFS and LFS were relatively high with no mock student group choosing "Disagree" or "Strongly Disagree" for those items (Figure 8). However, each student group did appear to more strongly agree that the objectives were clear for HFS as opposed to LFS. For example, 100% of nursing students chose "Agree" for this item. However, for HFS, 25% chose "Strongly Agree" and 75% chose "Agree." The pharmacy students shared similar rating for that item. The medical students showed the greatest change with 75% choosing "Agree" post LFS and 75% choosing "Strongly Agree" post HFS.

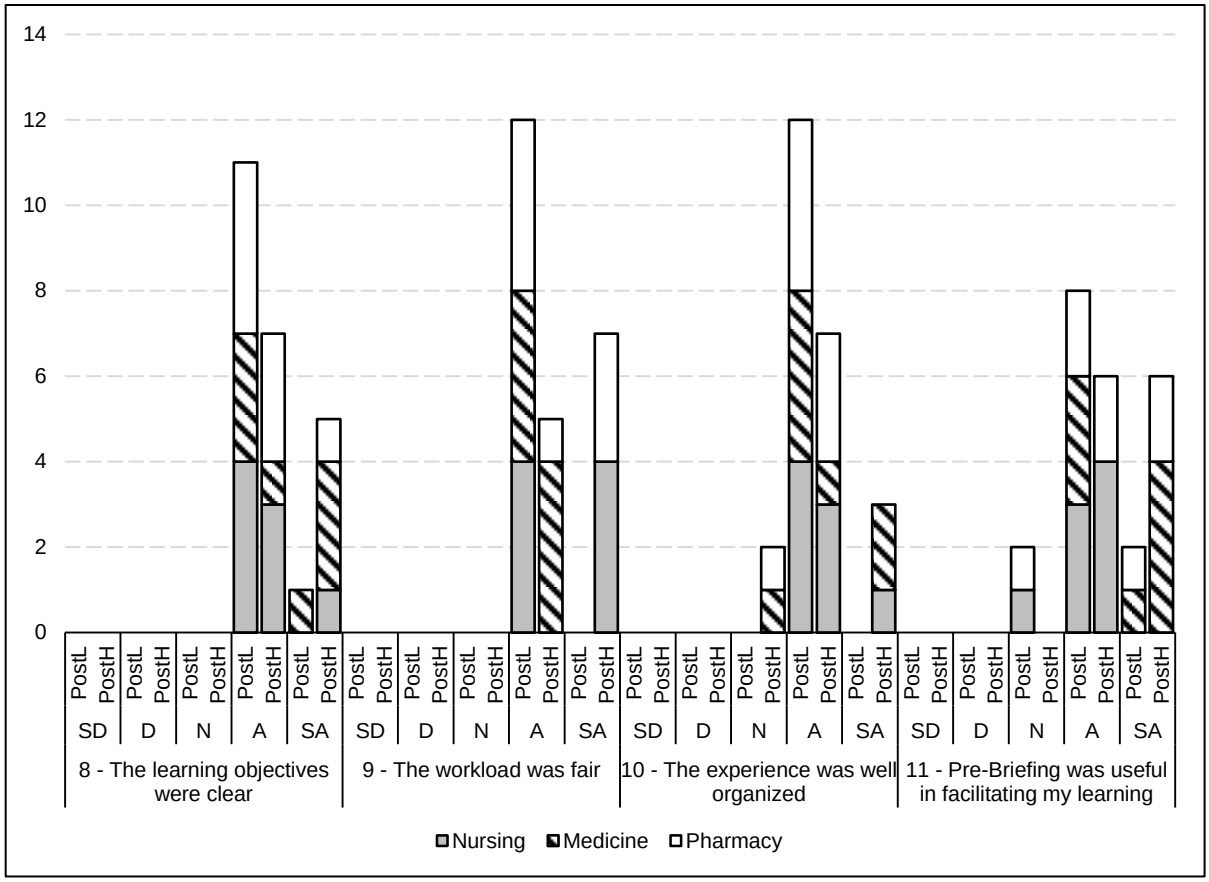

Figure 8. Satisfaction with Objectives, Workload, Experience and Pre-briefing

All the mock student groups chose "Agree" when rating the workload as fair (100%) post LFS. However, both nursing and pharmacy students were most likely to rate this item higher post HFS whereas the medical students remained unchanged. For the nursing students, 100% chose "Strongly" agree for this item whereas 75% of pharmacy students did the same. Like the previous item, all mock student groups chose "Agree" when rating the organization of the experience as well organized (100%) post LFS. However, post HFS, the medical students were more likely than the other student groups to choose "Strongly Agree" (50%). Interestingly, 25% of the medical and the pharmacy students chose "Neutral" post HFS, showing a downward trend from their LFS scores. Nursing

students did not show this trend with 75% choosing “Agree” and 25% choosing “Strongly Agree” post HFS. For pharmacy students 50% chose “Agree,” 25% chose “Strongly Agree,” and 25% remained “Neutral” as previously mentioned.

Unlike the mock medical students, 25% of nursing and pharmacy students chose “Neutral” when describing whether pre-briefing was useful in facilitating learning for LFS. The remaining 75% of nursing students chose “Agree.” This differed from the remaining pharmacy students who chose “Agree” (50%) and “Strongly Agree” (25%). The medical students largely chose “Agree” (75%) and “Strongly Agree” (25%) post LFS. These values did show a positive shift post HFS with no student group rating this item as “Neutral” or below (Figure 8).

However, unlike the other two mock student groups none of the nursing students chose “Strongly Agree” for this item. Instead, 100% of nursing students chose “Agree” compared to 100% of medical students chose “Strongly Agree.” The pharmacy students were evenly split between “Agree” (50%) and “Strongly Agree” (50%).

Debriefing post LFS appeared to be more beneficial for the mock nursing and pharmacy students, than it did for the mock medical students (Figure 9). Post LFS scores for both pharmacy and nursing students were identical with 75% choosing “Agree” and 25% choosing “Strongly Agree” for this item. Conversely, 25% of medical students remained “Neutral” on this item with the remaining 75% choosing “Agree.”

Remarkably, the mock medical students appeared to gain more benefit from debriefing post HFS than the other two student groups did. Post HFS, the medical students’ scores were split evenly between “Agree” (50%) and “Strongly Agree” (50%). Conversely, 25% of pharmacy students remained “Neutral” on the subject and 100% of nursing students

only chose “Agree.” The remaining pharmacy students chose “Agree” (50%) and “Strongly Agree” (25%).

When considering whether they would recommend this experience to others, all mock student groups chose “Agree” (100%) for LFS (Figure 9). However, post HFS all student groups chose “Strongly Agree” (100%), showing a greater inclination toward HFS but also not showing any difference between the student groups.

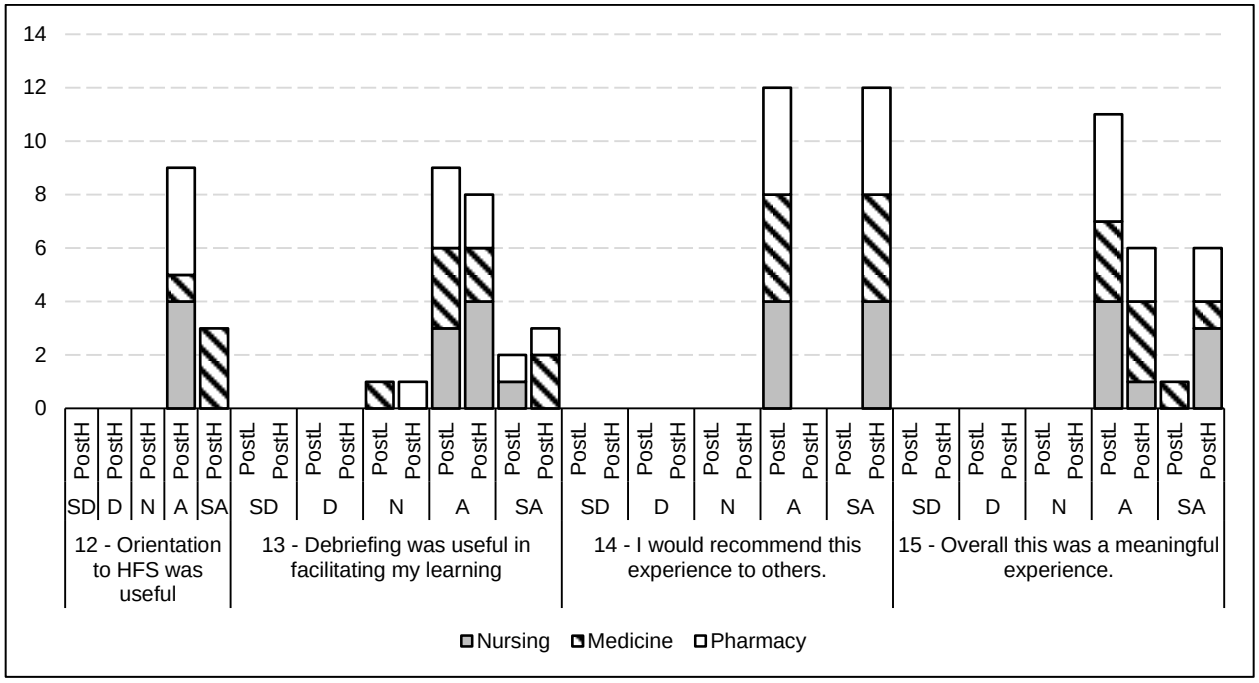

Figure 9. Overall Satisfaction Including Orientation and Debriefing

When considering whether the experience was meaningful, the pharmacy and nursing students had identical responses for LFS: 100% chose “Agree.” However, the medical students were more likely to rate the LFS experience higher with 25% choosing “Strongly Agree” for this item and 75% choosing “Agree.” It also appeared as though nursing students found the HFS experience more meaningful as 75% of this group chose

"Strongly Agree" post HFS with only 25% choosing "Agree." No other group showed such a strong trend toward "Strongly Agree," though the pharmacy students were in second place with 50% choosing that post HFS. Post HFS, the other 50% of pharmacy students chose "Agree." The medical students' scores remained unchanged from the post LFS period.

Inferential Statistics

Comparing individual and mean group scores is another common method of data analysis for the ordinal data collected with a Likert Scale (Table 3). The individual and total group means and SD from the mock data were further analysis using the Wilcoxon Matched Pairs Test to determine whether there were any significant differences between those means (Table 4). The Wilcoxon Matched Pairs Signed Rank Test was used to compare the individual and combined mean score of each item. Since so many significant results were found a Bonferroni correction was applied with a significant *p*-value being recalculated to 0.0167. The Independent Samples Kruskal-Wallis Test was also performed to determine whether the was any significant differences between the total group mean scores of each profession, but there were no significant differences, indicating there was no difference in mean scores between the professions.

Table 3.

Mean Scores on the Interprofessional Teamwork Questionnaire

			Pretest		Post-LFS		Post-HFS	
Item	Profession	n	Mean	(SD)	Mean	(SD)	Mean	(SD)
Function effectively								
	Nursing	4	2.75	(.957)	2.75	(.957)	4.50	(.577)
	Medicine	4	3.50	(1.291)	3.50	(.577)	4.25	(.500)
	Pharmacy	4	3.50	(.577)	3.00	(.000)	4.25	(.957)
	Total	12	3.25	(.965)	3.08	(.669)	**4.33***	(.651)
Importance of teamwork								
	Nursing	4	3.75	(.500)	2.50	(.577)	4.25	(.957)
	Medicine	4	3.75	(.500)	3.00	(.816)	4.50	(.577)
	Pharmacy	4	3.50	(.577)	2.50	(.577)	4.50	(.577)
	Total	12	3.67	(.492)	**2.67***	(.651)	4.42	(.669)
Role of Nursing								
	Nursing	4	2.75	(.500)	2.50	(.577)	4.00	(.816)
	Medicine	4	2.75	(.957)	3.00	(.816)	4.00	(.816)
	Pharmacy	4	2.50	(.577)	3.00	(.816)	3.50	(.577)
	Total	12	2.67	(.651)	2.83	(.718)	**3.83***	(.718)
Role of Medicine								
	Nursing	4	2.50	(.577)	3.00	(.816)	3.75	(.500)
	Medicine	4	2.75	(.957)	2.50	(.577)	4.00	(.816)
	Pharmacy	4	3.00	(.816)	2.75	(.957)	3.75	(.500)
	Total	12	2.75	(.754)	2.75	(.754)	3.83	(.577)
Role of Pharmacy								
	Nursing	4	3.00	(.816)	2.75	(.500)	4.00	(.816)
	Medicine	4	2.75	(.957)	3.50	(.577)	4.00	(.000)
	Pharmacy	4	2.50	(.577)	3.00	(.816)	4.50	(.577)
	Total	12	2.75	(.754)	3.08	(.669)	**4.17***	(.577)
Communicate Effectively								
	Nursing	4	3.00	(.816)	3.00	(.816)	3.75	(.957)
	Medicine	4	3.50	(1.291)	3.00	(.816)	4.00	(.816)
	Pharmacy	4	3.00	(.816)	3.00	(.816)	4.25	(.500)
	Total	12	3.17	(.937)	3.00	(.739)	4.00	(.739)
Collaborate Effectively								
	Nursing	4	3.00	(.816)	2.50	(.577)	4.25	(.500)
	Medicine	4	3.00	(.816)	3.00	(.816)	4.00	(.816)
	Pharmacy	4	2.50	(.577)	3.00	(.816)	4.50	(.577)
	Total	12	2.83	(.718)	2.83	(.718)	**4.25***	(.622)

**Significantly different from pretest at $p < .05$, CI 95%*

Table 4.

Analysis of Variance Between Total Group Mean Scores

Statement	PreTest PostLFS	PostLFS PostHFS	PreTest PostHFS
1. I can function effectively as a team member.	.344	.004	.010*
2. I understand the importance teamwork.	.003	.001*	.023
3. I understand the role of nursing.	.344	.001*	.002*
4. I understand the role of medicine.	.500	.001*	.010*
5. I understand the role of pharmacy.	.180	.002*	<.001*
6. I can communicate effectively with the team.	.383	.004*	.030
7. I can collaborate effectively with the team.	.603	<.001*	.002*

** Significant with Bonferoni Correction p < 0.0167*

Discussion of Mock Results

Analysis of the mock data revealed that the individual and total mean scores changed significantly after participation in the HFS in the areas of understanding the importance of teamwork, understanding the role of nursing and pharmacy and confidence in communicating in an inteprofessional team. Those same changes did not occur after participation in LFS. This would indicate that participation in HF-IPE did have a significantly positive impact on students' knowledge of teamwork and the roles of the interprofessional team. There was also evidence to suggest that students were more satisfied with HFS as compared to LFS. Wellmon et al., (2017) also noted statistically significant improvements in attitudes about the student's own discipline specific competency and autonomy post HFS when compared with a control group. Similarly, de Voest et al., (2013) also noted that post simulation with a real patient student participants became more comfortable with their communication skills. Rossler and Kimble (2016) also found that post HFS students had more positive attitudes about interprofessional learning.

Nursing students also reported less negative attitudes toward learning with other health care professionals after performing HFS, consistent with the results noted in this mock data.

Krueger et al., (2017) mirrored these results as well, noting that HFS produced a positive effect on nursing and medical students' attitudes toward interprofessional collaboration, shared educational experiences, and patient-centered care. These students also expressed that the HFS enhanced their respect for other providers. It increased the value these students placed on effective communication and collaboration. Furthermore, it increased the confidence they had in their communication skills. Like the other studies mentioned here, Paige et al., (2014) also noted immediate improvements in students' team-based attitudes and behaviours post HFS. This supports the findings from the analysis of the mock data.

Wellmon et al., (2017) presented a potential reason why HFS may have dominated in these areas. HFS may present an optimized, patient-centered care moment for these students. This hands-on experience could therefore better reinforce the value of interprofessional practice and help students understand their discipline-specific strengths and limitations. It could also reinforce the importance of working interdependently with other professions (Wellmon et al., 2017). Perhaps LFS does not provide this optimized care moment, leading to the differences noted in the mock data.

In relation to the mock data analysis, when the percentages were reviewed, it was noted that in a number of areas, the pretest percentage scores of the nursing students were lower than the other two professions. For example, this was true for pretest ratings of functioning effectively as a member of the interdisciplinary team, understanding the role of the nurse, and understanding the role of medicine. Sigalet et al., (2012) presented a possible

reason for this phenomenon. According to these authors, the lower pretest scores from nursing students may be due to a lack of exposure to teamwork concepts. This lack of exposure could therefore lead them to have a lower perception of or attitude toward these areas. Lin et al., (2013) added to this by stating that before licensure, these students may have yet to really develop their own professional identity because of this lack of professional qualification and clinical experience. Since these students had very little experience, it therefore makes sense they would rate themselves relatively low.

Unlike the mock nursing and pharmacy students, the medical students rated the item regarding functioning effectively as a member of the interdisciplinary team as higher after LFS than their nursing and pharmacy counterparts. In that case, 50% of them rating it as "Neutral" and 50% of them rating it as "Agree." Conversely, nursing students showed no difference from their pretest scores and 100% of pharmacy students rated the item as "Neutral" post LFS, indicating a decrease from pretest values. Lin et al., (2013) noted that, "Medical education has mostly focused on diseases and issues regarding patient-physician relationships. Little attention is paid to issues about interprofessional interaction that would occur in everyday clinical practice" (p. 510). Perhaps the interprofessional simulation experience was novel, regardless of whether it was LFS or HFS, leading them to consider it in a more positive light. This is in contrast to the nursing students who largely only found real benefit from HFS and they often rated items lower post LFS. This discussion of the proof of concept of the data analysis plan has several important limitations. One important limitation was that reliability and validity of the ITQ has not been established, therefore leading to the potential for error. A second limitation was the use of mock data to provide

proof of concept for the data analysis plan. In summary the proposed data analysis plan is appropriate to guide the data analysis of the quantitative data from the ITQ.

Conclusion

The overall goal of this research practicum was to develop advanced nursing competencies through participating in the data analysis phase of a research study. Through reviewing the literature and participating in relevant meetings, a data analysis plan was created for the ITQ that could be used to measure nursing, medicine, and pharmacy students' changes in knowledge and attitudes towards high fidelity interprofessional education. Although time did not permit the analysis of actual student data, this plan was applied and subsequently modified based on mock data. This proof of concept exercise resulted in modifications to the data analysis plan to ensure that it could be used to analyze, summarize, interpret and display the quantitative data collected from the ITQ.

References

Canadian Nurses Association. (2008). *Advanced nursing practice: A national framework.* Ottawa:Author

Curran, V. R., Mugford, J. G., Law, R. M. T., & MacDonald, S. (2005). Influence of an interprofessional HIV/AIDS education program on role perception, attitudes and teamwork skills of undergraduate health sciences students. *Education for Health, 18*(1), 32–44. https://doi.org/10.1080/13576280500042606

de Voest, M., Raguckas, S., Bambini, D., & Beel-Bates, C. (2013). Interprofessional teaching: An inter-university experience involving pharmacy and nursing students. *Currents in Pharmacy Teaching and Learning*, *5*(5), 450–457. https://doi.org/10.1016/j.cptl.2013.06.004

Garbee, D. 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 Perspectives*, *34*(5), 339–344. https://doi.org/10.5480/1536-5026-34.5.339

Gough, S., Hellaby, M., Jones, N., & MacKinnon, R. (2012). A review of undergraduate interprofessional simulation-based education (IPSE). *Collegian*, *19*(3), 153–170. https://doi.org/10.1016/j.colegn.2012.04.004

Hepburn, Tsukada & Fraser (1996). In E.L. Siegler, K. Hyer, T. Fulmer, M. Mezey, M. (1998). *Geriatric Interdisciplinary Team Training*. New York: Springer Publishing Company

Kardong-Edgren, S., Adamson, K. A., & Fitzgerald, C. (2010). A review of currently published evaluation instruments for human patient simulation. Clinical Simulation in Nursing, 6(1), e25-e35. https://doi.org/doi:10.1016/j.ecns.2009.08.004

Krueger, L., Ernstmeyer, K., & Kirking, E. (2017). Impact of interprofessional simulation on nursing students' attitudes toward teamwork and collaboration. *Journal of Nursing Education*, *56*(6), 321–327. https://doi.org/10.3928/01484834-20170518-02

Liaw, S. Y., Zhou, W. T., Lau, T. C., Siau, C., & Chan, S. W. (2014). An interprofessional communication training using simulation to enhance safe care for a deteriorating patient. *Nurse Education Today*, *34*(2), 259–264. https://doi.org/10.1016/j.nedt.2013.02.019

Lin, Y.-C., Chan, T.-F., Lai, C.-S., Chin, C.-C., Chou, F.-H., & Lin, H.-J. (2013). The impact of an interprofessional problem-based learning curriculum of clinical ethics on medical and nursing students' attitudes and ability of interprofessional collaboration: A pilot study. *Kaohsiung Journal of Medical Sciences*, *29*(9), 505–511. https://doi.org/10.1016/j.kjms.2013.05.006

Munro, B. H. (2005). Statistical methods for health care research (5th ed). Philadelphia: Lippincott Williams & Wilkins.

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. https://doi.org/10.1016/j.jamcollsurg.2013.09.006

Reising, D. L., Carr, D. E., Shea, R. A., & King, J. M. (2011). Comparison of communication outcomes in traditional versus simulation strategies in nursing and medical students. *Nursing Education Perspectives*, *32*(5), 323–327. https://doi.org/10.5480/1536-5026-32.5.323

Rossler, K. L., & Kimble, L. P. (2016). Capturing readiness to learn and collaboration as explored with an interprofessional simulation scenario: A mixed-methods research study. *Nurse Education Today, 36*, 348–353. https://doi.org/10.1016/j.nedt.2015.08.018

Sigalet, E., Donnon, T., & Grant, V. (2012). Undergraduate students' perceptions of and attitudes toward a simulation-based interprofessional curriculum: The KidSIM ATTITUDES Questionnaire. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *7*(6), 353–358. https://doi.org/10.1097/SIH.0b013e318264499e

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. https://doi.org/10.4212/cjhp.v68i4.1471

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

Wellmon, R., Lefebvre, K. M., & Ferry, D. (2017). Effects of high-fidelity simulation on physical therapy and nursing students' attitudes toward interprofessional learning and collaboration. *Journal of Nursing Education, 56*(8), 456–465. https://doi.org/10.3928/01484834-20170712-03

Zou, K. H., Carlsson, M. O., & Quinn, S. A. (2010). Beta-mapping and beta-regression for changes of ordinal-rating measurements on Likert scales: A comparison of the change scores among multiple treatment groups. *Statistics in Medicine, 29*(24), 2486–2500. https://doi.org/10.1002/sim.4012.

Appendix A

Data Analysis Plan For Likert Scales To Measure The Impact Of High Fidelity Interprofessional Education

Integrative Literature Review

Amy Warren

Memorial University of Newfoundland

School of Nursing

Introduction

Interprofessional education (IPE) is a collaborative approach that fosters teamwork among students in health-related fields such as nursing, medicine and pharmacy (de Voest, Raguckas, Bambini, & Beel-Bates, 2013; Garbee et al., 2013; Gough, Hellaby, Jones, & MacKinnon, 2012; Krueger, Ernstmeyer, & Kirking, 2017; Rossler, & Kimble, 2016). IPE encourages students to use their varied educational backgrounds to learn together for a defined period during their education programs. Simulation is a particularly useful teaching and learning approach for IPE, including the use of high-fidelity human patient simulation (HFS) to create a realistic patient scenario for active student engagement (Kardong-Edgren, Adamson, & Fitzgerald, 2010; Paige et al., 2014; Stewart, Kennedy, & Cuene-Grandidier, 2010). Yet, traditional problem-based methods involving low fidelity simulation (LFS) with roundtable discussions are also a useful teaching and learning approach for effective IPE (Curran, Mugford, Law, & MacDonald, 2005; Reising, Carr, Shea, & King, 2011). Although both HFS and LFS are useful as a teaching and learning approach for undergraduate IPE, there is some ambiguity regarding which is the best approach. One of the problems with determining the best learning approach is the need for reliable and valid instruments and the development of a data analysis plan for those instruments (Gough, 2012; Kardong-Edgren, 2010).

This literature review will analyze and synthesize existing literature to provide the evidence that will inform the development of a data analysis plan for the Interprofessional Teamwork Questionnaire (ITQ) designed to measure the impact of undergraduate interprofessional simulation education on teamwork. This review will explore current psychometric evaluation instruments used in interprofessional simulation education and in

particular will focus on self-report evaluation questionnaires utilizing Likert scales. This review will include descriptions of questionnaires, data analysis methods, and recommendations for the data analysis plan for the ITQ Likert scale questionnaire data.

The Literature Review Method

MUN University Libraries search service, CINHAL, PubMed, and Google Scholar data based were utilized in the identification of relevant articles for this literature review. The following search terms were used: Likert scale, data analysis plan, interprofessional teamwork questionnaire, quantitative questionnaire, simulation measurement, interprofessional, teamwork, high fidelity simulation, low fidelity simulation, simulated scenarios, education, undergraduate, pre-licensure, students, medicine, pharmacy, and nursing. This search generated approximately 30 pertinent abstracts and included searching the references of relevant articles.

The criteria used to screen the abstracts included: (a) the article related to high or low-fidelity simulation and undergraduate IPE, (b) the article was a research study, (c) the article included quantitative evaluation measures with Likert scales, and (d) the article included nursing students and at lease one other health-related student group. For the purposes of this literature review, Zou, Carlsson, and Quinn's (2010) definition of the Likert scale was used for article selection: "The Likert scale consists of 'an ordered set of discrete terms or statements from which patients are asked to choose the response that best describes their state or experience" (p. 2486).

Overview of Literature Review

That process yielded a total of twelve research articles from educational institutions around the work including Canada (Curran et al., 2005; Sigalet, Donnon, & Grant, 2012),

the United States (de Voest, 2013; Garbee et al., 2013; Krueger, 2017; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Wellmon, Lefebvre, & Ferry, 2017), Northern Ireland (Stewart, 2010), Singapore (Liaw, Zhou, Lau, Siau, & Chan, 2014), and Taiwan (Lin et al., 2013). The majority of studies also utilized high fidelity simulation (HFS), though two utilized the low fidelity simulation (LFS). The uncontrolled before-after research design was the predominant type utilized by the researchers with eight out of the twelve studies utilizing that design (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Paige et al., 2014; Rossler & Kimble, 2016; Sigalet et al., 2012). Only one study used a controlled before-after research design (Wellmon et al., 2017). One study utilized a randomized control trial design (Lin et al., 2013). One study employed a cross-sectional design (Stewart et al., 2010). Finally, one study applied a posttest-only design (Reising et al., 2011).

Simulation in Interprofessional Education

A total of nine studies conducted an evaluation of simulation as a teaching and learning approach in undergraduate IPE (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017). Three studies used actors portraying patients as either the sole means of simulation (de Voest et al., 2013), in conjunction with LFS (Curran et al., 2005), or in conjunction with HFS (Liaw et al., 2014). Interestingly, one study utilized both LFS and HFS for IPE (Reising et al., 2011). One study employed solely a low-fidelity PBM approach for IPE (Lin et al., 2013). The research studies using Likert Scale instruments will be discussed in relation to the student population sample, variables studied, data analysis plans utilized, and validity and reliability of instruments. These

results appear to highlight a lack research on the evaluation of the impact of HFS as compared to LFS, as a teaching and learning approach in undergraduate IPE

Undergraduate Interprofessional Education

A variety of undergraduate interprofessional student combinations all used in IPE by the authors, with all twelve studies including medicine and nursing students (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017), nine studies including medical students (Curran et al., 2005; Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Sigalet et al., 2012; Stewart et al., 2010), four studies including respiratory therapy students (Garbee et al., 2013; Krueger et al., 2017; Rossler & Kimble, 2016; Sigalet et al., 2012), three studies including pharmacy students (Curran et al., 2005; de Voest et al., 2013; Krueger et al., 2017), two studies including physical therapy students (Rossler & Kimble, 2016; Wellmon et al., 2017); two studies including nurse anesthesia students (Garbee et al., 2013; Paige et al., 2014), one study including radiography students (Krueger et al., 2017), and one study including health care administration students (Rossler & Kimble, 2016). These studies show that nursing students are often involved in IPE, most commonly with medical students but infrequently with other professionals such as pharmacy.

Research Variables Measured by Likert Scales

The Likert Scales used in these evaluation studies measured changes in several variables in an effort to evaluate the impact of simulation in IPE including: attitudes and perceptions; professional roles and scopes of practice; teamwork and collaboration;

confidence in collaboration; and group communication. Nearly all the studies had some sort of focus on the changes in attitudes and perceptions of IPE in the students who underwent the intervention, with ten out of the twelve articles utilizing this variable as an evaluation element (Curran et al., 2005; de Voest et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Paige et al., 2014; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017). Similarly, seven studies had some evaluation of changes in the student's understanding of professional roles and scopes of practice of their peers after participation in IPE (Curran et al., 2005; de Voest et al., 2013; Krueger et al., 2017; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017).

Aspects of communication among the groups were also evaluated by seven studies (Curran et al., 2005; Liaw et al., 2014; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Sigalet et al., 2012; Stewart et al., 2010). Teamwork and collaboration was evaluated by six of the authors (Curran et al., 2005; Garbee et al., 2013; Paige et al., 2014; Rossler & Kimble, 2016; Sigalet et al., 2012; Wellmon et al., 2017). The least predominant variable was confidence in collaboration, with only five sets of authors measuring this variable in their study (Garbee et al., 2013; Paige et al., 2014; Liaw et al., 2014; Lin et al., 2013; Wellmon et al., 2017).

Data Analysis Tools for Likert Scales

The data analysis tools used for the Likert Scales in five of the studies included t-tests (Curran et al., 2005; de Voest et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Stewart et al., 2010), six studies used paired t-tests (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Paige et al., 2014; Sigalet et al., 2012; Wellmon et al., 2017), and

six studies used analysis of variance (ANOVA) (Curran et al., 2005; Krueger et al., 2017; Lin et al., 2013; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017). As well, only seven of the research studies established internal consistency and reliability of the instrument using Cronbach's α (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010). This apparent lack of thorough reliability reporting is consistent with Kardong-Edgren et al., (2010) who reported a paucity of reliability and validity data in their review of evaluation instruments for human patient simulation, although this does not mean such evaluation did not take place. Two of the studies analyzed in this literature review did not directly report data on validity or reliability (Curran et al., 2005; Reising et al., 2011).

Five studies directly analyzed or to some degree discussed the reliability and validity of the instruments used (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Sigalet et al., 2012), four studies included only the reliability of the instruments (Paige et al., 2014; Rossler & Kimble, 2016; Stewart et al., 2010; Wellmon et al., 2017), and one study discussed only the validity of the instruments (de Voest et al., 2013). Of the six total studies that referred to the validity of the measurement instruments, three reported the validity of the existing measures (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014), two calculated the validity of the purpose designed instruments (Lin et al., 2013; Sigalet et al., 2012), and one referred to the content validity analysis of a panel of experts (de Voest et al., 2013).

Less common data analysis tools included: Tukey's honestly significant difference (HSD) (Lin et al., 2013; Wellmon et al., 2017); Cohen's *d* and component analysis with varimax rotation (Sigalet et al., 2012); and analysis of covariance (ACOVA) (Liaw et al.,

2014). Three articles utilized nonparametric methods for data analysis including the Mann-Whitney *U* test (Rossler & Kimble, 2016; Wellmon et al., 2017), chi-square test (Liaw et al., 2014; Rossler & Kimble, 2016), as well as Wilcoxon signed rank and a Kruskal-Wallis test (Rossler & Kimble, 2016). Of the 12 studies included in this review, only one (Stewart et al., 2010) included a confidence interval (CI) among the provided data. However, that study did not state a level of statistical significance. All other studies stated a level of significance of 0.05, or declared data significant that had a *p*-value of 0.05 or below. These studies show that there are several statistical tests that could be used in the data analysis plan for the ITQ data including, but not limited to; Cronbach's α, t-test, paired t-tests, ANOVA, HSD, Cohen's *d*, ACOVA, Mann-Whitney *U* test, Wilcoxon signed rank, and a Kruskal-Wallis test.

Likert Scales in the Evaluation of Interprofessional Education

One of the limitations of current psychometric evaluation instruments for IPE is that researchers are using original, purposely designed instrument that have not been previously used or validated (de Voest et al., 2013; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Sigalet et al., 2012; Stewart et al., 2010). While six studies utilized instruments from previous research (Curran et al., 2005; Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Rossler & Kimble, 2016; Wellmon et al., 2017), four used modified versions of existing instruments (Curran et al., 2005; Liaw et al., 2014; Lin et al., 2013; Rossler & Kimble, 2016).

The following Likert scales used in IPE will be discussed: Team Skills Survey, Readiness for Interprofessional Learning Scale (RIPLS), Health Professional Collaboration Scale (HPCS), Interdisciplinary Education Perception Scale (IEPS), Attitudes Toward

Health Care Teams Scale (ATHCTS), Student Satisfaction and Self-Confidence in Learning Scale (SCLS), Attitude Towards Teamwork in Training Undergoing Designed Educational Simulation (ATTITUDES) questionnaire, Jefferson Scale of Attitudes Toward Physician-Nurse Collaboration, Interprofessional Communication and Collaboration Questionnaire (ICCQ) and Teamwork Assessment Scale (TAS). There are also five unnamed surveys for IPE evaluation, which will be discussed. The following is a discussion of these Likert Scale instruments, including the context of their use, description of the instrument, and their data analysis plans.

Team Skills Survey (TSS)

Curran et al., (2005) adapted the TSS from Hepburn, Tsukuda and Fraser (1996) in their investigation of role perception, attitudes, and teamwork skills in a scenario involving a standardized patient for interprofessional HIV/AIDS education. In their pretest-posttest and time-series study, the Canadian nursing, medical, and pharmacy students completed the 15-item questionnaire as part of the post-evaluation. A total of 133 students completed this assessment, with 45 being from nursing, 62 being from medicine, and 26 being from pharmacy. Included in the TSS were Likert scale items ranging from 1, which indicated "Poor" to 5, which indicated "Excellent." The lowest possible total score was 15 and the highest possible score was 75 with higher scores denoting more positive self-assessment of team skills. The authors did not report on the reliability or validity of the tool. The data analysis plan utilized ANOVA to compare mean scores between the three professions therefore implying descriptive statistics were performed.

Readiness for Interprofessional Learning Scale

The Readiness for Interprofessional Learning Scale **(**RIPLS) is a relatively well-known instrument and was utilized as an evaluation tool for high fidelity simulation by Rossler and Kimble (2016) as well as by Wellmon et al. (2017). Both studies involved a sample of American health sciences students, who underwent a high-fidelity patient-simulator scenario. However, unlike Rossler and Kimble who only utilized one other Likert-based questionnaire, Wellmon et al., utilized three additional instruments. Rossler and Kimble utilized RIPLS as a pre-and post-evaluation with nursing, respiratory therapy, health administration, and physical therapy students. Their study sample included 50 students with 25 being from nursing, 10 being from respiratory therapy, five being from physical therapy, and 10 being from health administration. Similarly, Wellmon et al., also utilized RIPLS as a pre-and post-evaluation tool with a sample of 151 nursing and physical therapy students. However, Wellmon et al., (2017) utilized a much larger sample with 68 being from nursing and 83 being from physical therapy. Furthermore, Wellmon et al. (2017) included a control group whereas Rossler and Kimble (2016) did not include a control group.

According to Wellmon et al., (2017) RIPLS focuses on the attitudes of the student toward teamwork and collaboration in learning. It involves 19 items rated on a 5-point Likert scale from 1 to 5 with 1 denoting “Strongly disagree” and 5 denoting “Strongly agree.” The scores of each item are totalled with higher overall scores representing more positive attitudes toward collaboration and lower scores representing the opposite. These 19 items fell into three scale domains: professional identity, learning from others, and understanding the roles and responsibilities of other disciplines.

Rossler and Kimble (2016) reported a total Cronbach's alpha for RIPLS of 0.84 and 0.89. This is consistent with Wellmon et al., (2017) who noted the reliability of RIPLS was satisfactory. However, Rossler and Kimble enriched this assessment by providing Cronbach's alpha rating for subscales: teamwork and collaboration (0.88), negative professional identity (0.76), positive professional identity (0.81), roles and responsibilities (0.43). The Cronbach alpha rating of 0.43 for roles and responsibilities denoted lower internal consistency and reliability of that subscale. Unfortunately, this meant there might have been a relation between a lack of significant results for that subscale area and the lower alpha rating.

These two authors varied significantly in their data analysis plan, with Wellmon et al., (2017) choosing parametric measures while Rossler and Kimble (2016) utilized nonparametric measures to analyze RIPLS data. Wellmon et al., used a two-group by two-group repeated measures research design and statistical analysis using ANOVA, with one pair denoting learning intervention versus control, and the other pair denoting time pre-learning versus post-learning. Tukey's honestly significant difference test (HSD) was performed as a post hoc analysis. This method was also used to examine two other instruments reported later in this review.

Rossler and Kimble also noted that the subscale data from the RIPLS had a non-normal distribution thus breaking one of the foundational assumptions for parametric data analysis (Munro, 2005; Zou et al., 2010). Therefore, the change over time for pre-and post-scores was analyzed using a Wilcoxon signed rank test: the nonparametric equivalent of a paired *t*-test (Munro, 2005). Additionally, the Kruskal-Wallis test was utilized to assess the differences among the student groups. Nonparametrically, this is the equivalent of a one-

way ANOVA (Munro, 2005; Rossler & Kimble, 2016). For post hoc analysis, the Mann-Whitney *U*, the nonparametric analog of a *t*-test, was used (Munro, 2005).

Health Professional Collaboration Scale

A second Likert-based questionnaire that was utilized in pre-and post-testing by Rossler and Kimble (2016) was the Health Professional Collaboration Scale (HPCS). The HPCS consists of 12 items reflecting a five-point Likert scale, and measures collaboration perceptions within the student sample. The highest possible total score for HPCS is 60, denoting a highly positive perception toward collaboration. The lowest possible total score is 12, signifying a significantly less positive perception toward collaboration. The authors noted that previous research determined a Cronbach's alpha rating on the HPCS of 0.95, which indicates a relatively high internal consistency and reliability. The data analysis plan of Rossler and Kimble included using the Kruskal-Wallis test to explore the differences among the student groups in HPCS scores, and the Mann-Whitney *U* was utilized for post hoc analysis.

Interdisciplinary Education Perception Scale

One of the three Likert-based questionnaires utilized by Wellmon et al., (2017) was the Interdisciplnary Education Perception Scale (IEPS). That 18-item questionnaire included a six-point Likert scale with 1 signifying "Strongly disagree" and 6 signifying "Strongly agree." The IEPS assessed the students' perceptions regarding their profession's capacity to collaborate with others from different professions and included four subscales: perceptions of competency and autonomy, beliefs surrounding the need for cooperation with other disciplines, perception of actual cooperation, and understanding the values of other disciplines. Like other questionnaires reported, scores were totalled with higher

scores denoting more positive perceptions. No data regarding internal consistency or reliability were reported by the authors. Data analysis methods for the IEPS were similar to those reported under the RIPLS section of this review.

Attitudes Toward Health Care Teams Scale

Continuing with Likert-based questionnaires utilized by Wellmon et al., (2017) the Attitudes Toward Health Care Teams Scale (ATHCTS) is a 21-item questionnaire using a 6-point Likert scale with zero signifying "Strongly disagree" and 5 signifying "Strongly agree." The ATHCTS assessed the attitudes of students toward collaboration with their team members within three overarching domains: quality of care or team value, cost of team care or team efficiency, and shared leadership. Uniquely, the ATHCTS also assessed perceived care quality. Like the IEPS, the ATHCTS also can be totalled with higher values signifying more positive attitudes toward collaboration. No data regarding internal consistency or reliability were reported by the authors. Data analysis methods for the ATHCTS were similar to those reported under the RIPLS section of this review.

Student Satisfaction and Self-Confidence in Learning Scale

The fourth Likert-based questionnaire utilized by Wellmon et al., (2017), was the Student Satisfaction and Self-Confidence in Learning Scale (SCLS). Unlike the other assessments, this tool was only administered as a post-assessment and only to the learning intervention group. The SCLS is a 13-item questionnaire with a 5-point Likert scale with 1 signifying "Strongly disagree" and 5 signifying "Strongly agree." Though not related directly to interprofessional collaboration, this scale did assess the student's confidence level and satisfaction with the activity, therefore indirectly referencing interprofessional collaboration. The portion of the questionnaire that pertained to satisfaction with teaching

methods contained five items: learning materials, facilitation, motivation, and suitability of the simulation. The portion that focused on the student's self-confidence contained eight items: content mastery and necessity, skill development, availability of resources, and knowledge of how to obtain help to solve clinical problems in simulation. Like other assessments, the SCLS can be totalled with higher values signifying higher satisfaction and self-confidence levels.

Wellmon et al., (2017) conducted a descriptive analysis of the SCLS, including individual and total subscale items. This descriptive analysis included means and standard deviations. The Mann-Whitney *U* was used to differentiate levels of satisfaction between the two student groups on individual items. Interestingly, a parametric measure was also employed on both totalled subscale scores: the paired sample *t*-test.

Attitudes Towards Teamwork in Educational Simulation

Sigalet et al., (2012) developed and psychometrically evaluated the Attitude Towards Teamwork in Training Undergoing Designed Educational Simulation **(**ATTITUDES) questionnaire by using a Canadian sample of 127 nursing, 35 medical, and 34 respiratory therapy students. These students completed the ATTITUDES questionnaire before the three-hour IPE curriculum module with HFS and completed the survey afterward. The ATTITUDES questionnaire was based on a plethora of previous tools including the Cockpit Management Attitudes Questionnaire, Anti-Air Teamwork Observation Measure, RIPLS, Naval Training Attitudinal Survey, Human Factors Attitude Scale, and ATHCTS. The ATTITUDES questionnaire included 30 items and utilized a 5-point Likert scale, with 1 representing "Strongly disagree" and 5 representing "Strongly agree."

Like other questionnaires discussed, the ATTITUDES questionnaire focused on the participant's perceptions and was divided into five domains: relevance of IPE, relevance of simulation, communication, situation awareness, as well as roles and responsibilities. Higher total scores in each of these domains represented more positive perceptions toward IPE. Sigalet et al., (2012) performed additional psychometric analysis using varimax rotation for validity of constructs and Cronbach's alpha for internal reliability. Cronbach's alpha ratings ranged from 0.78 to 0.91 and data from the varimax rotation yielded a percentage of variance from 9.8% to 13.8% with a cumulative total of 61.8%. Overall, this showed the ATTITUDES questionnaire had strong reliability and validity.

The data analysis plan for the ATTITUDES questionnaire included descriptive statistics on each of the items, overall scale scores, and subscale scores. Parametric statistical measures were also utilized including the paired sample *t*-test for detecting pretest and posttest differences as well as ANOVA to analyze differences between the three student groups. Uniquely, Sigalet et al., (2012) also calculated effect sizes between the mean scores in pretest and posttest ATTITUDES data using Cohen's *d*.

Jefferson Scale of Attitudes Toward Physician-Nurse Collaboration

Krueger et al., (2017) utilized the Jefferson Scale of Attitudes Toward Physician-Nurse Collaboration (JSAPNC) questionnaire with American nursing and medical students who participated in a HFS. Included in the study by Krueger et al., (2017) were a total of 293 students with 268 being from nursing and 25 from medicine as well as an undisclosed number of students from pharmacy, respiratory therapy, paramedicine, and radiography. Interestingly, these students also participated in a pre-session online community one week prior to the three-hour simulation, and a pre-briefing session immediately prior to the HFS.

Though only medical and nursing students completed the assessments, students from a variety of professions such as respiratory therapy, paramedic, radiography, and pharmacy technicians also participated in the HFS.

The JSAPNC questionnaire included 15 items and utilized a 4-point Likert-type scale. These items were divided into four categories: shared educational and collaborative relationships; caring as opposed to curing; nurse's autonomy; and physician's authority. Like other questionnaires explored, a higher score indicated a more positive attitude toward interprofessional collaboration. A Cronbach's alpha of 0.85 was noted for medical and nursing students with these authors determining a reliability coefficient of 0.83 for the presimulation survey and subscale values of the following: shared educational and collaborative relationships (0.758); caring as opposed to curing (0.627); nurse's autonomy (0.599); and physician's authority (0.583). Post-simulation Cronbach alpha scores were as follows: shared educational and collaborative relationships (0.816); caring as opposed to curing (0.634); nurse's autonomy (0.715); and physician's authority (0.635).

For the JSAPNC questionnaire, Krueger et al., (2017) utilized descriptive statistics as well as parametric measures including the one-way ANOVA, *t*-tests, and paired *t*-tests. Totals were analyzed pre-and post-simulation as well as the pre-and post-simulation data from each of the four domains.

Interprofessional Communication and Collaboration Questionnaire

Lin et al., (2013) developed and utilized the Interprofessional Communication and Collaboration Questionnaire (ICCQ) for their pilot study involving Taiwanese medical and nursing students. A total of 36 students participated in the study with 18 being from nursing and 18 being from medicine. Unlike many of the studies discussed earlier, the ICCQ was

utilized as an evaluation tool post participation in a low-fidelity PBM. Post participation comparisons were made between groups of students that were of the same profession and groups that were mixed.

The 10-item ICCQ had four-point Likert scales and assessed the students' confidence and attitude toward interprofessional teamwork. The Likert scale ranged from 1, which represented "Strongly disagree" to 4, which represented "Strongly agree." Items on the ICCQ included the following: understand the role of the other professions in clinical situation; recognize and respect roles and contribution of other professions; recognize and respect competence in others; capable of working as a team with people from other professions; capable of communication, coordination, and conflict resolution; recognize and respect leadership in collaborative practice; capable of facilitating collaborative practice; confident in own ability as well as others'; capable of patient-centered collaborative practice; and willing to work as a team and share the same goal with people from other professions" (Lin et al., 2013, p. 509).

In addition to this IPE tool, Lin et al., (2013) also utilized the Self-Directed Learning Scale (SDLS), the Critical Thinking Scale, and a general performance questionnaire thus constituting four dimensions of interest. Though these also utilized the four-point Likert scale, they were not deemed as sufficiently pertinent to IPE to discuss in detail. When all four dimensions were taken together, internal consistency was found to be 0.70 using Cronbach's alpha. However, the ICCQ had a much higher alpha rating of 0.93.

In their data analysis plan, Lin et al., (2012) utilized descriptive statistics including percentages. Criterion-related validity was established using a one-way ANOVA. Associations were tested between the scores in the four dimensions and the participant

group, since multiple participant groups were used. The Tukey's honestly significant difference (HSD) comparison was also utilized to assess criterion-related validity.

Teamwork Assessment Scale

Garbee et al., (2013) utilized teamwork Assessment Scale (TAS) in their investigation of interprofessional teamwork in HFS with American nursing, respiratory therapy, medical, and graduate-level nurse anesthesia students. A total of 52 students participated with 28 being from nursing and nurse anaesthesia, 11 being from medicine, and 13 being from respiratory therapy. In addition to observational ratings for other tools performed by trained raters, all participants completed the TAS after undergoing the HFS. Participants also completed the TAS and the Communication and Teamwork Skills Assessment (CATS), during both high-fidelity simulations.

The TAS involved students rating the performance of team members by profession and the team's interactions using a six-point Likert-type scale ranging from 1 that denoted "Definitely no" to 6 which denoted "Definitely yes." Subscales within the TAS assessed aspects of team-based behavior (TBB), shared mental model (SMM), and adaptive communication and response (ACR). As noted by Garbee et al., (2013) "The subscales measure teamwork competencies such as communication, role clarity, flattened hierarchy, mental rehearsal, situational awareness, cross-monitoring, resource management, shared mental model (often referred to as 'on the same page' in a situation), and anticipatory response" (p. 341). These authors also noted that the TAS had previously been shown to have face validity, content validity, and convergent validity. Still, internal consistency coefficients were not reported.

In addition to descriptive statistics paired samples *t*-tests were tabulated for comparison of means between scenarios. However, the authors made no mention of further statistical tests or post hoc analysis.

Other Questionnaires

Five of the studies analyzed in this literature review had questionnaires that had no specified name by the authors, so they were called Unnamed Questionnaires one through to five.

Number One: Stewart et al., (2010) utilized Unnamed Questionnaire Number One (UQ1) to assess IPE within a sample of medical and nursing students in Northern Ireland. A total of 95 students participated in the study with 46 being from nursing and 49 being from medicine. The authors developed the UQI as a post-simulation evaluation method. It included 32 statements rated on a five-point Likert scale with 1 signifying "Strongly disagree" to 5 signifying "Strongly agree." In addition to assessing the attitudes toward shared learning, this questionnaire involved questions assessing the perception of the workshop, the skills acquired, and included open-ended questions that were analyzed qualitatively.

The 32 statements fit within three domains: the development of clinical knowledge and skills; communication and teamworking; professional identity and role awareness; and attitudes to shared learning. Cronbach's alpha was reported for each of these domains and were as follows: the development of clinical knowledge and skills (0.84); communication and teamworking (0.89); professional identity and role awareness (0.79); and attitudes to shared learning (0.69). Consequently, the domain of attitudes to shared learning had the lowest internal consistency.

Stewart et al., (2010) rescaled the Likert items so that they reflected a zero to 100 scale instead of one to five, with zero being coded as a negative response and 100 coded as a positive one. Specific statistical analysis of each domain for both profession was then performed using the Student's *t*-test and ANOVA. However, the authors made no mention of further statistical tests or post hoc analysis.

Number Two: De Voest et al., (2013) tackled IPE in a particularly unique way. Though the interprofessional element was closer to the PBM, the 14 American nursing and 15 American pharmacy students met with an older adult in their community several times over the course of several semesters. These home visits were interspersed among four interprofessional seminars that included the creation of interprofessional patient-teaching plans, and the pharmacy students also performed additional assignments for their course. A unique questionnaire was developed to evaluate the study. The Unnamed Questionnaire Number Two (UQ2) included the following elements: student demographics, perceptions of each profession, understanding of each profession's scope of practice, the significance of the endeavor as assessed through questions, and qualitative comments.

Only the first three elements were administered before the intervention began whereas in the posttest only the last four elements were assessed. The last four elements had Likert and Likert-type scales ranging from 1 to 4. For example, the comfort level when working with older adults ranged from 1 representing "Very uncomfortable" to 4 representing "Very comfortable." The status of each profession was assessed from 1 being very low to 4 being very high. Attitudes toward each profession were rated from 1 denoting "Very negative" to 4 denoting "Very positive." Items that compared the knowledge, academic training, and professional competence between the professions were rated from 1

denoting “strongly disagree” to 4 denoting “strongly agree.” Though no information regarding internal consistency was provided, content validity was ensured using expert reviewers. De Voest et al., (2013) utilized descriptive statistics including percentage values for student responses. The authors also used *t*-tests to analyze the quantitative questionnaire data. However, the authors made no mention of further statistical tests or post hoc analysis.

Number Three: Paige et al., (2014) utilized an operating room HFS with American cohorts of 18 nursing, 28 medical, and 20 junior-level nurse anesthesia students, created the Unnamed Questionnaire Number Three (UQ3) assessed student self-efficacy in teamwork competencies using 15 Likert-type items ranging from 1 denoting “Not confident at all” to 6 denoting “Completely confident.” It was administered as a pre-and post-assessment. Additionally, trained observers utilizing the Operating Room Teamwork Assessment Scales (ORTAS) also collected observational data. Paige et al., (2014) utilized descriptive statistics as well as paired *t*-tests to statistically analyze the UQ3 data. Additionally, a Bonferroni adjustment was performed with the paired *t*-test.

Number Four: Reising et al., (2011) provided a unique comparison between low-fidelity (LF) and high-fidelity (HF) simulation modalities with their American nursing and medical student cohorts. A total of 60 students participated with 41 being from nursing and 19 being from medicine. Participants were randomly assigned either to either a LF group or a HF group and after the simulation all groups completed the Unnamed Questionnaire Number Four (UQ4). The overall focus of the UQ4 was on interprofessional communication and affective elements. It included nominal “yes or no” level questions as well as Likert-type scales ranging from 1, denoting “Low” to 5 denoting “High.” The scaled items included questions about overall stress levels due to the intervention,

managing stress in group interactions, nervousness level at the end as opposed to be beginning of the simulation, and the level of respect the student exhibited when presenting ideas in the scenario. No data was provided regarding the validity or reliability of this instrument. Reising et al., (2011) performed descriptive statistics and reported *p*-values but no direct acknowledgement of specific statistical tests applied to the data was noted.

Number Five: Like Reising et al., (2011) Liaw et al., (2014) focused on the communication aspect of IPE. The authors' Singaporean nursing and medical student sample participated in a HFS with pre-and posttest evaluation. A total of 127 students participated in the study with 94 being from nursing and 33 being from medicine. The eight-item Unnamed Questionnaire Number Five (UQ5) assessed the perceptions of interprofessional learning using a five-point Likert scale. The authors reported a Cronbach alpha of 0.91 to 0.92 in their study.

Liaw et al., (2011) performed descriptive statistics as well as a chi-square tests and *t*-tests to determine associations between demographics and the two student professions. To determine the change between pre-and posttest data, a paired *t*-test was utilized. Additionally. ACOVA was also performed using pretest data as covariances to gauge the effect on post-test data.

Interprofessional Teamwork Questionnaire

The Interprofessional Teamwork Questionnaire (ITQ) extracts ordinal level data and includes seven statements that are rated by the participant on a six-point Likert scale from "Strongly agree" to "Strongly disagree." Those statements focus on areas of individual functionality in an interprofessional team, the significance of interprofessional collaboration in the simulation, comprehension of the role of each profession involved,

individual communication confidence in an interprofessional team, and confidence in collaboration for care planning. In addition to these initial statements, the post-test version of the ITQ also includes nine items that assess the participant's level of satisfaction with the experience.

Items on the ITQ include the level of clarity of the learning objectives, the fairness of the workload, the experience was organized, utility of each component in the simulation experience, (pre-briefing, orientation to the simulator and debriefing), whether they would recommend the scenario to others, and whether the experience was meaningful overall. One of the limitations of the ITQ is that it is a purposeful questionnaire and reliability and validity have not been established.

This literature review has shown that when creating data plans for instrument that use a Likert scales it is appropriate to consider both non-parametric (Wilcoxon, Kruskal-Wallis, Mann-Whitney *U*) and parametric (ANOVA, *t-test*, and HSD) measures. This literature review has also shown that the majority of the instruments had established internal reliability using Cronbach's alpha. These findings can be applied to the creation of the data analysis plan for the Interprofessional Teamwork Questionnaire (ITQ).

Likert Scale Data Analysis Plans

The data analysis plan for analysis of the Likert Scales used in the Interprofessional Teamwork Questionnaire (ITQ) will focus on statistics relevant to a pretest, posttest, repeated measures cross over research design and the nominal data collected by the ITQ. The pretest version of the ITQ is completed once, but the posttest version is completed twice: once after HFS and once after LFS. SPSS version 23 will be used to compute all statistics that will utilized in this study. Different types of data in different circumstances

can necessitate different statistical tests, therefore it is essential to consider whether the Likert-scales data is viewed as ordinal or nominal data. Zou et al., (2010) noted that the outcome variables on Likert scales can be interpreted as ordinal or being on a continuum, making regression analysis and parametric statistics applicable. This is assuming the assumptions for parametric statistical measures are met; independence, normalcy, and variance (Munro, 2005). The following section presents a discussion of the items to be considered when developing a data analysis plan including sample size, descriptive statistics, levels of significance, and non-parametric and parametric measures. A discussion of the data analysis plan for the ITQ is also presented.

The Sample Size

It is important to consider whether the study sample is convenient or random when selecting appropriate statistics. The ITQ is designed for a convenient sample of senior level, undergraduate nursing, pharmacy, and medicine students. However, participants can be randomly assigned to teams and sequences HFS or LFS first. In this way the assumption of randomization would be met for the use of parametric statistics. It is important to calculate a sample size that meets the needs of the data analysis because if the sample size is too small it will not be able to detect significant differences (Munro, 2005).

Several factors need to be considered when determining appropriate sample size including significance level, power, and effect size. The significance level correlates with reducing the probability of a type I error where a false positive occurs. This value will be discussed in later sections of this paper. Conversely, a higher power level means there is a greater chance of avoiding a type II error where a false negative occurs. An adequate power level of 80% is generally accepted.

The largest sample of students noted in the research studies accessed for this literature review was 293 (Krueger et al., 2017) and the smallest sample size was 29 (Voest et al., 2013). Between these limits, would be the study by Reising et al., (2011) which compared HFS to LFS to explore communication outcomes. The study used a sample of 41 nursing students and 19 medical students. However, Reising et al., (2011) performed descriptive statistics and only reported on *p*-values. The authors did not delineate the specifics of the tests they performed.

In terms of the professional groups surveyed, the sample utilized by Curran et al., (2005) utilized pharmacy, nursing, and medical student populations. In that case, Curran et al., (2005) had 45 nursing, 62 medicine, and 26 pharmacy students that participated in the study.

Descriptive Statistics Used in Likert Scales

A data analysis plan for a Likert Scale should include descriptive statistics such as measures of central tendency, consideration of normal distribution, measures of variability, and measure of symmetry. It must also consider how the data could be presented such as tables, graphs or figures. It is imperative that the initial data analysis of the ITQ begins with descriptive statistics, as this will inform the path of all other data analysis that must follow. Appropriate descriptive statistics for the ITQ include the measures of central tendency, measures of variability, and measures of symmetry.

Measures of central tendency. The measures of central tendency include the mean, median, and mode with the most telling of these three being the mean. Calculation of a mean for a dataset involves summing all the values then dividing the result by the number of values present in the dataset (Munro, 2005). This makes the mean a particularly sensitive measure,

because the mean can be influenced by very large and very small items in the dataset (Munro, 2005). This sensitivity to the dataset bodes well for better power in statistical tests, but is only appropriate when the distribution is normal or manipulated to be closer to normal. This is why parametric measures of data analysis hinge so strongly on the mean (Munro, 2005).

On the other hand, the median and mode are much more robust but not quite as powerful. No matter what shape the distribution holds, the median, or "middle," of the dataset will remain the same, hence why many nonparametric measures hinge on the median (Munro, 2005). With this statistic, 50% of the data will fall below and 50% of the data will fall above the median point (Munro, 2005). Unlike the mean, it has no algebraic expression but is instead calculated by looking at an arranged data table, counting the total then finding the data point that is in the middle of that table (Munro, 2005). If the total is even, the median will be the mean of the middle two values. The mode is simply the values that appear most frequently in a dataset. These values are important because in a normal distribution the mean, median, and mode all hover closely to the exact same number (Munro, 2005). Yet, why is a normal distribution important to choosing statistics?

Since parametric measures rely so heavily on means, one of the key assumptions to using them is that the data is normally distributed. As previously noted, the mean is a very sensitive and powerful measure, therefore it would not be appropriate if the data showed significant skewness (Munro, 2005; Zou et al., 2010). Conversely, the median and mode are more robust against these changes, though not as powerful. They form the basis of nonparametric measures. Such measures may become particularly important as Likert-scaled data is bounded, and therefore has a propensity at times to show skewness (Zou et

al., 2010). Therefore, the branching choice between parametric and nonparametric measures necessitates the determination of normalcy in the distribution (Munro, 2005; Zou et al., 2010).

Though it is possible to utilize data transformations to manipulate the distribution into appearing more normal, such measures would require more extensive consultation with statisticians to determine whether such actions would be appropriate as it would involve a fundamental shift in what most of the values appear as in the data tables. Similarly, a consultation would need to take place if significant data is missing and/or exhibiting a systematic pattern when SPSS Missing Value Analysis is performed on initial analysis. Inappropriate data deletion or replacement could significantly pollute the results from all tests that follow (Munro, 2005).

It is also important to consider whether the sample is homogenous or heterogenous. Homogenous samples have very little variability in data sets whereas heterogenous samples have much higher variability. Even if they have the same mean, there can be great variability within a heterogenous sample hence this is why measures of variability are important (Munro, 2005). Standard deviation is the most frequently used measure of variability due to the fact it creates an unbiased estimate of a population variance. Like the mean, it is sensitive to high and low values, and is therefore most appropriate for more normal distributions (Munro, 2005). However, it can tell us important information about whether a distribution is a normal, bell-shaped curve. Within three standard deviations of the mean over 99% of the data is covered. In two standard deviations, 95% is covered. In one standard deviation, 68% of the values are covered (Munro, 2005). Therefore, if the distribution does not show these characteristics, it is unlikely to be normal. Nevertheless,

there are more robust methods for determining whether a distribution is normal: measures of symmetry (Munro, 2005).

Measures of symmetry. There are three tests that can measure symmetry; Pearson's Skewness Coefficient, Fisher's Measure of Skewness, and Fisher's Measure of Kurtosis. In a perfectly normal bell-curve all three of these measures should produce values of zero (Munro, 2005). As mentioned before, in a normal distribution the mean should equal the median. Pearson's Skewness Coefficient subtracts the mean from the median then divides the result by the standard deviation (Munro, 2005). Therefore, in a normal distribution, the subtraction component would result in zero. If the result of the equation is a positive number, then the data is positively skewed (Munro, 2005). If the result of the equation is a negative number, then the data is negatively skewed.

Conversely, Fisher's Measure of Skewness does not have quite so neat a formula but is incredibly sensitive (Munro, 2005). The basis for this statistic is deviations from the mean in the third power. Similarly, Fisher's Measure of Kurtosis is based on deviations from the mean in the fourth power (Munro, 2005). While skewness referred to a shift in the curve side to side, kurtosis refers to making the curve higher at the highest point or flatter at the highest point. Like Pearson's Skewness Coefficient, if the computed result is positive, the curve is more pointed than it should be. If the computed result is negative, then the curve is flatter than it should be (Munro, 2005). Measures of symmetry should be performed on Likert Scales due to the propensity of the Likert data to sometimes show skewness (Zou et al., 2010).

Significance and Confidence Levels

For the purposes of all hypothesis testing that follows, a *p*-value of less than or equal to 0.05 has been chosen. This means that getting a significant result from a statistical test might only occur five out of 100 times, making it highly unlikely the result was due to chance alone. Therefore, a stated hypothesis could be considered compatible with the study sample (Munro, 2005). This is consistent with data from the previous literature review section of this paper where this same *p*-value range was chosen for most of the studies analyzed.

Confidence intervals give an entire range of values that are not distinguishable from the observed sample (Munro, 2005). A 95% confidence interval was also chosen as this was deemed appropriate by Zou et al., (2010). Additionally, in the only study that reported confidence levels in the literature review, those authors also utilized a 95% confidence interval (Stewart et al., 2010). If previous data analysis reveals significant skewness to the data distribution, the following nonparametric measures could be employed assuming no recommendations for data transformations come forth from the statistician.

Nonparametric Statistical Tests

Unlike parametric statistical tests, there is no assumption of normalcy in nonparametric statistics. Additionally, nonparametric statistical tests can be implemented on ordinal level data, such as Likert scales (Munro, 2005; Zou et al., 2010). Nonparametric measures can be used to determine whether there is a difference between two groups on an outcome measure. It can also be used to determine whether there is a relationship between two variables. This could be valuable in comparing whether the post-test scores of the pharmacy, nursing, and medicine students were higher for HFS as opposed to LFS. It could

also be particularly useful for determining how each group of students varied from one another in their questionnaire results therefore necessitating multiple analyses of variance.

In a repeated measures cross-over design, the students themselves serve as their own control group and are measured more than once. They are essentially measured before the first test, after the first test, and then after the second test. This is particularly relevant when choosing a nonparametric measure for statistical analysis. Munro (2005) noted that the Wilcoxon matched-pairs signed rank test and the Friedman matched samples test would be particularly applicable here because of the repeated measures and paired nature of a sample that served as its' own control.

Instead of comparing means as in parametric measures, the scores for participants are transformed into "ranks" for these nonparametric tests. From there, analyses can contrast the mean ranks for each group instead (Munro, 2005). To do this, several other assumptions must be met: the data must be paired from the same participant, the paired data must come from the same population, each pair must be selected randomly and independently, and the data must be at least on an ordinal scale (Munro, 2005). The data from the current study would fit these assumptions.

To come at it from a more parametric perspective, the Wilcoxon matched-pairs signed rank test is the nonparametric equivalent to a paired *t*-test therefore making it appropriate to analyze the difference between pretest and posttest scores (Munro, 2005). Furthermore, the Friedman matched samples test would be the nonparametric equivalent to a repeated measures analysis of variance therefore making it appropriate for analysis of differences between the groups. Since there will be at least three pairwise comparisons, there is a greater likelihood of a type one error (Munro, 2005). To compensate for that a

Bonferroni correction will be applied. This simply involves the basic division of the *p*-value by the number of comparisons made. Therefore, with this correction, if three comparisons were being made, the significance level would be 0.167 (Munro, 2005).

Parametric Statistical Tests

Predictably, the parametric tests to be performed on this data are the parametric equivalents noted above. In a *t*-test, distributions and means are evaluated. However, if data are paired, you may see similar scores (Munro, 2005). To more accurately discern a significant result, the paired *t*-test, makes a correction for this similarity. The paired *t*-test was also the most frequently used statistical test noted in the literature review, with six studies utilizing it. Therefore, it would appear to be appropriate for use in this study, assuming the appropriate assumptions are met.

There are three essential assumptions to using *t*-tests. The first assumption concerns the assumption of independence (Munro, 2005). Under this assumption a participant can only add one data point to one of the two groups, thus constituting two mutually exclusive groups of participants. The second assumption underlies why assessing normalcy was so important: the dependent variable must have a normal distribution (Munro, 2005). The third assumption has to do with variance, hence why it was important to determine the standard deviation. For the two groups, the variance of the dependent variable should be similar. In addition to this, a continuous dependent variable is required (Munro, 2005). However, as Zou et al., (2010) noted, the dependent variable can be viewed on a continuum for Likert-based data, which would fill this requirement for parametric statistical analysis. Since parametric tests have greater power and flexibility, they would be preferred over

nonparametric measures (Munro, 2005). Nevertheless, to determine their appropriateness for this study would require performing the initial part of the data analysis plan.

Repeated measures analysis of variance is appropriate for the data for similar reasons discussed previously: the students serving as their own control group with repeated measures of the same variables after the LFS and HFS (Munro, 2005). Though this may seem to vary from the data presented in the literature review where six studies utilized ANOVA, the study design for this inquiry does appear to be better suited to repeated measures of analysis. Furthermore, Wellmon et al., (2017) did use a two-group by two-group repeated measures ANOVA with one pair denoting learning intervention versus control and the other pair denoting time pre-learning versus post-learning. Therefore, repeated measures analysis of variance was utilized by at least one of the analyzed articles. Essentially, repeated measures analysis of variance can help to tackle individual differences, which would become a problem if analyzed by a regular ANOVA (Munro, 2005).

Repeated measures analysis of variance could be particularly useful in determining differences between the three student groups surveyed, what differences are present in ITQ scores from the first simulation to the second simulation the students complete, and whether there is an interaction between simulation type and time (Munro, 2005). Like the *t*-test, the same data levels and assumptions apply. However, an additional assumption does apply. This is an important one since violating this assumption significantly breaks down the ability to use the test. The assumption of compound symmetry has two parts: measurements are the same across correlations and across these measurements, the variances are equal (Munro, 2005).

When developing a Likert Scale data analysis plan, consideration must be given to calculating an adequate sample size, conducting descriptive statistics, determining levels of significance, and determining whether non-parametric or parametric statistical tests will be used, before analyzing any statistical significance of any changes in the data.

Summary

Existing literature on data analysis plans for Likert Scale instruments used in interprofessional simulation education were analyzed. The review showed that when creating data plans for instruments that use Likert scales it is appropriate to consider both non-parametric (Wilcoxon, Kruskal-Wallis, Mann-Whitney *U*) and parametric (ANOVA, *t-test,* and HSD) measures. The branching choice between parametric and nonparametric measures necessitates the determination of normalcy in the distribution of the data. So, appropriate descriptive statistics for Likert Scales would include measures of central tendency, consideration of normal distribution, measures of variability, and measures of symmetry to determine whether to use parametric or non-parametric measures.

If the study design meets the required criteria, the data analysis plan for Likert Scales could include parametric statistics, however it is critical to calculate the required sample size because if the sample is too small it will not be able to detect significant differences. An adequate sample size should be confirmed by consultation with a specialist in the area of sample size determination. A data analysis plan for a Likert Scale should also consider how the data will be presented, such as tables, graphs or figures.

Appropriate descriptive statistics for an instrument that uses a Likert Scale could include measures of central tendency, measures of variability, and measures of symmetry. In particular, measures of symmetry should be performed due to the propensity of Likert

Scale data to show skewness. If data analysis reveals significant skewness to the data distribution, the Wilcoxon matched-pairs signed rank test and the Friedman matched samples test could be used.

References

Curran, V. R., Mugford, J. G., Law, R. M. T., & MacDonald, S. (2005). Influence of an interprofessional HIV/AIDS education program on role perception, attitudes and teamwork skills of undergraduate health sciences students. *Education for Health*, *18*(1), 32–44. https://doi.org/10.1080/13576280500042606

de Voest, M., Raguckas, S., Bambini, D., & Beel-Bates, C. (2013). Interprofessional teaching: An inter-university experience involving pharmacy and nursing students. *Currents in Pharmacy Teaching and Learning*, *5*(5), 450–457. https://doi.org/10.1016/j.cptl.2013.06.004

Garbee, D. 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 Perspectives*, *34*(5), 339–344. https://doi.org/10.5480/1536-5026-34.5.339

Gough, S., Hellaby, M., Jones, N., & MacKinnon, R. (2012). A review of undergraduate interprofessional simulation-based education (IPSE). *Collegian*, *19*(3), 153–170. https://doi.org/10.1016/j.colegn.2012.04.004

Hepburn, Tsukada & Fraser (1996). In E.L. Siegler, K. Hyer, T. Fulmer, M. Mezey, M. (1998). *Geriatric Interdisciplinary Team Training*. New York: Springer Publishing Company

Kardong-Edgren, S., Adamson, K. A., & Fitzgerald, C. (2010). A review of currently published evaluation instruments for human patient simulation. Clinical Simulation in Nursing, 6(1), e25-e35. https://doi.org/doi:10.1016/j.ecns.2009.08.004

Krueger, L., Ernstmeyer, K., & Kirking, E. (2017). Impact of interprofessional simulation on nursing students' attitudes toward teamwork and collaboration. *Journal of Nursing Education*, *56*(6), 321–327. https://doi.org/10.3928/01484834-20170518-02

Liaw, S. Y., Zhou, W. T., Lau, T. C., Siau, C., & Chan, S. W. (2014). An interprofessional communication training using simulation to enhance safe care for a deteriorating patient. *Nurse Education Today*, *34*(2), 259–264. https://doi.org/10.1016/j.nedt.2013.02.019

Lin, Y.-C., Chan, T.-F., Lai, C.-S., Chin, C.-C., Chou, F.-H., & Lin, H.-J. (2013). The impact of an interprofessional problem-based learning curriculum of clinical ethics on medical and nursing students' attitudes and ability of interprofessional collaboration: A pilot study. *Kaohsiung Journal of Medical Sciences*, *29*(9), 505–511. https://doi.org/10.1016/j.kjms.2013.05.006

Munro, B. H. (2005). Statistical methods for health care research (5th ed). Philadelphia: Lippincott Williams & Wilkins.

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. https://doi.org/10.1016/j.jamcollsurg.2013.09.006

Reising, D. L., Carr, D. E., Shea, R. A., & King, J. M. (2011). Comparison of communication outcomes in traditional versus simulation strategies in nursing and medical students. *Nursing Education Perspectives*, *32*(5), 323–327. https://doi.org/10.5480/1536-5026-32.5.323

Rossler, K. L., & Kimble, L. P. (2016). Capturing readiness to learn and collaboration as explored with an interprofessional simulation scenario: A mixed-methods research study. *Nurse Education Today, 36*, 348–353. https://doi.org/10.1016/j.nedt.2015.08.018

Sigalet, E., Donnon, T., & Grant, V. (2012). Undergraduate students' perceptions of and attitudes toward a simulation-based interprofessional curriculum: The KidSIM ATTITUDES Questionnaire. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare, 7*(6), 353–358. https://doi.org/10.1097/SIH.0b013e318264499e

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

Wellmon, R., Lefebvre, K. M., & Ferry, D. (2017). Effects of high-fidelity simulation on physical therapy and nursing students' attitudes toward interprofessional learning and collaboration. *Journal of Nursing Education*, *56*(8), 456–465. https://doi.org/10.3928/01484834-20170712-03

Zou, K. H., Carlsson, M. O., & Quinn, S. A. (2010). Beta-mapping and beta-regression for changes of ordinal-rating measurements on Likert scales: A comparison of the change scores among multiple treatment groups. *Statistics in Medicine*, *29*(24), 2486–2500. https://doi.org/10.1002/sim.4012.

Appendix B

Data Analysis Plan for the Interprofessional Teamwork Questionnaire

By Amy Warren

2018

Memorial University of Newfoundland

School of Nursing

The Interprofessional Teamwork Questionnaire (ITQ) consists of a Likert scale with an ordered set of discrete terms or statements from which participants are asked to choose the response that best describes their state or experience with the simulation (Appendix B1). The data analysis plan proposed for analysis of the Likert Scales used in the ITQ will focus on statistics relevant to a pretest, posttest, repeated measures cross-over research design and the nominal data collected by the ITQ. There are several different types of data analysis for Likert scale data, but the majority of studies using Likert scales to evaluate interprofessional teamwork have focused on parametric methods of data analysis (Curran, Mugford, Law, & MacDonald, 2005; de Voest, Raguckas, Bambini, & Beel-Bates, 2013; Garbee et al., 2013; Krueger, Ernstmeyer, & Kirking, 2017; Lin et al., 2013; Paige et al., 2014; Reising, Carr, Shea, & King, 2011; Sigalet, Donnon, & Grant, 2012; Stewart, Kennedy, & Cuene-Grandider, 2010). Only three studies used nonparametric measures including the Mann-Whitney *U* test, the Wilcoxon Signed Rank test, the Kruskal-Wallis test, and the Chi Squared Test (Liaw, Zhou, Lau, Siau, & Chan, 2014; Rossler & Kimble, 2016; Wellmon, Lefebvre, & Ferry, 2017).

Descriptive Statistics

A good place to start is to summarize the ITQ data with descriptive statistics, percentages, and measures of symmetry with Pearson's Coefficient. This is supported by the literature as all of the studies reported some form of descriptive statistics (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017). Such endeavors would help determine whether the data is normally distributed and whether it is possible to proceed with

parametric methods. This method is similar to the procedure that Rossler and Kimble (2016) utilized when they performed descriptive statistics and a normality assessment, which consequently resulted in choosing nonparametric statistical analyses for their quantitative data.

Half of the 12 studies analyzed used paired *t*-tests (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Paige et al., 2014; Sigalet et al., 2012; Wellmon et al., 2017). Since this test was so prevalent within literature that exhibits similar characteristics to the current research study, it is reasonable to assume it would be useful in the analysis of the ITQ. It could potentially be useful for comparing the satisfaction scores for HFS to LFS to determine whether the participants were more satisfied with one than the other. This would also align with the literature review as Wellmon et al., (2017) also utilized paired *t*-tests to analyze student satisfaction on the SCLS.

Comparison of Multiple Variables

Comparisons of multiple variables could occur through the use of repeated measures analysis of variance or ANOVA (Curran et al., 2005; Krueger et al., 2017; Lin et al., 2013; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017) therefore ANOVA may be the most appropriate statistical test for comparing multiple variables in the ITQ. The repeated measures version of ANOVA may be particularly useful since the study sample who will complete the ITQ will be serving as their own control group and will be subject to two different interventions: LFS and HFS (Munro, 2005).

A three factor 2 x 3 x 2 (simulation type, profession, and time) table could be used with the ITQ to evaluate knowledge of teamwork, knowledge of team roles, and teamwork perceptions. In such a scenario, simulation type would have two levels: LFS and HFS.

Furthermore, profession would have three levels: nursing, pharmacy, and medicine. On top of this, time would have two levels: pretest and posttest. If the data proves to be too skewed for such parametric measures, the nonparametric equivalents of the Friedman Matched Samples Test for the repeated measures analysis of variance and the Wilcoxon Matched-Pairs Signed Rank Test could be performed. A Bonferroni correction could also be used, similar to how Paige et al., (2014) used it on the paired *t*-tests they performed. Ultimately though, these data analysis recommendations could change after further consultation with a statistician takes place.

Creation of the SPSS Data Analysis File for the ITQ

The first step in the creation of the SPSS data analysis file for the ITQ was to create a new file with the dataset in SPSS. This involved identifying the name and type of variables to be analyzed and the names and labels for every variable, and values for the data measurements (Table B1 and Table B2).

Name Attribute

Once the new file was created, it was switched to "Variable View". Under the "Name" attribute, fields from the ITQ were entered. The unique identification number given to each participant was named "ID". To designate the academic discipline for participants a variable named "Profession" was included. The "Name" attribute is limited to 64 bytes, so statement names used the following conventions "**PreTest_**" to denote statements from the pretest ITQ, "**PostLfs_**" for statements from the post-low-fidelity simulation ITQ, "**PostHfs_**" for statements from the post-high-fidelity simulation ITQ, as well as the statement number to provide sufficient differentiation, with an underscore

preceding the number to allow for easier reading. The names and labels for each variable in the pretest are shown in Table 1 and the posttest in Table 2.

Table 1

Pretest Variables, Names And Labels

Name	**Label**	**Variable**
ID	ID number	Unique participant identifier
Profession	Profession	Nursing. Medicine, Pharmacy
PreTest_1	Pretest 1 Function effectively	Function effectively as a member of the interprofessional team
PreTest_2	Pretest 2 Importance of teamwork	Understand the importance of interprofessional teamwork
PreTest_3	Pretest 3 Role of nursing	Understand the role of nursing on an interprofessional team
PreTest_4	Pretest 4 Role of medicine	Understand the role of medicine on an interprofessional team
PreTest_5	Pretest 5 Role of pharmacy	Understand the role of pharmacy on an interprofessional team
PreTest_6	Pretest 6 Communicate effectively	Confident can communicate effectively with the interprofessional team
PreTest_7	Pretest 7 Collaborate effectively	Confident can collaborate effectively with the interprofessional team

Label Attribute

The label for "ID" was "ID Number." The variable denoting the respective academic disciplines was labeled as "Profession". For the ITQ statements, the format of "statement # - two-word or three-word descriptor" was employed. For example, "Pretest 1 – Function effectively" was the label for the pretest statement "I can *function effectively* as a member of the interprofessional team when caring for a patient experiencing anaphylaxis", and "Post-HFS 10 - Experience was well organized" was the label for the post-HFS test statement "The *experience was well organized*."

Table 2

Posttest Variables, Names And Labels

Name	**Label**[a]	**Variable**
PostLfs_1 PostHfs_1	Post-LFS 1 Function effectively	Function effectively as a member of the interprofessional team
PostLfs_2 PostHfs_2	Post-LFS 2 Importance of teamwork	Understand the importance of interprofessional teamwork
PostLfs_3 PostHfs_3	Post-LFS 3 Role of nursing	Understand the role of nursing on an interprofessional team
PostLfs_4 PostHfs_4	Post-LFS 4 Role of medicine	Understand the role of medicine on an interprofessional team
PostLfs_5 PostHfs_5	Post-LFS 5 Role of pharmacy	Understand the role of pharmacy on an interprofessional team
PostLfs_6 PostHfs_6	Post-LFS 6 Communicate effectively	Confident can communicate effectively with the interprofessional team
PostLfs_7 PostHfs_7	Post-LFS 7 Collaborate effectively	Confident can collaborate effectively with the interprofessional team
PostLfs_8 PostHfs_8	Post-LFS 8 Learning objectives were clear	The learning objectives were clear.
PostLfs_9 PostHfs_9	Post-LFS 9 Workload was fair	The workload was fair.
PostLfs_10 PostHfs_10	Post-LFS 10 Experience was well organized	The experience was well organized
PostLfs_11 PostHfs_11	Post-LFS 11 Pre-briefing useful	Pre-Briefing was useful in facilitating my learning
PostLfs_12 PostHfs_12	Post-LFS 12 Orientation useful	Orientation to high fidelity simulator was useful in facilitating my learning
PostLfs_13 PostHfs_13	Post-LFS 13 Debriefing useful	Debriefing was useful in facilitating my learning
PostLfs_14 PostHfs_14	Post-LFS 14 Recommend experience	I would recommend this experience to others.
PostLfs_15 PostHfs_15	Post-LFS 15 Meaningful experience	Overall this was a meaningful experience.

[a]Labels shown for Post-LFS only, Post-HFS labels are similar with "H" replacing "L"

The largest difficulties for this format were posttest statements 11 to 13, as they were organized differently with the statement "The following activities were useful in facilitating my learning:" above the statements "Pre-briefing", "Orientation to high fidelity simulator (if applicable)", and "Debriefing." Theoretically, labeling these "Post-LFS 11 - Pre-briefing useful", "Post-LFS 12 - Orientation useful", and "Post-LFS 13 - Debriefing

useful" provided sufficient clarity as to which statement the label was referring to. Variable labels are shown in Table B1 and Table B2.

Values Attribute

The "Value = Label" entries for the "Profession" variable are as follows: "1 = Nursing"; "2 = Medicine"; and "3 = Pharmacy". The following "Value = Label" entries were utilized for the ITQ statement variables: "0 = Not Applicable", "1 = Strongly Disagree", "2 = Disagree", "3 = Neutral", "4 = Agree", and "5 = Strongly Agree". The value "0" was chosen because it is a single digit, which allows for faster entry, and it is somewhat visually distinct from the other values which will make it easier to recognize in entered data.

Missing Attribute

After selecting the "Discrete Missing Values" button, the value "0" was entered as this is a user entered value that should not be used in any calculations.

Columns Attribute

All values were left at their default.

Align Attribute

The alignment for "ID" was left at the default value of "Right". For all other variables the value was changed to "Center" to aid in readability during data entry.

Measure Attribute

The entry for "ID" was set to "Nominal". However, "ID" will not be used directly in calculations therefore this entry could also be "Ordinal" with no impact on the results. The "Profession" variable was set to "Nominal" as this variable contains discrete unranked

categories. For the statements ("PreTest_1" to "PostHfs_15") the "Measure" entries were changed to "Ordinal" as the entries consist of categorically ranked values.

Role Attribute

All values were left at their default. After saving, the database was ready for data to be entered in the "Data View" tab.

Interprofessional Teamwork Questionnaire

Research Questions

The ITQ will be used answer the research questions as follows: (1) Are participants more satisfied with high-fidelity simulation (HFS) as compared to low-fidelity simulation (LFS)? (2) Does participation in HFS result in a higher level of knowledge of teamwork and team roles, as compared to participation in LFS? (3) Does participation in HFS result in a higher level of teamwork, collaboration and communication behaviors, as compared to participation in LFS? and, (4) Does participation in HFS result in more improved attitudes towards teamwork, as compared to participation in LFS?

Study Design

The research questions will be answered using a within subjects, pretest, posttest, repeated measures design. The convenience sample of nursing, pharmacy, and medicine students will be randomized using a coin flip procedure so that they can be appropriately grouped. Each group will include nursing, medicine, and pharmacy students. Each group will participate in both simulation types therefore serving as their own control group. The

pretest ITQ will be completed by participants prior to the first simulation and the posttest ITQ will be completed after the first simulation and again after the second simulation.

Sample

Nursing, pharmacy, and medicine undergraduate students from an Eastern Canadian university will participate in this study. Each student group will be enrolled full time in either nursing, medicine or pharmacy undergraduate education programs. For the sample to be adequate, it must be large enough to detect significant differences (Munro, 2005). Several factors need to be considered when determining an appropriate sample size. These include significance level, power, and effect size. The significance level correlates with reducing the probability of a type I error where a false positive occurs. This value will be discussed in later sections of this paper.

Conversely, a higher power level means there is a greater chance of avoiding a type II error where a false negative occurs. An adequate power level of 80% is generally accepted (Munro, 2005). The literature review conducted for this practicum showed that sample size could vary from 30 to 300 subjects and include several health care professionals. The study from the literature review that had the most similar sample to the current study would be Curran et al., (2005). In that case, 45 nursing, 62 medicine, and 26 pharmacy students participated in the study. Therefore, with so many factors to take into account, adequate sample size for the study will need to be determined.

Level of Measurement

Ordinal level data collected within the ITQ include the seven items categorized into six-point Likert scales from “Strongly agree” to “Strongly disagree” in the pretest. Those statements focus on areas of individual functionality in an interprofessional team, the

significance of interprofessional collaboration in the simulation, comprehension of the role of each profession involved, individual communication confidence in an interprofessional team, and confidence in collaboration for care planning. The posttest also includes the seven initial statements but also collects satisfaction data through an additional nine items rated on the same scale. Items include the level of clarity of the learning objectives, the fairness of the workload, the experience was organized, utility of each component in the simulation experience, (pre-briefing, orientation to the simulator and debriefing), whether they would recommend the scenario to others, and whether the experience was meaningful overall. One of the limitations of the ITQ is that it is a purposeful questionnaire and reliability and validity have not been established.

Level of Significance

To reduce the potential for a type one error, a significance level of less than 0.05 was chosen for all hypothesis testing. This is consistent with data from the previous literature review section of this paper where this same *p*-value range was chosen for most of the studies analyzed (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Sigalet et al., 2012; Wellmon et al., 2017). Additionally, a confidence level of 95% was chosen as this was deemed appropriate by Zou, Carlsson, & Quinn, (2010). Additionally, in the only study that reported confidence levels in the literature review, those authors also utilized a 95% confidence interval (Stewart et al., 2010). As discussed earlier, an adequate power level of 80% is generally accepted and will be recommended (Munro, 2005).

Methods of Summarizing Study Data: Descriptive Statistics

All statistical analysis will take place using SPSS version 23. Descriptive statistics will be performed for ID and Profession using percentages, measures of central tendency, measures of variability, and measures of symmetry. Measures of central tendency will include the calculation of the mean, median, and mode. Measures of variability will include the calculation of standard deviations. Furthermore, measures of symmetry will include calculating Pearson's Coefficient to determine whether the data exhibits a normal distribution. This aligns with the literature review as all of the twelve studies analyzed performed descriptive statistics in some manner or another (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Sigalet et al., 2012; Sigalet et al., 2012; Wellmon et al., 2017).

Parametric measures of statistical analysis may be utilized if the data exhibits a normal distribution: repeated measures analysis of variance and paired *t*-test. However, the SPSS file and data analysis plan may require some refinement in order to perform repeated measures analysis of variance. These tests exhibit a greater power, however due to the bounded nature of Likert data, they are not always possible (Munro, 2005; Zou et al., 2010). This preference toward parametric measures also aligns with the results of the literature review as nine of the studies reviewed utilized parametric statistical tests (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Sigalet et al., 2012; Stewart et al., 2010). Furthermore, half used the paired *t*-test (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Paige et al., 2014; Sigalet et al., 2012; Wellmon et al., 2017). Additionally, six

studies utilized ANOVA (Curran et al., 2005; Krueger et al., 2017; Lin et al., 2013; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017) with the study by Wellmon et al., (2017) utilizing a repeated measures analysis of variance, as prescribed in the next section of this data analysis plan.

Overall though, this prevalence of parametric measures also underscores the importance of descriptive statistics and measures of symmetry. These measures will help determine whether parametric measures or nonparametric measures should be utilized (Munro, 2005). Rossler and Kimble (2016) support this approach with Likert Scales, to conduct descriptive statistics and normality assessment prior to choosing nonparametric statistical measures for their data analysis.

Methods for Interpreting Study Data: Inferential Statistics

The Friedman Matched Samples Test, the nonparametric equivalent to a repeated measures analysis of variance, will possibly be performed on a three factor 2 x 3 x 2 (simulation type, profession, and time) table to evaluate knowledge of teamwork, knowledge of team roles, and teamwork perceptions. Simulation type would have two levels: LFS and HFS. Profession would have three levels: nursing, pharmacy, and medicine. Furthermore, time also would have two levels: pretest and posttest. This analysis will attempt to address the second, third, and fourth study questions. To analyze the first study question, a Wilcoxon Matched-Pairs Signed Rank Test, the nonparametric equivalent to a paired *t*-test, will be performed on posttest data for HFS and for LFS. These matched pairs would compare posttest satisfaction data from each simulation type to determine which type had greater levels of satisfaction from the students.

Creation of the SPSS Data Analysis File for the ITQ

The ITQ consists of a Likert scale with an ordered set of discrete terms or statements from which participants are asked to choose the response that best describes their state or experience with the simulation. The data analysis plan proposed for analysis of the Likert Scales used in the Interprofessional Teamwork Questionnaire (ITQ) will focus on statistics relevant to a pretest, posttest, repeated measures cross-over research design and the nominal data collected by the ITQ. This literature review has focused on several different types of data analysis for Likert scale data, with the vast majority focused on parametric methods of data analysis (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Sigalet et al., 2012; Stewart et al., 2010). Only three used nonparametric measures like the Mann-Whitney *U* test, the Wilcoxon Signed Rank test, the Kruskal-Wallis test, and the Chi Squared Test (Liaw et al., 2014; Rossler & Kimble, 2016; Wellmon et al., 2017).

A good place to start is to summarize the data with descriptive statistics, percentages, and measures of symmetry with Pearson's Coefficient. This is supported by the literature reviewed as all of the studies reported some form of descriptive statistics (Curran et al., 2005; de Voest et al., 2013; Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Lin et al., 2013; Paige et al., 2014; Reising et al., 2011; Rossler & Kimble, 2016; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017). Such endeavors will help to determine whether the data is normally distributed and whether it is possible to proceed with parametric methods for the statistical analyses of the quantitative data from the ITQ.

Half of the 12 studies analyzed used paired *t*-tests (Garbee et al., 2013; Krueger et al., 2017; Liaw et al., 2014; Paige et al., 2014; Sigalet et al., 2012; Wellmon et al., 2017).

Since this test was so prevalent within literature that exhibit similar characteristics to the current research study, it is reasonable to assume it would be useful in the analysis of the ITQ. It could potentially be useful for comparing satisfaction scores for HFS to LFS to determine whether the participants were more satisfied with one than the other. This would also align with the literature review as Wellmon et al., (2017) also utilized paired *t*-tests to analyze student satisfaction on the SCLS.

Comparisons of multiple variables could occur through the use of repeated measures analysis of variance. In this literature review six studies used ANOVA (Curran et al., 2005; Krueger et al., 2017; Lin et al., 2013; Sigalet et al., 2012; Stewart et al., 2010; Wellmon et al., 2017) therefore it too may be the most appropriate statistical test for comparing multiple variables. The repeated measures version of ANOVA may be particularly useful since the study sample who will complete the ITQ will be serving as their own control group and will be subject to two different interventions: LFS and HFS (Munro, 2005). Furthermore, a repeated measures analysis of variance was used by one of the studies explored in this literature review. However, Wellmon et al., (2017) only a 2 x 2 table with time and the control was utilized. In the current research study, such a table would need to be expanded to include the three profession types.

A three factor 2 x 3 x 2 (simulation type, profession, and time) table could evaluate knowledge of teamwork, knowledge of team roles, and teamwork perceptions. In such a scenario, simulation type would have two levels: LFS and HFS. Furthermore, profession would have three levels: nursing, pharmacy, and medicine. On top of this, time would have two levels: pretest and posttest. If the data proves to be too skewed for such parametric measures, the nonparametric equivalents of the Friedman Matched Samples Test for the

repeated measures analysis of variance and the Wilcoxon Matched-Pairs Signed Rank Test could be performed. A Bonferroni correction could also be used, similar to how Paige et al., (2014) used it on the paired *t*-tests they performed. Ultimately though, these data analysis recommendations could change after further consultation with a statistician takes place.

Summary of Data Analysis Plan for ITQ

The research study will utilize a within subjects, pretest, posttest, repeated measures design to compare the effects of HFS and LFS on IPE. The convenience sample will consist of medical, nursing, and pharmacy undergraduate students. The ITQ will be administered to these participants prior to the first simulation, after it, and then again after the second simulation. The ITQ consists of a Likert scale with an ordered set of discrete terms or statements from which participants are asked to choose the response that best describes their state or experience with the simulation. Therefore, the ITQ will be used to collect nominal and ordinal level data.

To analyze this data, several parameters will be in place. The confidence interval will be set to 95% and a significance level of less than 0.05 will be applied. All descriptive and inferential statistics will be performed using SPSS version 23. Initial descriptive statistics will include percentages, measures of central tendency, and measures of symmetry with Pearson's Coefficient. If this analysis reveals a normal distribution, further inferential statistics will be performed using parametric tests: paired *t*-tests and repeated measures analysis of variance. However, it is likely the data will not be normally distributed, and the Friedman Matched Samples Test and Wilcoxon Matched-Pairs Signed Rank Test will be performed instead.

In either case, the analysis will occur on a three factor 2 x 3 x 2 (simulation type, profession, and time) table to evaluate knowledge of teamwork, knowledge of team roles, and teamwork perceptions. Simulation type would have two levels: LFS and HFS. Profession would have three levels: nursing, pharmacy, and medicine. Furthermore, time also would have two levels: pretest and posttest. This analysis will attempt to address the second, third, and fourth study questions. To analyze the first study question, a Wilcoxon Matched-Pairs Signed Rank Test, the nonparametric equivalent to a paired *t*-test, will be performed on posttest data for HFS and for LFS. These matched pairs would compare posttest satisfaction data from each simulation type to determine which type had greater levels of satisfaction from the students.

References

Curran, V. R., Mugford, J. G., Law, R. M. T., & MacDonald, S. (2005). Influence of an interprofessional HIV/AIDS education program on role perception, attitudes and teamwork skills of undergraduate health sciences students. *Education for Health*, *18*(1), 32–44. https://doi.org/10.1080/13576280500042606

de Voest, M., Raguckas, S., Bambini, D., & Beel-Bates, C. (2013). Interprofessional teaching: An inter-university experience involving pharmacy and nursing students. *Currents in Pharmacy Teaching and Learning*, *5*(5), 450–457. https://doi.org/10.1016/j.cptl.2013.06.004

Garbee, D. 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 Perspectives*, *34*(5), 339–344. https://doi.org/10.5480/1536-5026-34.5.339

Gough, S., Hellaby, M., Jones, N., & MacKinnon, R. (2012). A review of undergraduate interprofessional simulation-based education (IPSE). *Collegian*, *19*(3), 153–170. https://doi.org/10.1016/j.colegn.2012.04.004

Krueger, L., Ernstmeyer, K., & Kirking, E. (2017). Impact of interprofessional simulation on nursing students' attitudes toward teamwork and collaboration. *Journal of Nursing Education*, *56*(6), 321–327. https://doi.org/10.3928/01484834-20170518-02

Liaw, S. Y., Zhou, W. T., Lau, T. C., Siau, C., & Chan, S. W. (2014). An interprofessional communication training using simulation to enhance safe care for a deteriorating

patient. *Nurse Education Today*, *34*(2), 259–264. https://doi.org/10.1016/j.nedt.2013.02.019

Lin, Y.-C., Chan, T.-F., Lai, C.-S., Chin, C.-C., Chou, F.-H., & Lin, H.-J. (2013). The impact of an interprofessional problem-based learning curriculum of clinical ethics on medical and nursing students' attitudes and ability of interprofessional collaboration: A pilot study. *Kaohsiung Journal of Medical Sciences*, *29*(9), 505–511. https://doi.org/10.1016/j.kjms.2013.05.006

MacDonald, S., Manuel, A., Dubrowski, A., Bandrauk, N., Law, R., & Curran, V. (2017) *Measuring the effectiveness of high fidelity simulation in interprofessional education to foster teamwork among undergraduate nursing, medicine and pharmacy students.* Unpublished manuscript, Memorial University of Newfoundland, St. John's, Canada

Munro, B. H. (2005). Statistical methods for health care research (5th ed). Philadelphia: Lippincott Williams & Wilkins.

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. https://doi.org/10.1016/j.jamcollsurg.2013.09.006

Reising, D. L., Carr, D. E., Shea, R. A., & King, J. M. (2011). Comparison of communication outcomes in traditional versus simulation strategies in nursing and medical students. *Nursing Education Perspectives*, *32*(5), 323–327. https://doi.org/10.5480/1536-5026-32.5.323

Rossler, K. L., & Kimble, L. P. (2016). Capturing readiness to learn and collaboration as explored with an interprofessional simulation scenario: A mixed-methods research study. *Nurse Education Today, 36*, 348–353. https://doi.org/10.1016/j.nedt.2015.08.018

Sigalet, E., Donnon, T., & Grant, V. (2012). Undergraduate students' perceptions of and attitudes toward a simulation-based interprofessional curriculum: The KidSIM ATTITUDES Questionnaire. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare, 7*(6), 353–358. https://doi.org/10.1097/SIH.0b013e318264499e

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

Wellmon, R., Lefebvre, K. M., & Ferry, D. (2017). Effects of high-fidelity simulation on physical therapy and nursing students' attitudes toward interprofessional learning and collaboration. *Journal of Nursing Education*, *56*(8), 456–465. https://doi.org/10.3928/01484834-20170712-03

Zou, K. H., Carlsson, M. O., & Quinn, S. A. (2010). Beta-mapping and beta-regression for changes of ordinal-rating measurements on Likert scales: A comparison of the change scores among multiple treatment groups. *Statistics in Medicine*, *29*(24), 2486–2500. https://doi.org/10.1002/sim.4012.

Appendix B1

Interprofessional Teamwork Questionnaire

Pre Test

Name: _____ Date: _____

Use the following scale to rate your agreement with each of the statements: SD= strongly disagree: D= disagree: N= neutral: A=agree: SA= strongly agree: NA= not applicable

Statement	SD	D	N	A	SA	NA
I can function effectively as a member of the interprofessional team when caring for a patient experiencing anaphylactic shock.						
I understand the importance of interprofessional teamwork when caring for a patient experiencing anaphylactic shock.						
I understand the role of nursing on an interprofessional team caring for a patient experiencing anaphylactic shock.						
I understand the role of medicine on an interprofessional team caring for a patient experiencing anaphylactic shock.						
I understand the role of pharmacy on an interprofessional team caring for a patient experiencing anaphylactic shock.						
I am confident I can communicate effectively with the interprofessional team to develop a plan of care for a patient experiencing anaphylactic shock.						
I am confident I can collaborate effectively with the interprofessional team to develop a plan of care for a patient experiencing anaphylactic shock.						

Comments

ID # _____

Interprofessional Teamwork Questionnaire
Post Test

Name: _____ Date: _____

Use the following scale to rate your agreement with each of the statements: SD= strongly disagree: D= disagree: N= neutral: A=agree: SA= strongly agree: NA= not applicable

Statement	SD	D	N	A	SA	NA
I can function effectively as a member of the interprofessional team when caring for a patient experiencing anaphylactic shock.						
I understand the importance of interprofessional teamwork when caring for a patient experiencing anaphylactic shock.						
I understand the role of nursing on an interprofessional team caring for a patient experiencing anaphylactic shock.						
I understand the role of medicine on an interprofessional team caring for a patient experiencing anaphylactic shock.						
I understand the role of pharmacy on an interprofessional team caring for a patient experiencing anaphylactic shock.						
I am confident I can communicate effectively with the interprofessional team to develop a plan of care for a patient experiencing anaphylactic shock.						
I am confident I can collaborate effectively with the interprofessional team to develop a plan of care for a patient experiencing anaphylactic shock.						

Statement	SD	D	N	A	SA	NA
The learning objectives were clear.						
The workload was fair.						
The experience was well organized						
The following activities were useful in facilitating my learning:						
Pre-briefing						
Orientation to high fidelity simulator (if applicable)						
Debriefing						
I would recommend this experience to others.						
Overall this was a meaningful experience.						