Evaluation of Interrater Reliability for Coding of Types of Gazes in Nurse-Patient Dyads

Alexandria Martz

The Ohio State University

Presented in Partial Fulfillment of the Requirements for the
Bachelor of Science in Nursing With Honors

Honors Thesis Committee:
Dr. Celia E. Wills, Advisor
Dr. Mary E. Happ

Abstract

Introduction and Purpose

The purpose of this Honors thesis is to describe and evaluate processes of interrater reliability assessment based on 13 video recordings of ICU nurse-patient dyads that were collected as part of a prior study (Happ, et al., 2011). Interrater reliability is defined as the amount of agreement between different raters when assessing the same objects in the same data using the same scale, classification, instrument, or procedure (Burns, 2014). Interrater reliability is important to establish when two or more researchers are collecting data so the information is obtained consistently, as well as to verify the reliability of a codebook or tool for measurement as part of quality assurance.

Background

Research has been conducted regarding different forms of nonverbal communication that occur between nurses and ventilated patients in the ICU. Additional research is needed to further understand the types of nonverbal communication that occur between nurses and patients in relation to patient-oriented outcomes. One research team used video recordings of nurse-patient encounters on an oncology unit to identify four types of touch a nurse performs during patient care: comforting, connecting, working, and orienting touch (Bottorff and Morse, 1993). This conceptualization was adapted as the theoretical framework for this thesis research on coding the types of visual gazes between nurse and patient.

Methods

A gaze typology was initially developed during a secondary analysis of video recordings of nurse-patient interaction from the SPEACS study (Crighton, Swigart, Happ, 2008), which then led to the creation of a codebook for coding nurse-patient visual gazes, requiring the assessment for interrater reliability (Frier et al., 2016). After four types of visual gazes were coded from 13 video recordings of interactions between mechanically ventilated patients and their nurses in the Medical (MICU) and Cardiothoracic Intensive Care Units (CTICU), interrater reliability was established using raw percentage of agreement. Lombard et al. (2010) discussed the overall steps for assessing interrater reliability and these steps were followed during establishment of interrater reliability for the codebook.

Results

Overall, the coders agreed in the coding of the four visual gazes throughout the 13 video recordings 70% of the time, but with a substantial range of agreement from 58% to 90%. With a goal of achieving at least 75% agreement based on an established gold standard for interrater reliability (Haidet et al., 2009), the overall percentage was lower than the target goal. Only one gaze, “relating gaze,” achieved a percentage agreement (90%) of over 75%. Sources of disagreement during coding arose from formative clarifications of the codebook definitions, but evolved from 70% agreement for the first video recording coded to 100% agreement for the final video recoding, which demonstrated improved agreement for coding as the codebook was further defined.

Discussion
As definitions were clarified during the coding process, fewer disagreements on coding of the gazes were found over the 13 video recordings. Ways to improve the raw percentage agreement are discussed with the best method being to include more training videos to better define the codebook prior to the official coding of the videos. Further research needs to use the Kappa coefficient method of establishing interrater reliability, which adjusts for agreement occurring by chance.

Introduction and Purpose

Interrater reliability is defined as the amount of agreement between independent raters for assessing and rating the same objects in the same data using the same scale, classification, instrument, or procedure (Burns 2014). Interrater reliability is important to establish when two or more individuals are collecting or coding data, so that data are obtained or coded consistently for multiple purposes, including verifying the reliability of a codebook or tool for measurement. Procedures for establishing interrater reliability are a pre-requisite step to achieve control of error variability in measurement as a basis for valid measurement of phenomena of interest.

Interrater reliability is important for both research and for quality assurance and improvement initiatives in clinical nursing practice (Burns 2014). For example, establishing interrater reliability for making ratings is a standard procedure in observational research studies, including for coding data that are inherently subjective. Adequate interrater reliability helps to assure that the ratings produced by a given individual can be independently confirmed by at least one other individual. Likewise, in clinical nursing practice, there are a number of activities that require reliable, independently reproducible ratings to be made. For example, assessments of the extent of pain, sedation, risk for falling, skin assessment, and so forth require that individual nurses have substantial agreement in their ratings for the same patient at a given point in time, in order to provide appropriate and consistent nursing care. When there is substantial variability in the ratings among individuals making ratings, this may mark particular areas for quality improvement to establish standardized procedures for interrater reliability and additional training to improve interrater reliability.

The purpose of this Honors thesis is to describe and evaluate processes of interrater reliability assessment based on 13 video recordings of ICU nurse-patient dyads that were collected as part of a prior study (Happ, et al., 2011). The video recordings were coded for four types of nurse-patient visual gazes that occurred during nursing care in the Medical (MICU) or Cardiothoracic Intensive Care Units (CTICU). The four types of gazes that were coded included *listening, assessing, technical doing*, and *relating*. The process and results for assessing interrater reliability for coding gazes within individual video recordings and across video recordings are described. The advantages, disadvantages, and implications of the specific approaches that were used for interrater reliability assessment in this project will be analyzed in relation to implications for quality assurance and improvement in research and clinical nursing practice.

Background and Review of Literature

For the purpose of this Honors Thesis, an initial review of the literature was conducted to analyze nurse and mechanically ventilated patient interactions, which allowed for an understanding of the information in the codebook regarding the four types of nurse-patient gazes when completing data coding and interrater reliability assessment of the coding. An additional literature search was then conducted to examine how interrater reliability is performed and to identify the most appropriate methods for particular purposes.

Communication Training in Healthcare

The ability to appropriately communicate with both verbal and non-verbal patients is a skill that is mastered from experience, persistence, and empathy. Various research studies have explored this relationship by analyzing how the concept is implemented into education, experiences for health care professionals, and how people understand the techniques used to communicate. In 2008 a review of literature was completed by Levy-Storms to determine what needed to be implemented into training programs for nursing aides. The review shows that curriculum should include an emphasis on verbal and non-verbal communication behaviors, instructions led by a professional outside of the facility, have a focus on psychosocial features and direct care responsibilities, use staff to provide feedback to improve quality of care, and to work in multiple long-term care facilities. This research helped to clarify communication between healthcare workers and patients, as well as provide guidelines. Another study focused on defining the difference between nurse-patient interactions and nurse-patient relationships, which placed a focus on how the nurse facilitated the interactions or relationships without discussion of the patient's participation in the communication process (Morse, 1997). Additional research is needed to go beyond a focus on the nurse only to include the patient's responses to the communication process as a key influence on the nurse-patient interaction. This research would serve as a basis for interventions to create constructive communication between nurse and patient. When studying the relationships of nurses and mechanically ventilated patients in the ICU, results of interviews showed that the preferred method of communication for patients was through body language and touch, while nurses preferred lip reading (Wojnicki-Johansson, 2001). This study reported that 23% of patients stated their nurses were unable to understand to a high degree what their needs were. This study supports the conclusion that there is a need for improved training in communication approaches to support nurses to better understand the needs of their patients.

Communication Used When Conversing with Mechanically Ventilated Patients

There are multiple types of non-verbal communication techniques used by nurses and intubated patients to communicate in relation to patients' needs, wants, and emotions. Nonverbal communication includes gestures, eye blinks, writing, touching, pointing, and assisted speech (Happ, 2001). One study found the most common communication methods used by nonvocal patients in the ICU were head nods, mouthing words, gestures,

and writing. The purpose of these communication interactions were most commonly to discuss pain, physical needs, and feelings (Happ, et al., 2011). These studies have been important in developing methods in which to communicate with ventilated patients in an easy and comfortable way for them. However, there is significant room for research to improve the effectiveness of non-verbal communication techniques. The need for additional research is supported by data that show that despite their best efforts, only about 73% of patients stated their nurses were able to understand what they needed (Wojnicki-Johansson, 2001). An important related finding is that patients reported that the presence of a nurse or relative provided a feeling of security and wellbeing (Karlsson, Lindahl, & Bergbom, 2012), despite the existing communication challenges for which improved communication approaches are needed. The implication of this finding is that interventions for supporting effective communication could be extended beyond the nurse-patient dyad to others, including family members and other types of caregivers.

Importance of Nonverbal Communication and Human Connection

Previous studies have examined the importance of nonverbal communication between patients and nurses and how these interactions affect the patient's emotional state. One study explained the difference between nurse-patient interactions and nurse-patient relationships (Morse, 1997). Interactions were described to be short events that concentrate on what can be seen while nurse-patient relationships are examining details and how they change over time to see the entire nurse-patient communication process. This study placed a focus on how the nurse facilitates the interactions or develops the relationship without perspective on how the patient influences these relations. While providing important information about how the nurse contributes to the communication process, this study did not examine the patient contribution to the communication process.

Another study looked at the various types of non-verbal communication between nurse and patient, which included eye contact, facial movements, body posture, touching, and personal space (Davies, 1994). Davies discussed the importance of non-verbal communication in relation to previous studies that found many patients were dissatisfied with their care due to poor communication. The article explains the importance of a nurse's non-verbal behavior and how it impacts the patient's perception of the care given.

Another research study analyzed mechanically ventilated patients' communication based on video recordings (Karlsson, Lindahl, & Bergbom, 2012). The patient's experiences during ventilation, how they communicated with the interviewer, and what was communicated through the patient's facial expressions were examined in each of the interviews from the video recordings and then compared to each other for significant findings of pain and fear during mechanical ventilation treatment, frustration with not being able to communicate their thoughts, questions, and feelings appropriately to their health care providers, and limitations in their abilities to use nonverbal communication due to the equipment they were attached to or other conditions preventing or limiting the use of their hands and face. These aspects are important to study because the findings demonstrate a continuation of reports similar to previous complaints of pain

and discomfort with communication and lack of being able to communicate their needs. This demonstrates the need for further research on improving communication techniques between nurses and patients in order to meet patients' needs and achieve better patient outcomes.

Based on a metasynthesis conducted by Carroll (2004), fives themes across 12 studies were found for ventilated patients' perceptions of being understood during their care (Carroll, 2004). These themes were not being understood, loss of control, negative emotions, individualized care, and caring presence leading to the conclusion that the interaction between nurse and non-vocal, ventilated patient's is not adequate. Additional research is needed to extend existing understanding of types of nonverbal communication between nurses and patients and how these forms of communication help or hinder the healing process for the patient.

In order to understand how various forms of communication achieve the purpose of appropriate communication between nurse and patient, observational studies must be completed. Studies such as these require reliable methods to code the different forms of communication into clear categories, so as to understand what form is communicating the necessary information best and with the most ease for the patient. The coding methods for these studies must be examined for interrater reliability, which indexes the extent to which observations between multiple raters are in agreement.

Approaches to Interrater Reliability Assessment

Lombard et al. discussed the overall steps for assessing interrater reliability as follows (Lombard, Snyder-Duch, & Bracken, 2010):

1. Select one or more appropriate indices
2. Obtain the necessary tools to calculate the index or indices selected
3. Select an appropriate minimum acceptable level of reliability for the index or indices to be used
4. Assess reliability informally during coder training
5. Assess reliability formally in pilot test
6. Assess reliability formally during coding of the full sample
7. Select and follow an appropriate procedure for incorporating the coding of the reliability sample into the coding of the full sample
8. Report intercoder reliability in a careful, clear, and detailed manner in all research reports

For steps (4), (5), and (6), two common methods of establishing interrater reliability in clinical research and clinical practice are raw percentage agreement and the Kappa statistic (Burns, 2014). Raw percentage agreement is likely to be used at step (4) as a crude indicator of extent of agreement among raters. However, for formal reliability assessment of ratings that involve subjective judgments of behaviors or diagnostic judgments, reliance on raw percentage agreement alone is limited as independent raters can be in agreement some extent of the time based on random chance. Interrater reliability assessment methods such as the Kappa coefficient have been developed in

order to correct for the extent of agreement in ratings that would be expected to occur by random chance.

For example, previous literature in the ICU context explained the development of a tool to measure communication interaction behaviors between nurses and mechanically ventilated patients in the intensive care unit (Nilsen, Happ, Donovan, Barnato, Hoffman, & Sereika, 2014). The result was a 29-item instrument with validity and interrater reliability that had a 73-100% agreement for nursing behaviors and 68-100% for patient behaviors. Interrater reliability assessment methods such as the Kappa coefficient typically result in somewhat lower percentage agreement as compared to raw percentage agreement because of the correction of the percentage coefficient for the extent of agreement that could be expected to occur by random chance alone. Raw percentage is calculated where the total number of possible items to be agreed on divides the number of times the data collectors agree (Burns 2014). The kappa statistic, however, calculates the percentage of agreement by adding the number of times the data collectors agree on what is being measured from the data divided by the total number of data items and has been adapted to calculate for when the two coders agree by chance.

Summary, Purpose, and Theoretical Framework

Current research has begun to explore the complex relationship between a nurse and the ventilated patient and the methods in which they communicate. Research has shown the importance for quality communication methods, so that patients may be understood and have their needs met. Research has been conducted on what forms of communication patients prefer and how these forms can be implemented into education, but with over a quarter of patients reporting a disconnect in understanding of their needs by their nurse, investigations need to continue to generate enhanced interventions to support better patient outcomes. This research area is premised on appropriate use of interrater reliability assessment methods as a basis for valid measurement of communication behaviors and associated outcomes.

Therefore, the purpose of this Honors thesis is to describe and evaluate processes of interrater reliability assessment using 13 video recordings of ICU nurse-patient dyads collected as part of a prior study (Happ, et al., 2011) that were coded for four types of gazes (relating, assessing, technical doing, listening) during nursing care in the Medical (MICU) and Cardiothoracic Intensive Care Units (CTICU). The social theoretical framework of symbolic interactionism (Blumer, 1969) was the conceptual basis for the analysis of video recording communication, and the steps described by Lombard et al. constituted the overall framework for interrater reliability assessment of the coding of gazes for the video recorded nurse-patient dyads (Lombard, Snyder-Duch, & Bracken, 2010). Four types of touches a nurse performs during patient care have been identified in prior research: comforting, connecting, working, and orienting (Bottorff and Morse, 1993). From this study the theoretical basis was formed to research the types of gazes that relate to the types of touches between nurse and patient during patient care tasks, to more fully describe and codify the communication process.

These types of gazes identified were then defined in a codebook for analyzing a set of 13-video recorded interactions between nurse-patient dyads.

Methods

Sample

The gazes were identified from 13 video recordings that included 10 nurses from a medical ICU and cardiothoracic ICU randomly selected to participate in the usual care cohort of the Study of Patient-Nurse Effectiveness with Assisted Communication Strategies (SPEACS) (Happ, et al., 2014). The 13 patients were adults over the age of 21 years, awake, able to communicate by head nods, or mouthing words, and non-vocal due to endotracheal intubation or tracheostomy. Videos were purposively chosen for this analysis of gazes based on those that had the greatest proportion of observable gaze interaction between each of the video recorded dyads. The number of gaze occurrences per video ranged from 7 to 29 with a mean of 18 gaze occurrences. Video lengths ranged from 3:15 to 15:27 minutes with a mean of 6:36 minutes.

Definitions of Nurse-Patient Gazes

Crighton, Swigart, and Happ performed a secondary analysis from the SPEACS study that examined a subset of the video recordings from the usual care group to create a typology of nurse gaze (Crighton, Swigart, Happ, 2008). From this secondary analysis Frier further developed the definitions for the codebook (Frier et al., 2016). Interrater reliability was analyzed to verify the accuracy of the data codebook for classifying nurse-patient gazes for the video recordings based on coding for four types of visual gazes: *listening, assessing, technical doing*, and *relating* (Frier, et al., 2016). “Listening” gaze is a sustained look for the purpose of understanding and interpreting the patient’s communication. “Assessing” gaze occurs when the nurse watches or observes the patient to examine a physical condition or response and often completed with physical distance between the nurse and patient, such as the nurse standing across the room rather than at the bedside. “Technical doing” gaze is a shifting gaze that alternates from the patient to the task the nurse is completing and positions the nurse close to the patient. The “relating” gaze assists to establish or reciprocate a connection through direct eye contact and is usually associated with gestures like a smile or wink.

Procedures

The four types of visual gazes: listening, assessing, technical doing, and relating, were coded by two independent coders (AM, KF) from video recordings of interactions between mechanically ventilated patients and their nurses in the Medical (MICU) and Cardiothoracic Intensive Care Units (CTICU). Using 13 of these video recordings, interrater reliability was established by having the videos coded for the visual gazes based on the codebook developed, after four additional videos (not included in the analysis for this thesis) were initially used as a training mechanism for coding the types of gazes. The method of raw percentage was used to assess interrater reliability. Raw

percentage is calculated similarly where the total number of possible items to be agreed on is divided by the number of times the data collectors agree (Burns, 2014). As explained by Lombard et al., there are eight steps to establish interrater reliability (Lombard, Snyder-Duch, & Bracken, 2010). Each of these steps is listed below with an explanation for how they were completed during the development of interrater reliability for the codebook of visual gazes between nurse-patient dyads.

1. Select one or more appropriate indices

The 13 video recordings (10 videos originally selected with 3 videos being added after the second coder began coding) were selected from the non-intervention/usual care group based on their high percentage of observable gaze interactions between the nurse and patient dyad. Four additional video recordings were used as training videos for initial calibration of ratings and refinement of the codebook for coding gazes. Video recordings were purposively selected based on a high frequency of at least one of each of the four visual gazes for which coding occurred.

2. Obtain the necessary tools to calculate the index or indices selected

A codebook had been previously developed and used for the purpose of coding the video recordings. Each of the video recordings was viewed and the most appropriate gaze was coded for each incidence throughout the video recordings. Coding of gazes was compared within and across the coded video recordings to assess extent of agreement in each video and between the videos for each type of gaze.

3. Select an appropriate minimum acceptable level of reliability for the index or indices to be used

For this interrater reliability assessment activity, the minimum acceptable level of reliability was set at 75%. This criterion was identified based on the minimum percentage agreement criterion from prior research on methods to improve reliability of video recording data (Haidet et al., 2009).

4. Assess reliability informally during coder training

After coding the initial 4 sample video recordings for the purposes of coder training, calibrating rates, and refining the codebook, the percent of agreement was compared together with areas of discrepancy in coding. During this time, questions and clarifications were discussed for each gaze. Clarifications included making clearer the difference between “listening” and “assessing” gaze to include a measurement of distance, as well as adding a “transitional” code option for when the nurse was not displaying any of the 4 defined gazes while walking around the patient’s room. The process of clarifying consisted of the two coders reading through the definitions in the codebook and asking questions to find areas of the definitions that need further detail. From there, the two coders created a new definition for the gaze together with input from Dr. Mary Beth Happ.

5. Assess reliability formally in pilot test

After completion of coding the visual gazes in the first four video recordings, the percentage of agreement using raw percentage agreement was calculated to verify the reliability of the codebook before completing all the video recordings in the full sample.

6. Assess reliability formally during coding of the full sample

Thirteen video recordings were coded and using an Excel spreadsheet the percentage of agreement was calculated by the method of raw percentage agreement and extent of disagreement was also noted.

7. Select and follow an appropriate procedure for incorporating the coding of the reliability sample into the coding of the full sample

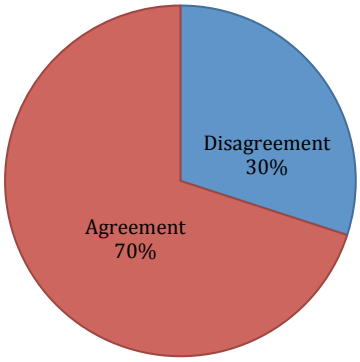
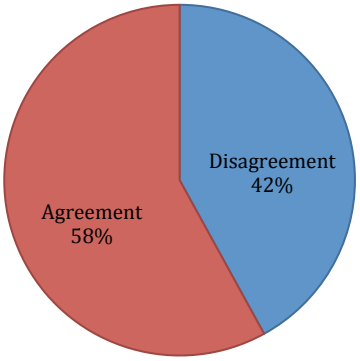
After coding each of the 4 pilot video recordings, engaging in discussion for calibration purposes, and refining the codebook, the coding was then applied to the remaining 10 video recordings. Three more video recordings were added to the study to increase the sample size, which increases the opportunities to assess for reliability of the codebook. With the three added video recordings the study has a total of 13 video recordings used.

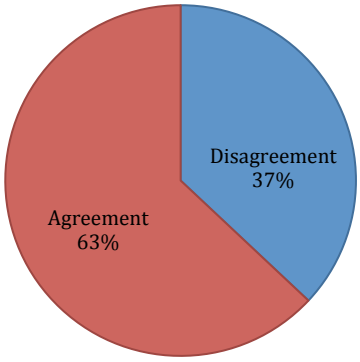
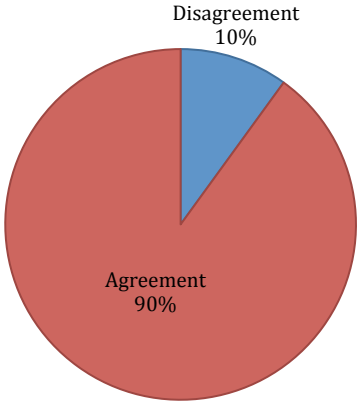
8. Report intercoder reliability in a careful, clear, and detailed manner in all research reports

Pie charts are provided in the results for the thesis to give a clear visual display of percentage of agreement and disagreement for each visual gaze, as well as the overall interrater reliability. Presentation of both agreement and disagreement data enables a balanced presentation of the results.

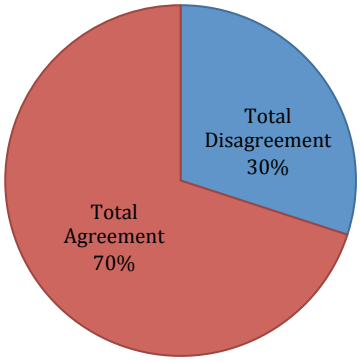
Results

Interrater Reliability Established for Each Type of Gaze

Gaze	Results
Technical Doing Gaze 	Seventy percent (70%) of the "technical doing" gazes were coded in agreement by both coders, while 30% of the time one coder chose another gaze for the coding of the video recorded interaction. Sources of disagreement often arose from confusion about the difference between the "technical doing" gaze and the "assessing" gaze. For the "assessing" gaze, a physical spacing between the nurse and patient was defined as approximately 3 to 4 feet apart, while the "technical doing" gaze spacing was defined as 2 to 3 feet apart.
Assessing Gaze 	"Assessing" gaze was more frequently incorrectly coded by one of the coders, with an agreement percentage of only 58%, due to clarifications that were needed in the codebook to distinguish "assessing" gaze from both "technical doing" and "listening." To distinguish from "technical doing" physical distance between the nurse and patient and focus of the gaze at the patient versus the task the nurse is completing was clarified. To differentiate the "listening" gaze, the intention of either examining a physical response from the patient or to understand patient communication was defined.

Listening Gaze  	"Listening" gaze was coded the same between the two coders 63% of the time and not coded with the same result 37% of the time. The key source of disagreement came from a need to clarify the definitions of "listening" gaze and "assessing" gaze. The reason the nurse is gazing at the patient is what distinguishes the two gazes.
Relating Gaze  	The coding of the "relating" gaze is the only coding category that exceeded the 75% consensus goal, with 90% agreement achieved between the two coders. The source of disagreement for coding this gaze was not a disagreement about the gaze category, but rather failure to identify the moment in the video recording, in which the gaze occurred. This gaze is different than the other gazes making it easily distinguishable for coding purposes.

Overall Interrater Reliability	Overall in the coding of all four of the gazes, the coders agreed in their coding 70% of the time (range of 58% to 90% agreement), which is just under the goal of 75% agreement. Sources of disagreement stemmed from necessary clarifications of the codebook definitions rather than the coders not being able to agree with what a certain gaze should be coded as.



Discussion

Prior to coding, the goal for total agreement was 75%, which was only achieved for one gaze, "relating," with 90% agreement. This gaze has the most discerning definition with the gaze being a mutual gaze that often established or reciprocated a connection between the nurse and the patient. "Relating" gaze could often be seen through a wink, smile, or during a hand shake making the gaze easily recognizable, where others took more focus on proximity of the nurse to the patient and the purpose of the nurse's gaze. The gaze most frequently coded incorrectly was "assessing" because the original codebook definition did not make a clear distinction between "assessing" and "technical doing," as well as "assessing" and "listening". During the process of comparing the coding performed by the two independent raters, the spacing or proximity criteria for the "assessing" gaze was redefined by making the space a greater distance, so as to make a clearer distinction between the "technical doing" and "assessing" gaze. "Assessing" was further clarified by adding that during this gaze the nurse may multitask by averting their eyes to equipment or a coworker for 1 to 6 seconds, but always maintaining a focus on the person and not the task. A focus on the task would then become a "technical doing" gaze. "Listening" gaze was clarified to define the nurse focusing on understanding the patient's communication techniques and does not include when a nurse is teaching and watching for signs of acknowledgment from the patient, which was placed in the definition for "technical doing." The definition for "listening" also help to separate the gaze from "assessing" gaze that was made more distinguishable by defining it as the nurse observing to assess a physical response or condition.

Once definitions were clarified and defined to better distinguish between categories, coding with more accuracy became easier and fewer disagreements on coding of the gazes were found over the 13 video recordings. For example, in the first video recording there were 23 incidences to be coded, with 7 of them coded differently by the

coders leading to a 30% disagreement in the first video. However, by the final two videos, the percentages of disagreement were 22% and 0%, respectively, demonstrating a decrease in disagreement, as the codebook was refined over the sequential process of coding the 13 video recordings. The best way to improve raw percentage agreement for the codebook would be to include more practice videos that allow for refining of the definitions for the coding categories, so when official coding begins, fewer questions and discrepancies arise. When creating a codebook it is important to think about what makes each category different from the rest and include that in its definition to make each category distinguishable. Another factor that affects percentage of agreement is the number of video recordings or samples used. With more instances for coding, the effect each video has on the overall percentage of agreement/disagreement is reduced. For this study, the length of the video and its number of incidences of gazes affected the percentage agreement, because each incidence is a sample within each video recording, so fewer incidences had a relatively greater impact on the overall percentage agreement.

This study used only raw percentage to assess interrater reliability. A next step in an extension of this pilot-coding project could be to use the Kappa coefficient approach to evaluating interrater reliability, which adjusts for the extent of agreement that is expected to occur by chance (random error). This method provides a more accurate estimate of interrater reliability in coding, but may result in a lower percentage agreement as the Kappa coefficient removes agreement that is expected to occur by chance.

Conclusions

In summary of this pilot coding development and testing of a codebook for a new observational measure of nonverbal nurse – patient communication in the ICU, nurse gaze, the steps of evaluating raw percentage agreement as written out by Lombard et al. were applied to help clarify the codebook and evaluate its accuracy. After completion of this study, various methods to improve interrater reliability were discussed, so that in the future these methods can be applied prior to coding in order to increase the percentage of agreement when interrater reliability is found. Interrater reliability is important not just for creating a codebook or tool of measurement that is reliable, but also for improving quality within a clinical setting. With its significant role in both research and the clinical setting, interrater reliability is important to establish when creating any codebook or tool of measurement.

References

Blumer, Herbert (1969). *Symbolic Interactionism: Perspective and Method*. New Jersey: Prentice-Hall, Inc.

Bottorff, J. L., Morse, J. M. (1993). Identifying types of attending: patterns of nurses' work. *Journal of Nursing Scholarship*, 26(1), 53-60.

Burns, M. K. (2014). How to establish interrater reliability. *Nursing, 44*(10), 56-58. doi:10.1097/01.NURSE.0000453705.41413.c6 [doi]

Carroll, S. M. (2004). Nonvocal ventilated patients' perceptions of being understood. *Western Journal of Nursing Research, 26*(1), 85-103; discussion 104-12. doi:10.1177/0193945903259462 [doi]

Crighton, M., Swigart, V., Happ, M.B. *The role of gaze in nurse-patient interactions in the ICU*. Council for the Advancement of Nursing Science 2008 State of the Science Congress. Washington, DC.

Davies, P. (1994). Non-verbal communication with patients. *British Journal of Nursing (Mark Allen Publishing), 3*(5), 220-223.

Frier, K., Martz, A., Crighton, M., Happ, M.E. The Role of Nurse Gaze in Nurse-Patient Interaction During ICU Care. American Academy of Hospice and Palliative Medicine/Hospice and Palliative Nurse Association Annual Assembly, March 9-12, 2016.

Haidet, K. K., Tate, J., Divirgilio-Thomas, D., Kolanowski, A., & Happ, M. B. (2009). Methods to improve reliability of video-recorded behavioral data. *Research in Nursing & Health, 32*(4), 465-474. doi:10.1002/nur.20334 [doi]

Happ, M. B. (2001). Communicating with mechanically ventilated patients: State of the science. *AACN Clinical Issues, 12*(2), 247-258.

Happ MB, Garrett KL, Tate JA, DiVirgilio D, Houze MP, Demirci JR, George E, Sereika SM.Effect of a multi-level intervention on nurse-patient communication in the intensive care unit: results of the SPEACS trial. Heart Lung. 2014 Mar-Apr; 43(2): 89-98.

Happ MB, Garrett K, Thomas DD, Tate J, George E, Houze M, Radtke J, Sereika S. Nurse-patient communication interactions in the intensive care unit. Am J Crit Care. 2011 Mar; 20(2): e28-40.

Karlsson, V., Lindahl, B., & Bergbom, I. (2012). Patients' statements and experiences concerning receiving mechanical ventilation: A prospective video-recorded study. *Nursing Inquiry, 19*(3), 247-258. doi:10.1111/j.1440-1800.2011.00576.x [doi]

Levy-Storms, L. (2008). Therapeutic communication training in long-term care institutions: Recommendations for future research. *Patient Education and Counseling, 73*(1), 8-21. doi:http://dx.doi.org.proxy.lib.ohio-state.edu/10.1016/j.pec.2008.05.026

Lombard, M., Snyder-Duch, J., Bracken, C. C. (2010, June 1). *Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects*. Retrieved from http://matthewlombard.com/reliability/

Morse, J. M. (1997). The comforting interaction: Developing a model of nurse-patient relationship. *Scholarly Inquiry for Nursing Practice: An International Journal*, 11(4), 321-343.

Nilsen, M. L., Happ, M. B., Donovan, H., Barnato, A., Hoffman, L., & Sereika, S. M. (2014). Adaptation of a communication interaction behavior instrument for use in mechanically ventilated, nonvocal older adults. *Nursing Research, 63*(1), 3-13. doi:10.1097/NNR.0000000000000012 [doi]

Wojnicki-Johansson, G. (2001). Communication between nurse and patient during ventilator treatment: Patient reports and RN evaluations. *Intensive & Critical Care Nursing: The Official Journal of the British Association of Critical Care Nurses, 17*(1), 29-39. doi:10.1054/iccn.2000.1547 [doi]