8-8-2014

Best Practices in High Fidelity Patient Simulation to Enhance Higher Order Thinking Skills

Kathryn S. Mock
University of South Carolina - Columbia

Best Practices in High Fidelity Patient Simulation to
Enhance Higher Order Thinking Skills

by

Kathryn S. Mock

Bachelor of Business Administration
University of South Carolina, 1986

Bachelor of Science in Nursing
University of South Carolina, 1997

Master of Science in Nursing
University of South Carolina, 2002

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Nursing Practice in

Nursing Practice

College of Nursing

University of South Carolina

2014

Accepted by:

Kathleen Scharer, Major Professor

Beverly Baliko, Committee Member

Lacey Ford, Vice Provost and Dean of Graduate Studies

DEDICATION

To Jeff, who did the laundry.

ACKNOWLEDGEMENTS

I wish to thank Dr. Kathleen Scharer, who first taught me classroom management as a new faculty member and then taught me just about everything else. Without her support and sometimes weekly phone calls this paper would never have been completed. In over ten years, she has never had an unkind word for all my foibles and idiosyncrasies. I am greatly saddened that she is retiring for her guidance and patience with me and other faculty and students has been immense. Her presence on the faculty has served teaching, research, scholarship, and friendship. May she cruise by me again one day.

I would like to thank Dr. JoAnne Herman for guiding me during the many early iterations of this project. Her guidance and advice were invaluable. I would also like to thank Dr. Beverly Baliko for her ready agreement to come on board. In addition, Dr. Bailiko's assistance in getting this project over the last few hurdles made this all possible.

ABSTRACT

Undergraduate nursing education has begun to use very expensive and time intensive high fidelity simulation activities without making full use of the ability to build higher order thinking skills in students. Current research in high fidelity patient simulation has tended to be subjective and focus on critical thinking. However, reflective thinking habits of mind must be in place before full use can be made of critical thinking skills. A comprehensive search of all reflective thinking literature used in conjunction with simulated patient experiences by healthcare students was undertaken. A guideline was created for nurse faculty to use that outlined current best practices in simulation to maximize reflective thinking. Though the research on which the guideline was based has been mainly subjective, several analytical studies were found that supported the findings. Policy changes to incorporate reflective thinking and the associated activities were recommended for nursing students and continuing nursing education. Nurse researchers and educators should incorporate reflective thinking exercises with their simulated patient undertakings to maximize higher order thinking skills.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AACN ..... American Association of Colleges of Nursing

CASP..... Critical Appraisal Skills Programme

CCTDI..... California Critical Thinking Dispositions Inventory

CCTST .....California Critical Thinking Skills Test

DML..... Debriefing for Meaningful Learning

HaPI ..... Health and Psychosocial Instruments

HFPS ..... high fidelity patient simulation

HSRT. ……Health Sciences Reasoning Test

LSCJR .....Lasater Clinical Judgment Rubric

OPT model .....Outcome Present state Test model

OSCE ..... Objective Structured Clinical Examination

pt/pts..... patient/patients

RJI..... Reflective Judgment Interview

SIGN ..... Scottish Intercollegiate Guidelines Network

WGCTA.....Watson Glasgow Critical Thinking Assessment

CHAPTER 1

INTRODUCTION

Background and Significance of the Problem

This introductory section presents the evidence for the adoption of reflection thinking exercises during high fidelity patient simulation (HFPS) in order to increase critical thinking in undergraduate nursing students. High fidelity patient simulation has been the most expensive type of simulated patient activity and during this century has been rapidly incorporated into undergraduate nursing education programs (Hoffmann, O'Donnell, & Kim, 2007; Medley & Horne, 2005; Rhodes & Curran, 2005). A large investment in equipment, manpower, and training to simulate patient situations has been required to undertake HFPS, over five times the cost of medium fidelity manikins (Lapkin & Levett-Jones, 2011). In order to make cost effective use of HFPS, objectives should include measures designed to increase students' higher level thinking skills and not focus solely on skills which could be more cheaply obtained using other methods (Lapkin & Levett-Jones, 2011).

Critical thinking. Critical thinking has been the higher level thinking skill that is the standard for undergraduate nursing programs (American Association of Colleges of Nursing [AACN], 2008; National League for Nursing Accrediting Commission, 2006). This emphasis has also been supported by the National Council of the State Boards of Nursing (2012). However, the evidence does not generally support that changes in critical thinking has been increased during the course of a nursing student's education

(Chau, Chang, Lee, Ip, & Wootton, 2001; Notarianni, 1991; Profetto-McGrath, 2003; Walsh & Seldomridge, 2006). A comprehensive definition of critical thinking, and the related skills and dispositions, was determined by the American Psychological Association's Delphi study (Facione, 1990). The consensus definition of critical thinking has been: "purposeful, self-regulatory judgment which results interpretation, analysis, evaluation, and inference, as well as the explanation of the . . . considerations upon which that judgment was based" (Facione, 1990, p. 2). Facione went on to describe the agreed upon skills and sub-skills that supported critical thinking (see Appendix A). Also described by the study, were affective dispositions that were seen as conducive to critical thinking (see Appendix B). The Delphi study opinion was that although critical thinking dispositions and skills transcend subject matter there may be additional knowledge, methods, or techniques needed to solve discipline specific problems (Facione, 1990). The APA definition has not been surpassed as the standard by which critical thinking is measured.

Facione (1990) believed it was not enough to teach a student logical analysis to promote critical thinking. Logic analysis was described by Dewey (1933), in his seminal work on critical and reflective thinking, as an abstract idea, while thinking has been based in context, such as a patient situation. Therefore, the teaching of formal logic has not been enough to allow the learner to apply logic to problems or situations (Dewey, 1933). This corresponds to Brookfield (1987) who believed that the process of critical thinking is supported by the processes of reflective analysis of the experienced situation. Brookfield proposed that reflection on assumptions and actions was a skill that needed to be developed in order to critically think.

Measures of critical thinking in nursing. The three most common objective measures of critical thinking in nursing students are the California Critical Thinking Dispositions Inventory (CCTDI), the California Critical Thinking Skills Test (CCTST) and the Watson-Glaser Critical Thinking Appraisal (WGCTA). The CCTST and the CCTDI were based on the APA (Facione, 1990) consensus definition of critical thinking. The CCTDI is a valid and reliable instrument made of seven subscales that describe the dispositions thought to be essential in order for a person to be able to critically think: truthseeking, openmindness, analyticity, systematicity, self-confidence, inquisitiveness, and cognitive maturity (Facione, Facione, & Sanchez, 1994). The CCTST is also a valid and reliable instrument and is comprised of five subscales that describe the skills needed to critically think about a situation or problem: analysis, evaluation, inference, deduction, and induction (Facione & Facione, 1994). The WGCTA definition of critical thinking has been frequently used by nursing schools (Vaughan-Wrobel, O'Sullivan, & Smith, 1997). The validity and reliability of the WGCTA has been well established in other undergraduate majors (Hassan & Madhum, 2007). There are five subscales contained within the WGCTA: inference/discrimination, recognition of assumptions, deduction, interpretation, and evaluation of arguments (Vaughan-Wrobel et al., 1997).

A systematic review of the changes in CCTDI, CCTST, WGCTA scores of undergraduate nursing students after a problem-based learning intervention revealed small improvements in the overall scores (Ling-Na, Bo, Ying-qing, Shao-yu, & Hui-Ming, 2014). However, the meta-analysis of the eight randomized controlled trials showed no significant changes in any of the CCTST and most of the WGCTA subscale

scores (Ling-Na et al., 2014). Additionally, two of the studies did not find any improvement in overall CCTST, CCTDI, or WGCTA scores (Ling-Na et al., 2014).

Profetto-McGrath's (2003) cross-sectional study measured critical thinking skills and dispositions in baccalaureate nursing students over four years using the CCTST and the CCTDI. In the sample of 228 volunteers, CCTST scores increased with each year of college, with the exception of the third year; however these increases were not statistically significant (Profetto-McGrath, 2003). The relationship between the students' critical thinking skills and dispositions was statistically significant (Profetto-McGrath, 2003). Eighty-five percent of the students had acceptable scores on the CCTDI; however, there was not a statistically significant difference in scores over the four years (Profetto-McGrath, 2003). The lack of a statistically significant progression in skills was felt to be related to the students' cognitive developmental level (Profetto-McGrath, 2003) as measured by Perry's (1970) schema of cognitive and ethical developmental levels.

Over the course of 15 years, Perry (1970) conducted reflective interviews of college students, at the end of their freshman, sophomore, junior, and senior years. Perry's model classified students as being in one of nine stages of intellectual and ethical development based on their reflective thinking processes. The first five stages dealt primarily with intellectual development, while the final four represented moral development and identify formation (Perry, 1970). Students were generally observed progressing from dualistic thinking, multiplicity, relativism, and possibly to commitment in relativism (Perry, 1970). Unfortunately, students might also have regressed, delayed, or escaped the commitment stage and avoided personal responsibility (Perry, 1970). The lowest levels of cognitive and ethical development, Positions 1, 2, and 3, are dualistic,

which are exemplified by dichotomous or right/wrong beliefs and thinking. The next set of positions, 4, 5, and 6, were defined as multiplistic or relativistic viewpoints that embraced the graduations of beliefs held by others and appreciated the effect of context on decision making (Perry, 1970). The highest levels of cognitive and ethical development, positions 7, 8, and 9, were defined by the students' level of commitment and personal responsibility in regards to their belief system (Perry, 1970). Also included within the model are positions describing: a retreat to an earlier level, a delay at one level, and an escape to negativity at position 4 or 5 (Perry, 1970).

Students in Profetto-McGrath's (2003) study were judged as being at the dualistic or multiplistic stage of cognitive development and had not progressed on to the relativistic or commitment stage of cognitive development. Cognitive development was seen as requiring more than four years of undergraduate education to optimally mature (Profetto-McGrath, 2003). The lowest sub-score on the CCTDI was truth-seeking and this was felt to reflect the lecture presentation of large volumes of material that needed to be memorized (Profetto-McGrath, 2003). Profetto-McGrath reported that another explanation for the low scores on the truth-seeking scale could have been that the students felt faculty did not welcome student questions or requests for clarification. Implications of this study include the recommendation that nurse educators learn about critical thinking skills and dispositions, and utilize strategies to develop critical thinking skills and dispositions in students (Profetto-McGrath, 2003). Some of the suggested strategies for improving critical thinking skills included reflective journals, papers, and the use of Socratic questioning. (Profetto-McGrath, 2003). Additionally, another researcher found no difference between student's scores on California CCTST after

participating in an educational intervention using videotaped vignettes even though knowledge scores improved (Chau et al., 2001).

McCarthy, Schuster, Zehr, & McDougal (1999) used the CCTST and the CCTDI to determine if there was a difference in the critical thinking of sophomore and senior BSN students. The large sample was comprised of 156 sophomore students and 85 seniors (McCarthy et al., 1999). The seniors scored significantly higher on the CCTST and the CCTDI, which were significantly correlated in the combined cohorts (McCarthy et al., 1999). The study was limited by possible differences in the cohorts and does not mention how many students were in the senior class cohort as sophomores and had failed to progress (McCarthy et al., 1999). A significant flaw in the choice of a cross sectional study is the failure to mention why the sophomore cohort was almost twice the size of the senior cohort (McCarthy et al., 1999). The sophomore and senior students had similar GPAs and scores on the American College Test (McCarthy et al., 1999). However, a longitudinal study would have revealed if sophomores who did not score well on the CCTST and CCTDI also failed to progress (McCarthy et al., 1999).

Colucciello (1997) also conducted a cross sectional study of nursing students using the CCTST and the CCTDI. A total of 328 students were in the sample: 94 second semester sophomores, 65 first semester juniors, 64 second semester juniors, 59 first semester seniors, and 46 second semester seniors (Colucciello, 1997).. As with McCarthy et al.'s (1999) sample, the cohort size decreases from sophomore to senior year (Colucciello, 1997). The first semester junior cohort had the highest overall CCTST score (Colucciello, 1997). The second semester sophomore cohort had the lowest CCTST score, but the students had not yet been admitted to the nursing program and did

not necessarily meet the requirements for admission (Colucciello, 1997). Similar to Profetto-McGrath (2003), Colucciello found that the truth-seeking subscale of the CCTDI was the lowest of all the subscales. Overall, a significant positive association between the CCTST and the CCTDI scores was found (Colucciello, 1997). The first semester juniors' and the first and second semester seniors' scores were significantly higher than the sophomores CCTDI scores (Colucciello, 1997). This study was limited by the fact that the sophomore cohort was not yet admitted to the program and was not equivalent to the other cohorts (Colucciello, 1997). There was not a progression in the CCTST cohort scores or a clear pattern to the CCTDI scores (Colucciello, 1997).

The evidence is mixed at best supporting the use of CCTST and the CCTDI to measure changes in undergraduate nursing students. Although McCarthy et al. (1999) found higher CCTST and CCTDI scores in senior versus sophomore students, no explanantion was given for the much smaller sample size of senior students. Therefore, the CCTST and CCTDI have not been shown to be appropriate measures of changes in the thinking skills of undergraduate nursing students (Chau et al., 2001; Colucciello, 1997; Profetto-McGrath, 2003).

When using a critical thinking instrument that was designed for nursing, the WGCTA, mixed results have been found (Gross, Takazawa, & Rose, 1987; Magnussen, Ishida, and Itano, 2000; Notarianni, 1991; Sedlak, 1997; Walsh & Seldomridge, 2006). Sedlak (1997) felt that measures such as the WGCTA might be less useful than longitudinal studies since the development of critical thinking is an ongoing process. The WGCTA has been thought to be a more accurate measure of metacognitive processes than the CCTSI and CCTDI, due to the combination of well- and ill-structured problems

(King & Kitchener, 1994). Gross et al. (1987) found improvement in WGCTA scores after students completed either the associate's or Bachelor's degree program at the University of Hawaii. Magnussen et al. (2000) found that after an inquiry-based learning intervention, low scoring students improved their scores on the WGCTA, but the high scoring students' average score declined and in medium scoring students there was no significant change in pre and post scores.

Notarianni's (1991) pre-test/post-test longitudinal study measured critical thinking in 321 associate's and bachelor's degree nursing students using two versions of the WGCTA. Neither first nor third year students in BSN programs showed statistically significant gains in WGCTA scores. There was a statistically significant drop in the WGCTA scores of second year BSN students. Forth year students also showed a drop in their scores but it was not statistically significant. Additionally, second year associate's degree students had a statistically significant drop in their scores. Overall the WGCTA scale showed insignificant or negative changes in the critical thinking of nursing students over the course of their studies. The WGCTA did not show that nursing students increased their critical thinking skills over a year of instruction or program of study (Notarianni, 1991).

Critical thinking as measured by the WGCTA in nursing students appears to be correlated to the students' ability to successfully complete a simulated patient scenario (Brooks & Shepherd, 1990). In their study of 200 nursing students, Brooks and Shepherd (1990) found a small but statistically significant positive link between WGCTA scores and clinical decision-making as measured by the Nursing Performance Simulation

Instrument. The Nursing Performance Simulation Instrument consists of four questions about 6 patients that ask the student to:

- choose whether a patient care activity is warranted,
- prioritize the needs of the patients,
- decide to whom the patients need to be referred, and
- choose between a pair of actions within the context of three clinical situations (Brooks & Shepherd, 1990).

Interestingly, although the generic BSN students had higher critical thinking scores, their Nursing Performance Simulation Instrument scores were identical to students in an associate's degree or diploma program (Brooks & Shepherd, 1990). Students enrolled in a RN to BSN program showed high clinical judgment scores (Brooks & Shepherd, 1990). The RN to BSN program students had completed three years of clinical practice before enrolling, and this was felt to have contributed to their significantly higher Nursing Performance Simulation Instrument scores (Brooks & Shepherd, 1990). Statistical significance was also demonstrated in the higher critical thinking scores of both the generic or RN to BSN students as compared to students in an associate's degree or diploma program (Brooks & Shepherd, 1990). Students may be self-selecting according to their critical thinking ability into a diploma, ADN or BSN program. Additionally, being enrolled in a BSN program may contribute to the development of critical thinking skills (Brooks & Shepherd, 1990). However, both the CCTDI and WGCTA were used from 1997 to 2002 in another undergraduate nursing program with no consistent findings and no explanations that seemed to fit the data (Walsh & Seldomridge, 2006).

Standardized measures have not been shown to measure improvements in the critical thinking of nursing students after educational interventions or over the course of their education (Cant & Cooper, 2010; Gross et al., 1987; Magnussen et al., 2000; Notarianni, 1991; Walsh & Seldomridge, 2006). This is not surprising since Gordon (2000) found that while nurse faculty agreed on the skills and dispositions of critical thinking as defined by Facione (1990), faculty did not agree on the concepts related to critical thinking. Additional concepts identified by nurse faculty included decision-making and clinical reasoning (Gordon, 2000). Another reason why standardized instruments do not tend to record differences in nursing students thinking about patients is that standardized instruments by their very nature tend to measure reductionist logic skills and not the holistic thinking desired in nursing (Walsh & Seldomridge, 2006). No best objective standardized method for measuring critical thinking in nursing students has been identified (Navedo, 2006).

The preceding objective evidence has been supported by the subjective opinions of leaders in nursing education. In 2001, Stone, Davidson, Evans, and Hansen surveyed the deans and directors of NLN-accredited nursing programs at the baccalaureate level or higher on their beliefs on critical thinking. Stone, et al. (2001) found that the deans and directors felt the CCTDI and CCTST contained skills and traits that were essential to the practice of nursing. However, the deans and directors did not believe that the CCTST was an appropriate measure of the critical thinking skills of a nurse (Stone, et al., 2001). Perhaps this is because the critical thinking skills of a nurse lead to clinical judgments that are not just composed of logical analysis. The highly valued clinical judgment of an experienced nurse has been context driven and developed through the application of

critical and reflective thinking to varied clients and situations (Kuiper & Pesut, 2004). Kuiper and Pesut (2004) postulated further that both critical and reflective thinking skills have been needed for the development of clinical judgment.

Chabeli (2007) concluded that although critical thinking is entwined with the nursing process, and it is difficult for nurse educators to encourage and measure critical thinking in nursing students. Measuring critical thinking in undergraduate nursing students has brought mixed results (Chau et al., 2001; Gross et al., 1987; Magnussen et al., 2000; McCarthy et al., 1999; Notarianni, 1991; Profetto-McGrath, 2003; Sedlak, 1997; Walsh & Seldomridge, 2006). Perhaps the reason for these mixed results has been that critical thinking measures tend to use well defined problems, while patients are unique and their clinical presentation may be ambiguous and might not match a textbook case (Chabeli, 2007).

Sedlak's (1997) qualitative study found that sophomore nursing students' reflective writing journals showed evidence of critical thinking after exposure to critical thinking content. An additional benefit gained by students from reflecting on their experiences has been that reflection promotes critical thinking and self-directed learning (Sedlak, 1997). A primary difficulty in measuring critical thinking changes in nursing students has been that the experiences and the lessons learned through reflection are inherently unique to the individual and not easily quantified or compared (Boyd & Fales, 1983). Perhaps this has been the reason why standardized objective measures have not been conclusively shown to be useful measures of critical thinking in nursing students and may not be valid measures of meaningful learning for this population (Boyadjian-

Samawi, 2006). However, it does appear that critical thinking can be encouraged through the use of reflective techniques (Sedlak, 1997).

Reflective thinking. Reflective thinking has been proposed as a precursor to critical thinking (Brookfield, 1987). Changes in reflective thinking have been successfully measured in undergraduates, including nurses (King & Kitchener, 1994). Little evidence has been found that examines reflective thinking with HFPS (Decker, 2007; Stirling, Smith, & Hogg, 2012). However, there exists a large reservoir of evidence examining reflective thinking during other simulated patient exercises. The question remains: can undergraduate nursing faculty make use of the evidence in other simulated patient experiences to improve the reflective thinking abilities of students using HFPS?

Since the current evaluative instruments used for critical thinking have not measured changes in the thinking of nursing students over the course of their education, perhaps measuring gains in reflective thinking would stand as a proxy. Dewey's book *How We Think* (1933) framed the arguments for the teaching of thinking as the mission of formal education. The term critical thinking was not used, but instead the term reflective thinking was used to describe what educators should teach. Dewey's delineation of the term reflective thinking, laid the foundation for both critical and reflective thinking of other authors. Dewey believed that reflective thinking involves "a careful comparing and balancing of evidence and suggestions, a process of evaluating what occurs. . ." (p. 76). Reflective thought is the method by which critical thinking is carried out. "The function of reflective thought is, therefore, to transform a situation in which there is experienced obscurity, doubt, conflict, disturbance of some sort, into a

situation that is clear coherent, settled, harmonious" (Dewey, 1933, pp. 100-101). What Dewey called reflective thinking is the "turning a subject over in the mind and giving it serious and consecutive consideration" (p. 3). Consecutive, in this usage meant that the thoughts are determined by the outcome of the preceding ideas, in the sense of consequences. Thoughts are linked as in a chain and are stronger than the usual sort of stream of consciousness thinking. Reflective thought has two stages: a state of doubt, hesitation, or controversy, and the mental searching for meaning to resolve the doubt. Therefore, reflective thought is driven by perplexity. The next step in the reflective process is the selection and weighing of evidence that is applicable to problem. Then, the choice of principles and their application is considered. The last step is the formation of a decision which closes the problem (Dewey, 1933).

Schon's (1983, 1987) work on reflective practice was rooted in Dewey's theory. Schon believed that reflection was poorly understood by those involved in the education of professionals that instead relied upon the technical-rational approach. He believed that the technical-rational approach that has prevailed in nursing, where procedure lists and textbook cases dominate, has been inappropriate for the training of professionals who work in ill-defined, complex, muddled situations (Schon, 1987). This thought is echoed by Grunwald & Corsbie-Massay (2006), who posits that behaviorist theory has promoted the use of a technical-rational approach that does not focus on internal thought process but concentrates instead on the use of memory. Behaviorists have seen critical thinking as a method to be applied to a problem in order to solve it, rather than an approach that encompasses the recognition of an ill-defined problem and examination of the underlying assumptions (Grunwald & Corsbie-Massay, 2006). To a behaviorist, simulation is best

used to assist the student to have a successful experience and that causes the student to replicate the behaviors that led to the successful experience (Grunwald & Corsbie-Massay, 2006). However, the goal of the reflective process has been to promote cognitive and affective changes after an experience and not merely a honing of recognition and psychomotor skills (Boyd & Fales, 1983; Sedlak, 1997).

Schon (1983) defined two different types of reflection that occur at different time in reference to an encounter. The names for the different types of reflection have been called various things by different authors. For clarity, Schon's concept of reflection during action is defined as reflection takes place while the practitioner is in the midst of caring for a patient. Reflection takes place after the encounter is finished, will be referred to as reflection after action. Greenwood (1993) expanded Schon's (1983, 1987) work to include the concept of reflection before action which is thought direct at planning for future situations. Dewey (1933) wrote that one of the advantages of reflecting before action has been that once an action is undertaken, it cannot be undone. Reflection before action involves thinking through the anticipated problem, planning intended actions, and considering the consequences (Greenwood, 1993). Reflection before action has allowed students to organize their thinking, problem solve, and mentally rehearse the scenario (Greenwood, 1993). Reflection before action may occur while completing the research for a simulation, after the briefing, or at any point before the student begins to take action.

Boud (2001) also included a preparatory reflective thinking stage he called, reflection in anticipation of events (reflection before action). There are three main foci of Boud's reflection before action: the learner, context, and learning skills and strategies.

The learner aspect is concerned with the intentions, goals, and expectations of the learner. Additionally, the learner aspect encompasses the strength of these concerns, and the bearing these concerns may have on steering the learner away from other possibilities. The second focus is on the context of the event. The context includes all features of the situation, including any briefing or preparation on the part of the learner. The last focus is on learning skills and strategies which consists of: what the learner plans to notice, fall back plans, and rehearsal for the cognitive, psycho-motor, and affective domains, (Boud, 2001). All of these aspects must be taken in account when planning simulated patient experiences.

Since Dewey (1933) first wrote about critical and reflective thinking, authors have been teasing out the relationship between the two. Three types of reflection have been identified: reflecting during action, reflecting after action, and reflecting before action (Boud, 2001; Greenwood, 1993; Schon, 1983, 1987). Healthcare professional education needs to include reflective thinking activities in order to prepare students for solving the ill-defined problems that they will encounter in their work (Boud, 2001; Greenwood, 1993).

Measuring reflective thinking. King and Kitcherer (1994) applied Perry's (1970) model of cognitive and ethical development to reflective judgment and continued to assess students through reflective interviews. The seven stage reflective judgment model is summarized in Table 1.1. In summary, students in the pre-reflective stages believed that knowledge was established and did not recognize the difference between well-defined and ill-structured problems. In the quasi-reflective stages (4 and 5), the difference between well-defined and ill-structured problems are recognized. Judgments

Table 1.1 Seven Stages of King and Kitchener's Reflective Judgment Model

Stage	Source of Knowledge	Justification of Beliefs
1 Pre-reflective	Absolute, concrete, through direct observation	No justification needed
2 Pre-reflective	Direct observation, authority figures	Not examined, one correct answer
3 Pre-reflective	May be temporarily unable to be verified, generally acknowledged	Based on authority, personal opinion used in the absence of concrete evidence
4 Quasi-reflective	Uncertain, claims to knowledge may be based on variables that are incorrect	Citing of evidence, reasoning, knowledge and beliefs are unique to individual
5 Quasi-reflective	Dependent on the situation, unique to each individual, subjective interpretation of events	Based on situation, weighed against other explanations.
6 Reflective	Individual experiences with prior ill-structured problems, highly regarded sources	Synthesis of evidence and expert opinions, variety of perspectives, weighting of evidences, utility of solution, perceived need for action
7 Reflective	Reasonable inquiry, evaluation of plausibility, reevaluated when new data or methods are available, analysis of wide range of explicatory factors, including risk of being wrong and possible consequences	Exhaustive investigations resulting in comprehensive, credible, or convincing evidence based on current research and experience

Note. Adapted from "Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking Adolescents and Adults," by P.M. King and K.S. Kitchener, 1994.

in ill-structured problems were challenging and the students did not know how to deal with making a decision when all the elements were not well defined. In the reflective judgment stages (6 and 7) the students recognized that data must be appraised and that the absolute truth may be unknown. Reflective judgment has been seen both as developing progressively and the key to solving ill-structured problems. King and Kitchener believed that reflective thinking was developed through the "interaction between the individual's conceptual skills and environments that promote or inhibit the acquisition of these skills (p. 7)." The Reflective Judgment Interview (RJI) developed by King and Kitchener was designed to allow interviewers to code student responses to open ended questions. The RJI measured the student's level of knowledge development and belief justification about ill-structured situations based on the student's use of evidence, experience, reason, and inquiry. In the King and Kitchener's original longitudinal study, the RJI was given to 20 high school, 20 college, and 20 doctoral students. The students were followed for 10 years and tested up to four times (in 1977, 1979, 1983, and 1987). Scores on the RJI were directly correlated to the seven stages of reflective judgment. The average reflective judgment score on the RJI tended to rise from 2.77 to 5.29 in the original high school student sample over the ten years of the study. The original college juniors' RJI scores also rose, from 3.76 to 5.05. Doctoral students' scores did not change significantly over the same time period, but did rise from 5.67 to 6.21. This was possibly due to the ceiling affect, since the doctoral students' scores were approaching seven, although no student had a perfect score. In seven other longitudinal studies reviewed by King and Kitchener (1994), 241 individuals, ranging in age from teens to middle-aged adults were interviewed according to the RJI protocol. The individuals' educational

levels varied from high school to graduate school. Individuals, who had completed at least two RJI, had either stable or rising RJI scores. The amount of time between interviews was positively correlated with a rise in scores.

Additionally, King and Kitchener (1994) reviewed the results of 25 cross-sectional studies. These studies had results that correlated with the students' educational level and the scores were moderated by academic ability. Compilation of the cross-sectional studies revealed average an RJI score of 3.2 for high school students, 3.8 for college students, and 4.8 for graduate students. Twenty of the 25 studies measured RJI scores in a total of 966 college students under the age of 25. In these twenty studies, the average freshman score was 3.6 and the average senior score was 4.0; demonstrating a rise in reflective thinking scores over the course of college education. The rise in scores may have been affected by many factors other than classroom, lab, and clinical experiences, with the most obvious being age. However, 137 adult learners' scores, as measured in five of the cross-sectional studies reviewed by King and Kitchener, were very similar to the traditionally aged students, demonstrating a rise in scores from freshman to senior year. The six studies of adults not currently in an educational program provided a control. Adults, who had previously earned a college degree, scored an average of 4.3 and adults who had not completed a college degree scored an average of 3.6. Overall, the higher RJI scores appeared to be correlated with increasing educational attainment. However, individual scores also revealed regressions and stalls that demonstrated considerable variability in how a person passes through the stages of reflective judgment. Reflective judgment typically follows the Reflective Judgment

Model, enrollment in an educational environment either as student or faculty at any point in life resulted in higher overall RJI scores (King & Kitchener, 1994).

As part of a larger study, Navedo (2006) evaluated seven senior nursing students using two of the standardized dilemmas from the RJI, truth in news reporting and the safety of chemical additives, and two additional researcher created nursing dilemmas. The two researcher-created dilemmas involved post-operative pain relief with narcotics, and early hospital discharge. Students were rated independently by two reviewers and were either given a single stage score of 1-7 or a range of two adjacent scores. Inter-rater reliability was calculated on 80 out of 85 scores to be from 85.7 to 89.5 percent on the dilemmas and 90 percent or greater on individuals except for one student where there was 50 percent agreement. Using a two tailed test the Pearson's Product-moment Correlation was .505 ($p<0.01$). After calculating reliability, the reviewers met and were able to resolve any differences in scoring. Individual scores on specific dilemmas and composite scores both ranged from 3-4 to 5-6. Navedo found that the two researcher-developed nursing dilemma scores correlated best with each other ($r=.823$, $p<0.05$). The postoperative narcotic use dilemma had significant correlations with the truth in news reporting ($r=.706$, $p<0.05$), but was not correlated as highly ($r=588$, $p<0.01$), with the safety of chemical additives dilemma. However, the early discharge scenario was not significantly correlated with either of the standard RJI scenarios. The overall mean student score was 4.43. Using both standard and researcher developed dilemmas, senior nursing students were able to show comparable scores on the RJI to other traditionally aged undergraduate students. Since other traditionally aged undergraduate students have been able to show gains in the RJI over the course of their education; then perhaps the

RJI or similar dilemmas can be used to evaluate changes in the thinking of baccalaureate nursing students (Navedo, 2006).

Kataoka-Yahiro and Saylor (1994) developed a critical thinking model specifically for nursing based in part on Perry's (1970) work. The model categorized critical thinking to three levels: basic, complex, and commitment (Kataoka-Yahiro & Saylor, 1994). The components of critical thinking that were thought to lead to nursing judgment consisted of: competencies in critical thinking, attitudes for critical thinking, standards in critical thinking, experience in nursing, and specific knowledge base in nursing (Kataoka-Yahiro & Saylor, 1994). Competencies in critical thinking while considered overlapping were further broken down into general critical thinking, specific to patient situations, and specific to nursing process (Kataoka-Yahiro & Saylor, 1994). Each level of critical thinking in Kataoka-Yahiro and Saylor's model corresponds to three of Perry's (1970) positions. Basic level thinking was considered comparable to dichotomous thinking or dualism (Kataoka-Yahiro & Saylor, 1994). Complex level thinking encompassed the multiplistic and relative thinkers, who had the ability to think about their thinking (Kataoka-Yahiro & Saylor, 1994). Commitment level was used as the top level of intellectual development in both models (Kataoka-Yahiro & Saylor, 1994). The model on nursing judgment, while consisting of many subcomponents not listed here, was considered a simpler way for nurse educators to classify student's critical thinking (Kataoka-Yahiro & Saylor, 1994). No studies were found that used the Kataoka-Yahiro and Saylor model with nursing students. Rapps (1998) used the model in a study of graduate nurses. Critical thinking level was not directly measured and years of experience as a proxy measure of critical thinking level (Rapps, 1998). The findings

of the study did not support a model of critical thinking and cognitive development (Rapps, 1998). This was not surprising since an inappropriate proxy was used. The Kataoka-Yahiro & Saylor's model has yet to be tested in undergraduate nursing students and, therefore, was of limited utility for this review.

Patient simulation. Alinier (2007) arranged simulation methods into five categories. The lowest level of simulations, Level 0, does not involve manikins but is a passive cognitive experience such as case studies (Alinier, 2007). Level 1, commonly called low fidelity primarily involves psychomotor skills, may be a task trainer such as an IV arm or a basic manikin (Alinier, 2007). A basic manikin is one that does not interact with the student but is designed to allow the student to practice skills such as: catheterization, giving enemas, starting IV's, and dressing wounds (Alinier, 2007). Level 2 simulations are computer simulations of patients and do not involve a manikin (Alinier, 2007). Level 3 simulation uses standardized patients portrayed by actors or volunteers and is a psychomotor, cognitive, and interpersonal activity (Alinier, 2007). Level 4 is considered medium level fidelity and involves manikins that are programmable and partially interact with the student (Alinier, 2007). The highest level of simulation uses fully interactive manikins and is an immersive experience involving psychomotor, cognitive, and interpersonal aspects (Alinier, 2007).

Low, medium, and high fidelity patient simulators have been recent additions to the gamut of simulated experiences which include: clinicals, virtual patients, standardized patients, case studies, and task trainers (Alinier, 2007; Magee, 2006; Nehring, 2008). The term fidelity has referred both to the physical and cognitive fidelity of the experience (Goettl, Ashworth, & Chaiken, 2007). Physical fidelity has most

commonly been thought of as how closely the manikin and room set up resembles real patient care situations (Goettl et al., 2007). Cognitive fidelity has been described as the way in way a situation resembles the type of choices that must be made in order to solve the problem (Goettl et al., 2007). High fidelity patient simulation can allow for both high physical and cognitive fidelity without the use of human patients or actors playing the role of standardized patients. Within nursing education, HFPS has been seen as a solution to many problems, including the following:

- Lack of clinical space (Medley & Horne, 2005);
- Inability to collaborate with other disciplines (Medley & Horne, 2005; Reese, Jeffries, & Engum, 2010);
- Limited opportunities to present high acuity and low frequency events (Lasater, 2007a);
- Concerns about patient safety when cared for by student nurses (Medley & Horne, 2005); and
- Unnoticed gaps in students' understanding, clinical practice, and skills (Lasater, 2007a).

Supervised clinical practice should be the best place to apply the principles of nursing and learn technical procedures. However, the reality has been that the sometimes too rapid pace of patient care has not been the best environment for learning (Goettl et al., 2007; Sedlak, 1997). Another problem has been that to encourage pattern recognition skills, constellations of patient presentations should be presented many times (Goettl et al., 2007). Finding patients that fulfill the pattern requirements may not be possible (Dewey, 1933). Dewey (1933) believed that a patient of the right kind could be the basis

for reflection that could be applied to many other patient situations but that patients of this kind did not occur frequently. Clinical instructors have often tried to find patients for their students that had conditions which were being covered in class. For example, due to the nature of human morbidity patterns, there may have been many pneumonia and COPD patients in the winter and fewer patients with other problems. With patient simulation, instructors could have presented an appropriate clinical case whenever needed that could have been linked to the course content.

Human patient simulation has been a bridge between the theoretical learning in the classroom and practice learning taking place during clinical experiences (Leigh & Hurst, 2008). Simulation has allowed learners to employ their understanding of principles to new situations. The application of principles to new situations has been the best way for students to demonstrate what they have learned (Dewey, 1933). However, the real strength of HFPS has been the ability to assist the students in forming habits of mind that can improve their practice over time by the incorporation of reflective techniques, before, during, and after their simulation experiences. Reflection has been identified as an essential conduit between theory and practice (Jones & Alinier, 2009) and critical to the experiential learning process (Boud, 2001; Boyd & Fales, 1983).

Experiential learning such as clinical practice and simulation has been based on the theory that ideas are not unchangeable but re-formed through experience (Kolb, 1984). Students have learned by processing their experience during a post experience analysis and creating new memories and meanings (Lederman, 1992). Connections have been made and developed through extended reflection and new understandings formed that allow for a more holistic understanding (Fonteyn & Cahill, 1998). The rapidly

developing knowledge base of nursing students has made reflective practice an integral part of their development (Fonteyn & Cahill, 1998).

According to Jeffries (2007) current best practice in education as well as simulation has consisted of opportunities for active learning, specific constructive feedback, student-faculty interactions, and collaboration with fellow students. When creating defined scenario roles for students, the potential for self-directed learning must be incorporated into the design. Faculty must be able to support learning and provide appropriate cues, prompts and questions to stimulate thinking and reflection. Specific learning objectives, a defined level of complexity, maximum fidelity, and a debriefing strategy that includes guided reflection must all be delineated in scenario planning. Outcomes that should be measured during simulation include skills, knowledge, student satisfaction, self-confidence, self-efficacy, critical and reflective thinking (Jeffries, 2007). However, not all of these outcomes may be good proxies for changes in critical or reflective thinking.

A well designed simulation activity should have five distinct parts: briefing, simulation, debriefing, extended reflection, and evaluation (Henneman & Cunningham, 2005; Jeffries, 2007). The briefing is defined as including faculty rehearsals as well as conveying to the students information concerning the scenario, directions, and expectations (Jeffries, 2007). The scenario planning should anticipate many possible student actions and include appropriate scripting (Jeffries, 2007). Debriefing should consist of the time spent with the simulation group, instructors, evaluators, and observers in which the scenario is reviewed and meaning is explored (Jeffries, 2007). Extended reflection refers to any activities designed to have the student further reflect on what

happened during the scenario and how things could have been done differently (Henneman & Cunningham, 2005). The evaluation phase should be completed by both students and faculty with an eye toward improving the simulation experience in addition to measuring learning outcomes and skills (Jeffries, 2007).

High fidelity patient simulation has been the newest form of experiential learning employed by nurse educators (Alinier, 2007; Magee, 2006; Nehring, 2008). There have been many educational and practical advantages to using HFPS (Medley & Horne, 2005; Reese et al., 2010). However, authors have not noted that reflective techniques have been used to enhance critical thinking associated with HFPS (Medley & Horne, 2005; Reese et al., 2010). Further, since the basis of experiential learning, such as a simulated patient experience, has been that new meaning is created by analysis and evaluation of the event through reflection (Boud, 2001; Boyd & Fales, 1983), the most effective use of HFPS has not been used reported in the literature.

Critical thinking and simulated nursing experiences. The evidence examining the effect of simulation on critical thinking has been either poorly supported or conflicting. Cant and Cooper's (2010) performed a systematic review of 12 nursing simulation studies and reported on 11 assessed critical thinking (Alinier, 2007; Birch et al., 2007; Brannan, White, & Bezanson, 2008; Brown & Chronister, 2009; Griggs, 2003; Howard, 2007; Jeffries & Rizzolo, 2006; Linden, 2008; Ravert, 2004; Ruggenberg, 2008; Scherer, Bruce, & Runkawatt, 2007; Shepherd, Kelly, Skene, & White, 2007). However, seven of these studies used proxy subjective measures such as the student's self-reported confidence in their capacity to make clinical decisions (Alinier, 2007; Birch et al., 2007; Brannan et al., 2008; ; Griggs, 2003; Jeffries & Rizzolo, 2006; Ruggenberg, 2008;

Scherer et al., 2007). Linden (2008) used 23 knowledge and application questions to measure cognitive knowledge, which was seen as a precursor of critical thinking. There was a statistically significant change in the knowledge scores (Linden, 2008), however increased knowledge does not necessarily correlate to an increase in critical thinking ability. Howard's (2007) study showed a significant difference between the posttest HESI critical thinking scores of the simulation group at the $p = 0.051$ level but not at the $p < 0.05$ level. The control group watched a recorded presentation reviewing the care of a patient and worked through two case studies either alone or in small groups over the course of two hours (Howard, 2007). The reason that Howard found borderline significance may have been because the mean pretest critical thinking score of the control group was higher than the mean posttest score of the control group, the simulation group, and the adjusted posttest score of both groups.

Two of the three remaining studies in Cant and Cooper's (2010) review used objective measures of critical thinking but found no differences in critical thinking between the control groups and the experimental groups (Brown & Chronister, 2009; Ravert, 2004). Brown and Chronister's study used the critical thinking score from the ECG SimTest, which uses questions at the application level or higher. Ravert (2004) used both the CCTST and the CCTDI. Only one study showed a statically significant improvement for the patient simulation trained group and that study used clinical assessment scores as a proxy for critical thinking (Shepherd et al., 2007). In summary, Cant and Cooper's (2010) systematic review did not find that an HFPS intervention that used objective standardized tools that was able to measure significant improvements in critical thinking.

Levett-Jones, Gersbach, Arthur, and Roche (2011) found that critical thinking and clinical reasoning were associated with the ability to make sound clinical judgments as measured by the Structured Observation and Assessment of Practice. The Structured Observation and Assessment of Practice was designed to assess clinical competence using student narrative during their skills check off, and to encourage critical and reflective thinking. Students were assessed during two 3 hour patient care blocks (Levett-Jones et al., 2011). The evaluation of each of their care activities was structured according to the situation, action, and outcome. The situation, action, outcome model placed the student thinking and activities in context with actual patients and examined their knowledge, values, and attitudes through open-ended questions after completion of the observation period. The questions were designed to elicit "intentions, knowledge, rationales, attitudes and values" (p. 66) and support for claims of critical thinking, and clinical reasoning. The student's behaviors were then compared to competency standards for RNs. Both formative and summative feedback were given to the student during a 2 hour debriefing directly following the assessment. The focus of the formative feedback was on providing "individualised, detailed and non-threatening feedback" (p. 66) that identified strengths, weaknesses, and strategies for improvement. Students were encouraged to reflect and plan for improvement. Summative feedback was that the student had either been judged competent, competent once specific remediation had been completed, or not competent and requiring both remediation and reassessment (Levett-Jones et al., 2011). Although the situation, action, outcome format is both time consuming and educator intensive, it could be adapted to a HFPS scenario.

Critical thinking and reflective practice have been inexorably woven together (Dewey, 1933; Brookfield, 1987). The focus on critical thinking skills and dispositions has ignored that critical thinking is contextual and supported by reflective analysis (Brookfield, 1987; Kuiper and Pesut, 2004). Essentially, the critical thinking of nursing students must always be evaluated in the context of the unique patient and has not shown consistent improvement as measured by standardized testing (Boyd & Fales, 1983, Sedlak, 1997, Boyadjian-Samawi, 2006, Chabeli, 2007). Therefore, in order to promote the critical thinking of nursing students through the use of HFPS, reflective techniques and appropriate tools for measurement must be incorporated into the practice of nurse educators.

Use of reflective thinking activities in simulation. While reflection techniques in conjunction with nursing practice and clinical experiences have been extensively reported, there has been a paucity of articles describing the use of reflection with simulation in undergraduate programs. In a small pilot study of new graduate nurses, Stirling, Smith and Hogg (2012) used a training log to record directed reflections prior to beginning the simulation and answer a different set of reflective questions after the simulation. These logs were then used to guide the debriefing session (Stirling et al., 2012). Usually reflection has first taken place during the simulation itself, when students evaluate the results of their actions while the scenario was being run. The next time students used reflection was during the facilitator led debriefing activity that took place soon after the simulation was completed. Debriefing has been the most common faculty guided reflection activity; however, little research and fewer resources have been available for faculty to learn how to debrief to maximize student reflective learning

(Dreifuerst, 2009). The third time students used reflection was during extended reflection activities that occurred hours or days after the scenario was completed. Extended reflection has been a crucial but often neglected component of simulation activities (Jeffries, 2007). The ways to increase reflective thinking have been documented but have not been effectively used in simulation activities (Jeffries, 2007). The main difference between debriefing and extended reflection activities was that the information exchange between the student and the facilitator takes places hours to days after the simulation experience and the exchange was usually written down. This difference was significant because critically reflective writing encourages the development of metacognitive skills which are necessary when developing critical thinking (Fonteyn & Cahill, 1998).

Reflective thinking consisted a set of skills that has been used to build critical thinking abilities and promoted through the use of specific activities (Fonteyn & Cahill, 1998; Jeffries, 2007). Critical thinking as measured by objective tests has been used as a logical method of problem solving (King & Kitchener, 1994). Critical reflective thinking has been used as the process of reviewing an experience and making decisions about future actions based on lessons learned (Dewey, 1933). By participating in a reflective review of the external experience, internal thought, and emotive processes that took place during an experience, learners have been building their ability to critically think (Boyd & Fales, 1983; Sedlak, 1997). Therefore, each subjective contextual experience has the ability to "teach" through reflection.

Purpose of the Project

The purpose project of this project was to develop a guideline for designing HFPS to promote higher order thinking skills through the use of teach strategies and activities

designed to enhance student reflection. Exercises to optimize reflective thinking, and methods and mechanisms for evaluating reflective thinking will be extrapolated from all health professions' education simulation research. Due to the paucity of current research findings in the area of high fidelity patient simulation, additional sources of data will come from reflection activities used to shape other lower level simulated experiences.

Many authors believe that reflective thinking has been the basis for critical thinking (Boyd & Fales, 1983; Brookfield, 1987; King & Kitchener, 1994; Navedo, 2006; Sedlak, 1997; Wallace, 1996). Therefore, the focus for advancing critical thinking should be on encouraging reflective thinking during all five phase of the simulation: briefing, the running of the scenario, debriefing, extended reflection, and during evaluation by promoting reflection before the next experience. Exposing students to the reflective process increases awareness, and may result in the student using the process intentionally and discovering its value as a learning tool (Boyd & Fales, 1983). Teaching students to use reflective thinking assists in their developmental progression, and over time leads to even more effective use of this tool (Boyd & Fales, 1983).

PICO Question

The PICO question format was used to guide the search for evidence. The PICO question to be answered was: What is the best way for nursing faculty to maximize undergraduate students' reflective thinking in the course of high fidelity human patient simulation activities as compared to current practice in simulated patient experiences in healthcare pre-professional programs? The P in PICO stood for population. The I stood for intervention. The C in PICO stood for the comparison intervention, while the O stood for the outcome.

Population description. The population for this project was defined as the instructors of nursing students who have not yet completed their first nursing degree. Nursing faculty in traditional and accelerated baccalaureate as well as associate's and diploma programs were included.

Intervention description. The intervention was defined as reflective thinking activities before, during, and after a HFPS experience. Reflective thinking has been defined in a variety of ways and different authors emphasize different parts of the process. For the purposes of this paper, reflective thinking was the habitual process of intentional and unintentional mental examination, either in the midst of reacting to an event, processing a past event, or for planning of responses to future events. Reflective thinking activities were any instructor designed event, activity, or assignment which was meant to encourage reflective thinking in the student, before during, or after the simulated patient experience. Examples of reflective thinking activities used with simulated events have been: Socratic questioning, thinking aloud on the part of the student, pausing the simulation, journaling, blogs, wikis, and role playing. Simulated patient care experiences have taken many forms: case studies; interactive computer programs; standardized patients; task trainers; low, medium, and high fidelity patient simulators; and supervised student experiences (Alinier, 2007). Any form of simulated patient experience that has employed strategies to motivate students to reflect on their thoughts, feelings, and actions was reviewed.

Comparison intervention description. The comparison intervention was defined as simulated patient experiences, that have taken place in classroom, lab, or clinical and that did not specifically incorporate reflective thinking activities. These

simulated patient activities were: case studies, virtual patients, standardized patients, task trainers, supervised clinical experiences, or low, medium, or high fidelity patient simulation (Alinier, 2007). Research from healthcare pre-professional programs were evaluated including: nursing, medicine, dentistry, physician assistants, pharmacy, and physical, occupational, speech, music, and respiratory therapy.

Outcomes description. The defined end result desired outcome was an improved ability to think reflectively about simulated patient situations. Unfortunately, the desired outcome may take years to be realized. Firstly, reflective thinking has been defined as a partially developmental process that takes many years to hone (King & Kitchener, 1994). Secondly, this guideline concerns student nurses, who have had only limited opportunities to experience patient situations, in which to develop their reflective thinking ability. Therefore, the critical outcome was the ability to demonstrate reflective thinking before, during, and after a HFPS experience. Due to difficulty in measuring thought processes, researchers have used proxy measures of reflective thinking processes to determine progress towards the critical outcome. Proxy measures have been: interviews, transcripts, writing samples, behavior checklists, and audio- or video taped simulations or debriefings. In addition to the critical outcome, other important outcomes have been measured by researchers. Many of the important outcomes have been subjective measures of the student's or instructor's opinion. Examples of subjective measures that researchers have used are: either the instructor's belief or the student's improved self-confidence in the student's enhanced ability to make clinical decisions or clinical judgments, or to reason clinically.

Evidence Search Process

I began the search for evidence relevant to this HFPS higher order thinking guideline in a graduate course in 2009. A concurrent search of critical thinking in new graduate nurses revealed that critical thinking seemed to be the province of the competent and/or proficient nurses as their thinking and knowledge development was described by Benner (1982, 1984). Changes in critical thinking would then be out of reach for the student (novice) or new graduate nurse (advanced beginner). Continued research into the area revealed that reflective thinking has been considered to be a stepping stone for critical thinking. I decided to refocus the guideline on reflective thinking after evidence was found that reflective judgment improves measurably during the course of undergraduate education and was part of the foundation for critical thinking (Brookfield, 1987; Dewey, 1933).

Determining the Depth and Breadth of the Literature Review

An EBSCHO search of CINAHL, MEDLINE, ERIC, and PsychINFO for articles with the subject headings *simulation* and *nursing education* revealed a plethora of evidence (1,343 articles). However, a paucity of evidence was identified that had *reflection* (8) as an additional subject heading. Since there were so few pieces of evidence found on this initial search, I decided to remove *Nursing education* as a search term. The search was expanded search to include all evidence concerning first time professional health related programs designed to work with students, whether graduate or undergraduate.

Then the EBSCHO databases were searched using reflecti*, education, and simulation as subject terms, without a date limit, and the first source that involved

health professions was dated 2005. The earliest result of the next EBSCHO search of the same four databases using reflecti* and simulation as subjects was also published in 2005. Therefore, a publication date delimiter of thirteen years was used because it represents the approximate span of time HFPS has been studied in undergraduate nursing education and includes a five year margin of error for the earliest found items in the preliminary searches.

Summary

High fidelity patient simulation has been an expensive and time consuming teaching tool in undergraduate nursing education. Best practices must be used in order to warrant the cost in time and money needed to run HFPS. Task trainers, and low and medium fidelity patient simulators justify their expense by teaching nursing students psychomotor skills and rule-governed behaviors. However, HFPS must show a return on investment that justifies their greater expense. Improving the critical thinking skills of nursing students has been one goal mandated in undergraduate nursing education (AACN, 2008; National League for Nursing Accrediting Commission, 2006). Sadly, the evidence has not objectively proven that HFPS improves students' critical thinking skills (Cant & Cooper, 2010). Whether this has been due to, not having a standardized tool that measures changes in nursing students' critical thinking, successful interventions, student developmental levels, or another reason, is not currently known. Perhaps it is time to concentrate on the higher order thinking skills that build critical thinking. Reflective thinking has been believed to be a precursor to improvement in critical thinking (Brookfield, 1987; Dewey, 1933; R. A. Kuiper & Pesut, 2004; Sedlak, 1997). Activities that would improve reflective thinking in HFPS have either been omitted or received

little attention when compared to the scenario experience (Jeffries, 2007). A search of articles for nursing simulation studies that were centered on improving reflection in simulation revealed only eight out of over thirteen hundred articles. Therefore, the search was expanded to pull together evidence from all health professional education profession programs that used reflection activities in simulated patient activities. This guideline would show how to best use and measure reflective thinking within HFPS in order to build undergraduate nursing students' higher order thinking skills and train them in reflective techniques that could potentially advance their professional practice.

CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

Higher order thinking skills, including critical and reflective thinking, in college students was first extensively examined by Dewey (1933). While Schon (1983, 1987), Dewey's student, explored the reflective thinking piece of Dewey's work; Facione (1990) looked at the other component of Dewey's work and developed a comprehensive definition of critical thinking. The National Council of the State Boards of Nursing (2012) and AACN (2008) have chosen to promote the critical thinking side of Dewey's work as an integral part of nursing education. Unfortunately, an improvement in critical thinking of nursing students over the course of their education has not been consistently documented. Part of the problem with trying to measure critical thinking in undergraduate nursing students may be because changes in critical thinking are out of reach of the beginner and novice nurse (Benner, 1984). Equally concerning about promoting critical thinking in undergraduate nursing education has been that the most widely used standardized objective measures of critical thinking, the CCTSI, CCTDI, and WGCTA, may not be the best measure of critical thinking as it is used by nursing students to support clinical reasoning and clinical judgment (Stone et al., 2001).

By comparison reflective thinking has long been embraced by nursing education and over time many methods have evolved to encourage or record students' reflective thinking. Reflective thinking changes have been consistently measured in undergraduate

students (King & Kitchener, 1994) as well as in nursing students (Navedo, 2006) but methods of objectively evaluating reflective thinking have not been fully utilized in HFPS. Further, it is believed that reflective thinking promotes the clinical reasoning and clinical judgment capabilities of nurses (Kuiper & Pesut, 2004). Therefore, this literature review has included reflective thinking literature concerning the education of health professionals in simulated patient experiences that can be utilized by the nurse educator in conjunction with HFPS. This search has sought to gather together the best practices in motivating and guiding students to reflectively think, and assessing reflective thinking in nursing students.

Inclusion and Exclusion Criteria

Since I do not speak another language, all evidence not written in English was excluded. Items about reflection by physician residents, practicing nurses, or nursing graduate students were excluded since they have already been licensed to practice. The decision not to include reflection by licensed nurses may appear in conflict with the inclusion of articles from students of other healthcare programs, particularly those at the post baccalaureate level, for example: medicine and dentistry. The reason for this delineation was that this guideline was focused on the facilitation of reflective thinking in students preparing for professional practice.

After preliminary review of the body of evidence, many types of items were also excluded from the evidence table. Items by healthcare professionals who typically did not have direct patient contact, for example health information management or health administration, were excluded since the focus of this guideline was in the area of patient simulation. Items that were about theory construction and concept analysis were

excluded, so that the guideline could be formulated on evidence that was drawn either from the reported direct experience of the authors, or a review of the evidence that could be replicated. Expert opinion items were excluded for the same reasons.

Literature Review

The search of the EBSCHO databases using reflecti*, education, and simulation as subject terms turned up 83 items of which 14 were retained after abstract review. An EBSCHO search of the same four databases using reflecti* and simulation as subject terms and patient as a text term found 34 articles and six were retained. Fifty-two articles were found in a search of the ProQuest Health and Medicine databases using the subject headings reflection or reflective thinking and simulation. Three of the ProQuest articles were not duplicate findings and were suitable for further consideration. A search of PubMed using the MeSH terms patient simulation and thinking with reflecti* in the text identified 10 entries. Five entries were retained for further investigation. A second search of PubMed using the MeSH terms patient simulation and reflecti* as a title or abstract word identified 96 entries. Thirteen articles were retained after abstract review.

A search the Joanna Briggs Institute website turned up no results for reflection or reflective thinking. Since the application of critical and reflective thinking to nursing situations results in clinical reasoning and clinical judgment, these search terms were added. This was so that studies would be included that focus on the use of higher order thinking skills to solve nursing problems. However, no results were included for further review. The search of Health and Psychosocial Instruments (HaPI) was conducted in a similar fashion as the Joanna Briggs Institute. As before, no results were found using reflective or reflection as subject terms. When searching HaPI using reflective or

reflection as the search terms in the abstract, 79 results were found. However, many of the results concerned the reflections of patients or teaching counselors to use reflective prompts with their patients and no tools were selected for inclusion. In an effort to find any relevant articles, the search of HaPI was expanded using clinical judgment as a search term, 41results were returned. However, just four tools measured clinical judgment in the education of health professionals and none were suitable for inclusion. Only three results, were returned when clinical reasoning was used as a search term in HaPI, and the one tool worthy of inclusion was a repeat from previous searches. A search of HaPI using clinical decision making as the search term had 26 results but no new tools were found.

Preliminary reviews of abstracts contained in the nursing, education, medical, and psychological databases revealed enough relevant articles to form the basis for a guideline. An additional 500 articles were skimmed or read and 21 were retained for further review. Many of those 500 articles were the result of hand searches of non-indexed journals, Google Scholar searches, reviews of citations, and related references in articles. An additional source of articles was the use of a PubMed's function that allows the researcher to find additional PubMed articles that have cited the source article. This was especially useful since reviews of citations allow the researcher to look back from the publication date and the PubMed function to look forward from the publication date of the source article. A preliminary search was conducted using the terms reflective thinking, reflection, and simulation. Some types of simulated patient experience had not used the term simulation when the articles had been indexed in databases. Therefore, alternate terms describing simulated patient activities were used: case study/studies, task

trainer, virtual patient, standardized patient, standardised patient, clinical, and clinical supervision. All of these types of experiences have been used to simulate a portion of the patient experience and should be considered simulation (Alinier, 2007).

The CINHAL, Medline, ERIC, Psych Info, ProQuest Health and Medicine, and Pub Med databases were searched again using alternate terms for simulated patient experiences. No further review of the Joanna Briggs Institute or the HaPI database was needed since the search term simulation was not used as a delimiter in the prior searches of those resources. The second search of CINHAL, Medline, ERIC, Psych-Info using reflecti*, education, and case study as subject terms had 409 results. However, only one article, by Ladyshewsky and Gardner (2008), was retained after abstract review. Only one additional article was found when task trainer was used as a subject term in addition to reflecti* and education, but it was not retained. Using the term standardized patient as the additional subject term revealed 16 results, of which four were retained. Using the British spelling of standardized patient found one article, by Plack, Dunfee, Rindflesch and Driscoll (2008), which was retained. The term clinical had the most results, 580 items, and 96 were retained for additional review. Using the term clinical supervision resulted in 76 articles and nine were retained.

The ProQuest Health and Medicine database was searched with clinical and reflection or reflective thinking as subject terms and an additional 85 entries were found. Eight entries were retained after reviewing the abstracts. When searching using clinical supervision and reflection or reflective thinking as subject terms, two articles were found but neither was retained. No results were found when reflection or reflective thinking

was used as subject terms and any of the following subject terms: case study, virtual patient, standardized patient, and standardized patient.

The PubMed database was searched again using reflecti* as a title/abstract word and other terms that might reveal different types of simulated patient experiences. When case study or case studies were used as MeSH terms, no articles were found. Task trainer, virtual patient, standardized patient, clinical and clinical supervision were not listed as a MeSH term so the database was searched using these terms as a title/abstract words in addition to reflecti*. No articles were found when task trainer was used. Three articles were found when virtual patient was used but were not retained. When standardized patient was used 27 articles were found and five were retained. One article was found using standardized patient as a search term but the article was not retained. When clinical or clinical supervision was used 708 articles were found and 38 were retained.

Overall, 2,337 entries were found in the multiple searches, although many were duplicate items. A total of 225 times retained for additional review. Multiple searches of the higher order thinking literature in healthcare uncovered over 500 pieces of evidence of which 21 were retained for further consideration as part of the evidence base for this paper.

Development of Evidence Table

An evidence table was created to systematically and critically appraise the articles. The table allowed significant elements to be reviewed, rated, and recorded. By organizing the evidence in a table all the articles could be easily compared and contrasted

using the same criteria. Different criteria checklists were created for different types of articles.

Rating the evidence. All evidence was then rated using the Scottish Intercollegiate Guidelines Network (SIGN, 2011) levels of evidence contained in Appendix C. The rating of each study is contained in the evidence table in Appendix D along with the review and summary. Evidence was rated on a scale of 1++ to 4 (SIGN; n.d.). A level 1++ indicated a high quality meta-analysis, systematic reviews of randomized controlled trials, or randomized controlled trials that had a low risk of bias (SIGN; n.d.). The lowest level of evidence was expert opinion which was rated as 4 (SIGN; n.d.). No expert opinion evidence was used in the evidence table. Therefore, the lowest level of evidence that was used in the table was level 3, which consists of non-analytic reports. The bulk of the evidence found was descriptive studies that contained primarily subjective opinions of the students and faculty. I used different types of criteria to consider and rate the different types of study. Next, I will discuss in detail the methods I used to evaluate each of the different type of studies: systematic review, randomized controlled trials, cohort studies, non-analytic studies and mixed methods studies, and qualitative studies.

Systematic review appraisals. The Critical Appraisal Skills Programme (2013c) systematic review appraisal tool was used to evaluate the systematic reviews. Systematic reviews were considered for evaluation if they contained a clearly-focused purpose that addressed the PICO question (CASP, 2013c). Next, I determined if all the relevant studies could have been found using the search methods that the researcher described and if the researchers assessed the quality of each of the studies (CASP, 2013c). After that, I

considered whether the results of the studies reviewed had been combined and whether this was an appropriate measure (CASP, 2013c). Then, I looked at how the results were organized, determined how important the results were, and wrote a synopsis of the results (CASP, 2013c). Next, I considered how precise were the results and how confident I was that the study achieved the correct conclusion as a result of their findings (CASP, 2013c). Then I decided whether all important outcomes had been considered for the student, faculty, school, clinical sites, and patients and if these result could be applied to traditional nursing students in the United States (CASP, 2013c). Lastly, I considered whether current practice should be changed by the findings of the systematic review (CASP, 2013c). Based on the results of my review of the study, I rated the systematic review as: 1+, 1-, or 2+. The 1++ designation was not used since no systematic reviews of randomized controlled trials was found in the literature review (SIGN, 2011).

Randomized controlled trials appraisal. Randomized controlled trials were rated according to the Critical Appraisal Skills Programme (CASP, 2013b). First, I decided whether the trial addressed an issue that was closely aligned to my PICO question, if the intervention assignments were randomized, and if the outcomes of all participants in the trial were analyzed. These screening questions determined if I continued with analyzing the trial. Next, I looked at how the study was conducted. Was blinding used to screen the students, instructors, and researchers from the intervention? Did the researchers determine if the groups were similar at the beginning of the trial and were any attempts made to try to balance the control and intervention groups? The last question in this section was if the results of the trial were valid and if the control and intervention groups were treated as similarly as possible. The next five questions I used

to help me gauge the effect of the results. I looked at how large the effect of the outcome was, what the confidence limits were, whether the results were applicable to an undergraduate school of nursing in the United States, whether all practice and educational outcomes were considered, and if the benefits of the intervention was worth the time, effort, and costs (CASP, 2013b).

Quasi-experimental studies. The checklist created by Downs and Black (1998) was used to guide the evaluation of the quasi-experimental. Since the 27 questions in the checklist would create an unmanageable evidence table, the findings from using the checklist were recorded under appropriate headings in the evidence table. Questions such as: "Was an attempt made to blind study subjects to the intervention they have received?" and "Was compliance with the intervention/s reliable?" were skipped because they were not appropriate for this type on intervention (Downs & Black, 1998).

Cohort studies appraisal. The Critical Appraisal Skills Programme (CASP, 2010) cohort checklist was used to evaluate cohort studies. First, I determined whether the study addressed my PICO question and whether a cohort study was an appropriate method to use. If I was able to answer these two questions in the positive then I continued to evaluate the study. Next, I looked at how the cohort was recruited and whether it was a representative sample of the population. After that, I looked at the measurement tools' validity and reliability. This impacted the next question to be answered, whether the outcome was measured in such a way as to minimize bias. Confounders were the next factor that I considered. I looked to see if the authors had identified important confounding factors and attempted to control or minimize the confounders. Then, I considered whether the follow up period was an appropriate length

and if members of the cohort lost to follow up were different from the sample. The last four items I examined concerned the results of the study. Basically, I first looked for what was the result and how strong the association was between or among the factors. I noted the size and range of the confidence interval. I looked for other possible explanations for the results: bias, chance, confounding, poor methods, inappropriate design or other flaws. Then I determined whether these results could be used in a HFPS with undergraduate nursing students. Lastly, I explored how these results fit with the other available evidence (CASP, 2010).

Mixed methods, non-analytical, and quantitative descriptive studies appraisal. Mixed methods, non-analytical, and quantitative descriptive studies were evaluated using the Evaluative Tool for Mixed Method Studies (Long, 2005). The tool allowed me to review the parts of the study that were included and skip areas that were not addressed. First, an overview of the article was established by answering five questions. The next set of questions concentrated on the type of study, the intervention(s), the level of detail, and the relationship of the study to my PICO question. Then, the setting, sample, and outcome were described and evaluated. The ethics of the study were then evaluated. If the study used groups, then the comparability of the groups was investigated. If there was a qualitative component, the data collection and data analysis methods, and potential researcher bias were reviewed. The implications of the study for education and practice were determined. Lastly, in other comments deemed important or unique to the study were recorded (Long, 2005).

Qualitative studies appraisal. If the study was a qualitative study, then the Qualitative Research Checklist (CASP, 2013a) was used to evaluate the study. First the

study was evaluated for appropriateness and applicability to the PICO question. If I determined that the evaluation was worth continuing, then the suitability of the research design was assessed. Then the recruitment methods were examined and compared to the aims of the research to decide if the two were well matched. The methods of data collection were reviewed to decide if they addressed the research issue. After that, I tried to determine the relationship between all members of the research team and the participants and decide if there were any concerns about bias or influence. I looked at how the researchers handled potential ethical issues and if an ethics committee or similar oversight had been sought before beginning the research. Then, I examined whether the data had been thoroughly analyzed, if contradictory findings were addressed, how data was organized into themes or categories, and if the researcher(s) examined their own input for possible sources of bias. I determined if the findings were explicit and clear, and explained in relation to the original aims of the study. Lastly, I looked to see if the researchers placed their finding in context with the current evidence, identified new areas of research, and discussed how the research could be used in other contexts (CASP, 2013a).

Summary

Although critical thinking has been mandated in baccalaureate nursing education (AACN, 2008; National League for Nursing Accrediting Commission, 2006), the focus should be on reflective thinking skills that lead to critical thinking in the professional nurse. Evidence examining reflective thinking in simulated patient experiences has been compiled from a variety of pre-professional healthcare programs. However, reflective thinking exercises have not been fully utilized within HFPS. A review of the relevant

literature using databases in the areas of nursing, allied health, medicine, education, and psychology was undertaken. No new tools were found in the HAPI database or within the Joanna Briggs Institute collection. To maintain consistency in the evaluation of the evidence, the CASP (CASP, 2010, 2013a, 2013b, 2013c) tools were used whenever possible. However, the mixed methods studies were evaluated using the Long (2005) instrument. The SIGN (2011) criteria were used to rate the evidence on a standardized scale. The evidence was winnowed to 83 studies that were compiled in the evidence table (see Appendix D) based on Downs and Black's (1998) quantitative guidelines and CASP's (2013a) qualitative guidelines. Much of the evidence was not from high fidelity simulation experiences but was extrapolated from other forms of simulated patient experiences that run the gamut from case studies to supervised clinicals.

CHAPTER 3

LITERATURE ANALYSIS AND SYNTHESIS

Introduction

There are many opportunities for nurse faculty to maximize undergraduate nursing students' reflective thinking with HFPS activities. Jeffries (2007) stated that a simulation activity should consist of: briefing, scenario, debriefing, extended reflection, and evaluation. However, a separate pre-briefing or orientation can prepare students for what to anticipate from and expect of the simulation experience (Lasater, 2007b). Jeffries (2007) included the orientation and instructor rehearsals in the category of briefing, while I use the term briefing to refer to scenario specific directions and reminders. Although the major focus has been on reflective exercises that take place during the debriefing and extended reflection phase of HFPS activities; many articles were reviewed that discussed elements that are necessary to creating a learning experience that enables critically reflective thinking. This chapter examined the evidence that reported the best practices in simulated patient activities as related to the promotion of critical reflection.

Preparation of the Student for HFPS

Faculty need to prepare students for simulations including an assessment of the students' knowledge that is needed for the scenario and an estimate of the students' reflective thinking abilities. Decker's (2007) mixed method study found that the students' level of reflective thinking correlated with their ability to complete the

simulation (2007). The descriptive evidence posits that it is necessary to prepare and assess students for the simulation experience and the planned reflective tasks (Cahalin, Markowski, Hickey, & Hayward, 2011; Corrigan & Hardham, 2011; Delany & Watkin, 2009; Hatlevik, 2012; Lasater, 2007b; McMahon, Monaghan, Falchuk, Gordon, & Alexander, 2005; Perera, Mohamadou, & Kaur, 2010; Thompson et al., 2010; Tofil, Benner, Worthington, Zinkan, & White, 2010). Students reported that a general orientation, that went over what to expect during the simulation and what was expected of the students, was seen as bringing all students to the same level of readiness for the HFPS (Lasater, 2007b). More specific preparation designed to insure that students had the skills and the ability to recall and understand the knowledge needed in the simulation has taken the form of: a review of material, interactive exercises, testing, training, handbooks, or guidelines (Cahalin et al., 2011; Corrigan & Hardham, 2011; Delany & Watkin, 2009; McMahon et al., 2005; Perera et al., 2010; Thompson et al., 2010; Tofil et al., 2010).

While many authors have documented the need to prepare and assess students for a simulation (Cahalin et al., 2011; Corrigan & Hardham, 2011; Delany & Watkin, 2009; Hatlevik, 2012; Lasater, 2007b; McMahon et al., 2005; Perera et al., 2010; Thompson et al., 2010; Tofil et al., 2010) the body of evidence for this was generally descriptive. The only report of a correlation between the level of students' reflective thinking and the students' ability to successfully complete a simulated scenario was in Decker's (2007) mixed methods study. The evidence did not contain any studies that examined student assessments and that determined if higher scores were associated with greater learning from the simulated experience. A multifactorial correlational analysis would be helpful

in deciding which preparation activities would be most beneficial and if an assessment of students' knowledge base, skills, or reflective ability would be useful in determining if students' preparation had primed them for optimal learning during the HFPS.

Simulation Design

Faculty need to carefully design all aspects of the simulation with the goal of maximizing the opportunity for student reflective thinking. In Blatt, Plack, Maring, Mintz, and Simmens' (2007) cohort and Lasater's (2007b) descriptive studies, students were able to improve their performance by either revisiting the same or similar patients. Blatt et al. used a convenience sample of 149 third year medical students, but not all students chose to revisit the patient in an attempt to improve their patient satisfaction or skill score. However, the students that did revisit a standardized patient showed an inverse association between initial score and the amount of improvement (Blatt et al., 2007). To clarify, if students performed poorly on an assessment and chose to revisit the standardized patient, they had a large increase in their scores (Blatt et al., 2007). The average change in scores was much smaller than the standard deviation and the association may have been due to regression to the mean or to self-selection bias in the students' choice to revisit a standardized patient (Blatt et al., 2007). Strengths were that the skills and patient satisfaction checklist had face validity and several researchers reviewed and coded the data (Blatt et al., 2007). This study was marred by letting students choose which patients to revisit, which effectively skewed the second score. Possibly, students who felt that they could score much better on a revisit chose to revisit a patient, while students who felt they could not improve their score did not revisit. Using

a control group or having a random assignment of revisits would have made this study more rigorous.

A retrospective cohort study conducted by Cook (2010) used reflective journals throughout clinical courses. The journals had no stipulations on content and students received little guidance. A three level reflection rating was used to score journals. Seventy-five records of physical therapy students who had graduated and taken the National Physical Therapy Exam were examined. Over 900 journal entries were reviewed by three coders. Inter-rater reliability was .849. No correlations were found between reflective writing levels and the National Physical Therapy Exam or scores on the Clinical Performance Instrument.

Numerous studies described how the scenario and reflective experiences were built on information the students already knew and experiences the students had already had (Blatt et al., 2007; Bruce, Parker, & Herbert, 2001; Cahalin et al., 2011; Corrigan & Hardham, 2011; Daly, 2010; Ertmer et al., 2010; Lasater, 2007b; McMahon et al., 2005; Thompson et al., 2010). Another common practice was to have additional students observing or participating in scenarios and many researchers have found that when two or more students participated in the simulation, students: learned more, practiced team building, and practiced working with simulated professionals and family members (Bruce et al., 2001; Cahalin et al., 2011; Corrigan & Hardham, 2011; Daly, 2010; Ertmer et al., 2010; Lasater, 2007b; Lindgren & Athlin, 2010; Perera et al., 2010; Thompson et al., 2010).

The strongest evidence shows that students were able to improve skill and patient satisfaction scores when they chose to revisit the same (Blatt et al., 2007) or similar

patient scenarios (Lasater, 2007b). The evidence was generally non-analytical and described the roles of student observers or participants in the scenario and the importance of linking a scenario to the curriculum. A higher level of evidence for designing the whole simulation experience to promote reflective thinking and not just during debriefing or an extended reflection activity is needed. In order to assist students in becoming reflective practitioners, the habit of reflecting before, during, and after patient experiences must be ingrained during their education. Analytical studies are needed that demonstrate what factors in simulation preparation, execution, and follow-up are most promising in raising students' reflective levels and inculcating the reflective mindset.

Recording the Process

Faculty need to videotape or otherwise record the simulation processes. Maloney, Stoor, Morgan, and Ilic's (2013) randomized controlled trial of 60 third year physiotherapy students found that students who reviewed simulation videos were able to reflect and monitor their progress and performed better on a related Objective Structured Clinical Examination. Both students and peer reviewers believed that video recording was helpful for review and analysis and that the review assisted students to identify errors and areas for improvement in both verbal and non-verbal communication (Maloney et al., 2013). Review of the videotaped scenario has been helpful to students both when they had a role in the scenario and when they were observers (Corrigan & Hardham, 2011; Daly, 2010; Hulsman, Harmsen, & Fabriek, 2009; Hussin, 2013; Kalish, Dawiskiba, Sung, & Blanco, 2011; Lasater, 2007b; Maloney et al., 2013; Thompson et al., 2010).

Jarris, Saunders, Gatti, and Weissinger's (2012) quasi-experimental pre-test post-test study found no significant difference between the control and intervention groups'

clinical skills assessments. The control group was comprised of 153 first year medical students who completed two clinical assessments three months apart. The intervention group of 47 students: reviewed a videotape of their first clinical assessment of a standardized patient, completed a self-assessment, received feedback from the standardized patient, and instructor verbal comments (Jarris et al., 2012). Faculty were able to review the video recordings and provided students with additional written feedback (Jarris et al., 2012). The researchers felt that the lack of difference between the groups may have been due to a lack of guidelines and instruction on critical reflection (Jarris et al., 2012). A limitation of this study was that there was no discussion of how the students were assigned to the intervention and control groups (Jarris et al., 2012). Several descriptive studies have used a review of the taped debriefing to assess student reflection levels as well as to evaluate the facilitator (Brown, 2011; Delany & Watkin, 2009; Duggan, Bradshaw, Carroll, Rattigan, & Altman, 2009). Additionally, preserved material from student completed activities could have been used to establish baselines, gauge progress, note missing skills or knowledge, and identify gaps in the curriculum (Cahalin et al., 2011; Flanagan, Nestel, & Joseph, 2004; Harrison & Fopma-Loy, 2010).

Two strong studies had conflicting evidence on the value of student reviewing the videotaped scenario (Jarris et al., 2012; Maloney et al., 2013). The review of the videotape needs to be accompanied by instruction and guidelines on how to critically reflect (Jarris et al., 2012). However, there is a large body of descriptive evidence supporting videotaping the scenario either for student or faculty review (Corrigan & Hardham, 2011; Daly, 2010; Hulsman et al., 2009; Hussin, 2013; Kalish et al., 2011; Lasater, 2007b; Thompson et al., 2010). Fewer studies focused on taping the debriefing

(F. S. Brown, 2011; Delany & Watkin, 2009; Duggan et al., 2009). No studies described documenting the orientation or briefing for later analysis. Videotaped orientations and briefings could assist in standardizing the student experience and preserving the most helpful elements. Although several studies mentioned the value of preserving documentation of students' work, all were descriptive in nature (Cahalin et al., 2011; Flanagan et al., 2004; Harrison & Fopma-Loy, 2010). More analytical studies are needed that concentrate on the value of retaining recordings and documentation for later analysis that may reveal areas that need improvement as well as previously successful strategies.

Safe Environment

Faculty need to conduct all simulation activities in a psychologically safe environment. Epp's (2008) systematic review examined the use of reflective journaling in undergraduate nursing education. One hundred and fifty abstracts were reviewed from articles indexed in the OVID, EDSCO, or Blackwell Synergy database and published from 1992 to 2007 and nine studies were analyzed (Epp, 2008). One article that Epp reviewed reported trust was a key part of reflection; for not only did journaling raise levels of trust, but as levels of trust rose so did the students' self-disclosure (Landeen, Byrne, & Brown, 1995). In addition to the findings of the systematic reviews, a number of descriptive studies found that a psychologically safe environment made for a good reflective environment (Becherer, 2011; F. S. Brown, 2011; Donovan, 2007; Harrison & Fopma-Loy, 2010; Lutz, Scheffer, Edelhaeuser, Tauschel, & Neumann, 2013; Manning, Cronin, Monaghan, & Rawlings-Anderson, 2009) or a good learning experience (Ekebergh, 2007; Ladyshewsky & Gardner, 2008; Lutz et al., 2013; McMahon et al., 2005).

The highest level evidence for a psychologically safe environment was a systematic review of undergraduates (Epp, 2008) that contained a study that specifically looked at undergraduate nursing student' levels of trust as it related to reflective journaling (Landeen et al., 1995). There was a plethora of descriptive studies stressing the importance of a psychologically safe environment; but there are no analytical studies that showed which interventions correlated with the students' feeling of safety. Analytical research studies are needed that are designed to test interventions that may increase trust within the simulation laboratory and during reflection exercises. Otherwise simulation laboratory procedures and practices will be based on opinion and observation without definitive evidence.

Facilitator Training and Evaluation

Faculty need to provide education, training, and materials; and evaluate facilitators that conduct the scenario, debriefing, and extended reflection activities. Hallmark's (2010) quasi-experimental study used either trained or untrained faculty for debriefing. A convenience sample of 84 nursing student volunteers, out of a cohort of 157 third year nursing students, was randomly assigned to either the intervention or control group. The groups showed no difference in HESI scores. Although the HESI is a valid and reliable tool, it was not designed to measure reflective thinking and may not have been the best measure of a change in reflective thinking. Hallmark noted that having a trained faculty debriefer resulted in higher student satisfaction scores after controlling for: age, gender, grades, and educational level. Additionally, students of the trained faculty rated themselves significantly higher on the Reflective Learning Continuum Likert scale (Hallmark, 2010).

In a pre-test post-test study, Ip et al. (2012) used 4.5 hours of interactive teaching that covered both the theory and application of reflective learning and small group discussion of a videotaped vignette to prepare students for reflective writing assignments. Following the education intervention students participated in 4 weeks of clinicals with instructor facilitation of self-reflection Ip et al., 2012). One hundred and seventy-three nursing students participated in the interventions (Ip et al., 2012). Only 38 students completed all the learning activities and turned in all three reflective journals after the educational intervention, 2 weeks of clinicals, and 4 weeks of clinicals (Ip et al., 2012). A post-test survey revealed that the students considered the role of faculty very important to gaining self-reflective ability (Ip et al., 2012). One of the barriers mentioned by students' in their open ended responses was that faculty were not always available to assist with self-reflection (Ip et al., 2012). Overall, students were able to significantly improve their level of reflective writing after two weeks of faculty facilitation in the two weeks from pre-test to post-test one (Ip et al., 2012). However, students did not significantly improve after two additional weeks of facilitation at post-test two (Ip et al., 2012). There was 95% inter-rater reliability between the two coders on the level of reflection: non-reflective, reflective, or critically reflective (Ip et al., 2012). The Friedman test was used to prove the statistical difference between the mean scores (Ip et al., 2012). The Wilcoxon signed-ranks test was used to compare an individual's scores over the three measurements: pre-test, post-test 1, and post-test 2 (Ip et al., 2012). The Friedman test and the Wilcoxon signed-ranks test were appropriate measures since the population was not normally distributed (Green & Salkind, 2008). Many other descriptive studies supported the premise that facilitation is a skill that needs to be taught

and assessed (F. S. Brown, 2011; Delany & Watkin, 2009; Ekebergh, 2011; McMahon et al., 2005; Murphy, 2004; Skovsgaard, 2004) and that facilitators are key to debriefing and reflection (Decker, 2007; Donovan, 2007; Ker, 2003; Lasater, 2007b; Manning et al., 2009; O'Donovan, 2006).

Hallmark's (2010) highly rated study had objective evidence that did not support faculty training and subject evidence that did support faculty training. However, the objective measure used by Hallmark (2010), the HESI, may not have been a good proxy measure of reflective thinking. Ip et al.'s (2012) equally highly rated study used an evaluation of the students' writing, a more appropriate measure of reflective thinking, and found student improvement with trained faculty. The coding of student's writing used by Ip et al., while a more subjective measure, may have reached a closer approximation of the students' reflective thinking level. Both measures are an improvement on students' self-rating on scales and the opinions of student and faculty that comprise the bulk of the evidence for using trained faculty (Brown, 2011; Decker, 2007; Delany & Watkin, 2009; Donovan, 2007; Ekebergh, 2011; Hallmark, 2010; Ker, 2003; Lasater, 2007b; Manning et al., 2009; McMahon et al., 2005; Murphy, 2004; O'Donovan, 2006; Skovsgaard, 2004). More quasi-experimental studies that use control groups, and pre and post testing are needed to evaluate the effect of trained faculty on the students' reflective experience. Additionally, changes in the students' reflective thinking should be measured through evaluations of the students' writing, speech, and behaviors and not by standardized tests designed to measure related concepts.

Debriefing

Facilitators should conduct an immediate debriefing, in a different area than the scenario, which should include: simulation anomalies; affective and cognitive content; a summary; and a focus on student learning, gaps in knowledge, learning process, and goals for future improvement. In a quasi-experimental pre-test post-test study, Dreifuerst (2012) used the Debriefing for Meaningful Learning method to implement guided reflection. Students were randomly assigned to the intervention or control group based on their clinical group. The control group received the standard debriefing method. The researcher provided debriefing for all of the intervention groups. The Debriefing for Meaningful Learning method begins with addressing the affective component and then moves to analysis of the scenario. To assemble 240 participants, student volunteers from three consecutive semesters were recruited. Statistical analysis showed that the three sets of students were homogeneous and able to be combined into one sample. The study was limited by self-selection bias. Only two students were lost to follow-up, making the final sample 238 (Dreifuerst, 2012).

The Health Sciences Reasoning Test (HSRT) was used pre-test and post-test, and given three weeks before and after the HFPS (Dreifuerst, 2012). Two additional instruments were given post-test to the intervention group to measure student satisfaction with additional elements in the simulation: the Debriefing Assessment for Simulation in Healthcare-Student Version and the Debriefing for Meaningful Learning Supplement Questions. The control group was not given the survey questions. This made comparison of the two groups on those two measures impossible and might have introduced a Hawthorne effect. Students who were in the intervention group had a

significant increase in their HSRT scores when compared to students in the control group. However, the difference in scores may have been due to the researcher being a more skilled facilitator than the other faculty conducting the control group debriefings. Interestingly, when students had higher scores on the HSRT, they highly rated the debriefing on the Debriefing Assessment for Simulation in Healthcare-Student Version. Overall, students gave higher scores to debriefing elements associated with the Debriefing for Meaningful Learning method. While the HSRT has established reliability, the Debriefing Assessment for Simulation in Healthcare-Student Version does not. Conversely, the Debriefing Assessment for Simulation in Healthcare-Student Version has content and criterion validity, but the HSRT has no criterion validity and is not specific to nursing. This study emphasizes that a trained facilitator was able to assist students in achieving greater reasoning abilities by using a method of debriefing that focuses on the students' affective and cognitive needs (Dreifuerst, 2012).

The Debriefing for Meaningful Learning method was also used by Mariani, Cantrell, Meakim, Prieto, and Dreifuerst (2013) as the intervention in their mixed method quasi-experimental study. A convenience sample of 86 out of 90 students enrolled in a medical surgical nursing course were randomly assigned to clinical groups that were used for the control and intervention groups. A power analysis was calculated and a moderate effect size would be detectable with 54 students at the $p < .05$ level and a power of .80. Students participated in the same simulation followed by either standard debriefing or a debriefing using the Debriefing for Meaningful Learning method (Mariani et al., 2013). All students were evaluated by the clinical faculty using the Lasater Clinical Judgment Rubric based on their simulation performance prior to the debriefing. The students then

completed a second HFPS and a Lasater Clinical Judgment Rubric was completed on them by the research team. After the second HFPS, all students then participated in the intervention method, Debriefing for Meaningful Learning. There was no significant difference in rubric scores between the intervention and control groups. The Lasater Clinical Judgment Rubric score was determined by the clinical faculty for the first scenario and the researchers for the second scenario. The researchers' scores were used for the second scenario to blind the researchers to whether the students were in the control or intervention group. The researchers also scored the first scenario to determine an inter-rater reliability for the study. The Lasater Clinical Judgment Rubric is a valid and reliable instrument and the research team had an inter-rater reliability of 0.92 with the clinical faculty on the ratings for the first simulation. This study was limited by possible history and maturation effects since students were in their clinical groups for either four or five weeks between simulations (Mariani et al., 2013). Additionally, the Lasater Clinical Judgment Rubric does not measure changes in reflective thinking and may not be a good proxy measure for reflective thinking.

In addition to the quantitative analysis, seven volunteers representing both the control and intervention groups participated in either a focus group or an individual interview (Mariani et al., 2013). Student believed that Debriefing for Meaningful Learning was a more learner focused holistic approach, that promoted figuring problems out, assisted students in making connections, and improved student learning (Mariani et al., 2013). Students thought that the standard debriefing method was a more instructor focused method that did not look at the whole picture, concentrated on what was right versus wrong, and was not as helpful to learning as the Debriefing for Meaningful

Learning method (Mariani et al., 2013). Although this study did not find a relationship between the Debriefing for Meaningful Learning method and increased scores on the LCJR; it gathered more evidence on the aspects of debriefing that students valued. Several other researchers supported the assertion that the debriefing needs to be focused on the students' affective and learning needs, and experiences (Boyd, 2002; Chou et al., 2011; Delany & Watkin, 2009; Dreifuerst, 2012; Dye, 2005; Ekebergh, 2011; Honey, Waterworth, Baker, & Lenzie-Smith, 2006; Lasater, 2007b; McMahon et al., 2005).

A convenience sample of 19 speech language pathology students were randomly assigned to clinical groups and used to test two different ways of receiving feedback on clinical skills and motivation (Ho & Whitehill, 2009). T-tests reveal no significant differences between the two groups in terms of: age, and sophomore or junior year GPA. However, three of the four male students were assigned to same group. One group gave a verbal self-evaluation and received immediate verbal group feedback as has been the standard in scenario debriefing. The other group submitted a written self-evaluation and received delayed written individual feedback. The immediate verbal feedback group received significantly higher ratings on their clinical skills, although the median score was 3 for both groups. Sixteen of the nineteen students received a score of 3 and the remaining scores were 2, 3.5, and 4. Overall, students received higher scores at the end of the course that during the mid-term evaluation. The immediate feedback students rated themselves higher than the delayed feedback students on the Motivated Strategies for Learning Questionnaire. The Motivated Strategies for Learning Questionnaire is a reliable survey tool but it was modified for this study so that it could be used by students

to evaluate clinical learning. The students tended to score themselves very similarly with the median being 5 for both groups (Ho & Whitehill, 2009).

The study was limited by the small sample, and homogenous scoring by the faculty and ratings by the students (Ho & Whitehill, 2009). Possible confounding was introduced by examining three sets of variables at once: immediate versus delayed, verbal versus written, and group versus individual feedback (Ho & Whitehill, 2009). Students in the immediate verbal feedback group felt that they learned from participating with other students but that the debriefing process was time consuming (Ho & Whitehill, 2009). The group of control students, who received individual delayed feedback, felt that they were better able to reflect and that writing and receiving written evaluations was more time efficient (Ho & Whitehill, 2009). However, only three of the ten students who received delayed group feedback and none of the students who received immediate individual feedback preferred the written feedback method (Ho & Whitehill, 2009). The reason for preferring a verbal exchange may have been, as two students in the written feedback group commented, that it was more difficult to write a self-evaluation (Ho & Whitehill, 2009). Several other studies highlighted the importance that students placed on receiving immediate feedback (Corrigan & Hardham, 2011; Dye, 2005; Flanagan et al., 2004). A follow up study testing each set of variables separately is needed to figure out what is the optimal way to receive reflections from students and give feedback to students. The evidence from Ho and Whitehill (2009) suggested that students perceived the benefits of both verbal and written reflective exercises. Therefore, it may be that in order to gain the most from a simulation, students should participate in both an immediate group verbal debriefing and a delayed individual reflection assignment that

receives written comments from the facilitator. Delayed individual feedback and many other forms of extended reflection assignments are discussed in the next section.

Exams were used by Tofil et al. (2010) to measure changes in pharmacy students' knowledge and application skills after a case study and two HFPS with reflective debriefings. The study was a pre-test post-test no control group design. Although 42 out of 45 of the eligible students participated, the sample suffered from self-selection bias since the students were recruited from an elective course. Additionally, two samples were combined from students enrolled in the course over two years without any analysis of whether the two samples were congruent. There were significant increases in student exam scores from pre-test to post-test when analyzed using paired t-tests. Ninety-five percent of the students improved their scores when compared using a chi-square analysis. Since there was no control group, it is difficult to state whether the increase in the researcher designed exam was due to the intervention or maturation. The exam was a test of knowledge and application related to the content of the case study and simulations and had face validity. Questions on the exam that addressed the application of knowledge showed the greatest amount of improvement, which may support the assertion that the intervention influenced the increase in scores. Additionally, students reported that they liked reflecting and the instructors reported that they believed the students benefitted from reflecting (Tofil et al., 2010). Ultimately, this study needed a control group to prove that the reflective debriefings caused the significant rise in application ability.

The studies that examined the effect of reflective debriefings had significant flaws. Although Dreifuerst (2012) found a significant difference in the HSRT scores of students debriefed using the DML method; the results may have been due to her ability as

a facilitator. The strongest evidence for using a DML method of debriefing shows no difference in the LCJR (Mariani et al., 2013). However the LCJR may not have been an appropriate proxy for reflective thinking. Another study looked at the time of the debriefing and found that students performed better on clinical skills and motivation scoring, after an immediate verbal group reflective debriefing than when receiving delayed individual written feedback (Ho & Whitehill, 2009). However, the study grouped multiple variables together making any claim of significance to the timing of the feedback suspect. Although the strongest studies purport the importance of reflection and reflective thinking to debriefing and ultimately to practice; all of these studies used proxy measure for a change in reflective thinking (Ho & Whitehill, 2009; Mariani et al., 2013; Tofil et al., 2010). What was ultimately gained from these mixed method studies comes from the non-analytical portion: the recognition on the part of students and faculty of the value of student reflection (Ho & Whitehill, 2009; Mariani et al., 2013; Tofil et al., 2010). The body of descriptive evidence supports Mariani et al.'s (2013) assertion that debriefing methods need to focus on the needs of the students (Boyd, 2002; Chou et al., 2011; Delany & Watkin, 2009; Dreifuerst, 2012; Dye, 2005; Ekebergh, 2011; Honey et al., 2006; Lasater, 2007b; McMahon et al., 2005). Additional analytical studies are needed to examine the difference in reflective writing, speech, and subsequent student behaviors after exposure to differing methods of reflective debriefing. Measuring related concepts such as changes in knowledge, clinical skills, or clinical judgment without also measuring changes in reflection does not help to tease out the relationship between the concepts. Control groups are needed to detect changes that might be due to history or maturation and are especially important since multiple simulations with reflective

debriefings might be needed before measurable changes in students' reflective thinking may develop.

Extended Reflection

Faculty should give students both guidelines and allotted time to undertake one or more extended reflection activities: essay, journal writing, taped log, care planning or mapping, related case studies, transcribed scenario, online or face-to-face discussions, or group Wiki. Students used the Outcome, Present state, Test Model (OPT) to frame patients encountered in clinical and reflected on the process in a log (Kautz, Kuiper, Pesut, Knight-Brown, & Daneker, 2005). There was a two week period of class during which students were trained how to use the self-regulation prompts and the OPT model. A purposive sample of 23 junior nursing students and their clinical faculty were used to implement the OPT model after each clinical. In the reflective logs describing the use of the OPT model, students addressed their behaviors, metacognition, and worked through problems. Students were compared to a previous student sample and: showed greater self-observation, self-judgment, knowledge work, and use of personal resources but were significantly less self-efficacious and used fewer environmental structuring strategies. Over the ten weeks, the student logs showed progression in framing of patient situations and choice of interventions (Kautz et al., 2005).

The Learning from your Experience as a Professional (LEaP) critical reflection guidelines designed by Aronson, Kruidering, Niehaus, and O'Sullivan (2012) were used along with different forms of feedback to examine their effect on the reflection level of students' writing (Aronson, Niehaus, Hill-Sakurai, Lai, & O'Sullivan, 2012). A quasi-experimental pre-test post-test design was used with a cohort of 167 third year medical

students (Aronson, Kruidering, et al., 2012). Students were randomly assigned with one group receiving the definition of reflection and the other receiving both the definition and the LEaP guidelines (Aronson, Niehaus, et al., 2012). The students were then randomly assigned to either receive feedback on the content of their reflections or to receive feedback on both the content and their ability to reflect (Aronson, Niehaus, et al., 2012). Unfortunately, the study's four arms were uneven due to 18 students that were excluded since they only participated in part of the course and did not complete both assignments (Aronson, Niehaus, et al., 2012).

A research assistant de-identified each reflection so that the raters would be blinded as to the identity of the students (Aronson, Niehaus, et al., 2012). Previously, the raters had been trained in the use of the Reflective Ability Scoring Rubric devised by O'Sullivan, Aronson, Chittenden, Niehaus, and Learman (2010) and had obtained an inter-rater reliability of 0.91. The Reflective Ability Scoring Rubric is a valid and reliable instrument (Aronson, Niehaus, et al., 2012). Four researchers gave student feedback according to a protocol and during training and practiced giving feedback until the feedback was similar (Aronson, Niehaus, et al., 2012). During the course of the study, the researchers gave and compared feedback on the same reflective piece to check for sustained continuity (Aronson, Niehaus, et al., 2012). Students that used the guidelines performed significantly better than students who received only the definitions (Aronson, Niehaus, 2012). Additionally, students that received feedback on both content and their reflective ability scored higher than students who received feedback only on the content of their reflective writing (Aronson, Niehaus, et al., 2012). However, there was no interaction between having the guidelines and being given additional feedback

(Aronson, Niehaus, et al., 2012). This study reveals the importance of both a guideline and feedback on reflective ability to the success of students in reflective writing assignments,

In an earlier study, Aronson, Niehaus, Lindow, Robertson, and O'Sullivan (2011) used a cohort of third year medical students to test the LeaP reflective learning guide. The guide was given to the intervention group before reflection and used by faculty to provide feedback (Aronson et al., 2011). The control group received a short prompt to guide their reflective writing (Aronson et al., 2011). Both raters had been previously trained and obtained a 0.89 for inter-rater reliability on the scoring rubric (Aronson et al., 2011). Five essays were unable to be fully analyzed and two students did not complete the course, resulting in a sample of 115 out of the cohort of 122 (Aronson et al., 2011). The essays were an ungraded assignment which may have led to having five essays that were not able to be scored (Aronson et al., 2011). All third year students rotated through the course and were assigned to either the control or intervention groups based on the timing of their rotation (Aronson et al., 2011). The first two rotations were controls and the last three were intervention groups (Aronson et al., 2011). The researchers believed that there was not a maturation affect since the third and fourth rotation scores did not significantly differ from the fifth and sixth rotation scores (Aronson et al., 2011). However, the study would have been more rigorous if the intervention and control groups had alternated rotations. Since the scores from the third, fourth, fifth, and sixth rotations were homogeneous, the scores were combined into one group (Aronson et al., 2011). The 78 students in the combined intervention group scored significant higher on their post-test writing than the control group (Aronson et al., 2011). Neither gender nor

learner satisfaction was correlated with a higher reflection score (Aronson et al., 2011). The researchers believed that reflective ability would improve with practice (Aronson et al., 2011). In summary, the quantitative part of the study found that the use of LEaP guidelines assisted students in writing higher level reflective pieces as did feedback that included comments on the students' reflective writing ability.

Fakude and Bruce's (2003) quasi-experimental study did not find a significant difference in reflective writing scores between students who had practiced reflective journaling and students who had not. Forty-three first year nursing student volunteers, out of a cohort of fifty-three, participated in the study. The students were assigned to groups based on which campus they attended. Although not random, this method reduced the possibility of contamination between groups. However, there was no comparison of demographic variables between the groups or pre-testing scores to ensure the groups were comparable. The 20 students in the intervention group had faculty support and used guidelines to write between one and four reflective entries over eight weeks. The voluntary ungraded reflective entries were combined and scored as one piece. At the end of the eight weeks, all students were required to write a reflective paper. Both the journals and the paper were evaluated by two researchers using a tool that had content validity. The reflective paper scored showed an improvement over the intervention groups' journal entries but the improvement was not significant. The non-significance may have been due to the combining all of the student's journal entries written over eight weeks into one writing sample. The 5%-20% difference in the experimental groups journal and paper scores was possibly due to experience, maturation, history, or the effort put into a graded assignment versus an ungraded one. Additionally,

the researchers felt that lack of discussion may also have contributed to the lack of a significant rise in the intervention students' scores (Fakude & Bruce, 2003).

Overall, there was no difference between the intervention and the control group reflective writing scores on the paper (Fakude & Bruce, 2003). However, both the intervention and the control group scored 100% in three areas of reflection: description, affective, and evaluation, (Fakude & Bruce, 2003). This may have resulted in a ceiling effect. The overall reflective writing scores combined the scores in all six areas: description, affective, evaluation, analysis, alternatives, and reflection before action (Fakude & Bruce, 2003). Reflection before action was considered the highest level of reflection (Fakude & Bruce, 2003). The reflection before action scores were significantly higher in the intervention group (Fakude & Bruce, 2003). Although Fakude and Bruce's (2003) study did not find significance; the evidence supporting the use of reflective writing guidelines was reported by later more rigorous studies (Aronson, Niehaus, et al., 2012; Aronson et al., 2011). One reason that Fakude and Bruce may not have found a significant difference in the overall scores might have been that a different method of scoring was used than in the other studies (Aronson, Niehaus, et al., 2012; Aronson et al., 2011). An evaluation rubric was used in the studies with significant findings (Aronson, Niehaus, et al., 2012; Aronson et al., 2011). Additionally, the LEaP guidelines and study methodology were fine-tuned by Aronson et al.'s (2012) study from Aronson et al. (2011). The problem Fakude and Bruce's small sample size was also overcome, when later studies used cohorts of third year medical students (Aronson, Niehaus, et al., 2012; Aronson et al., 2011). The most recent studies found a significant positive effect when students were provided with detailed guidelines (Aronson, Niehaus, et al., 2012; Aronson

et al., 2011). This may have been due to the sample size, specific interventions, experimental design, or rubric (Aronson, Niehaus, et al., 2012; Aronson et al., 2011). The most rigorous of the four studies, Aronson et al.'s 2012 quasi-experimental cohort study, also found that having faculty provide feedback on the student's reflective ability assisted the student in improving their reflective writing. Although, the intervention group was given instruction, assistance, and a guide in reflective journaling, Padden's quasi-experimental pre-test, post-test design with control group did not find a significant rise in reflective ability, insight, or perceived clinical decision making. However, only 33 out of 60 (55%) of the intervention group completed, while 79 out of 93 (85%) of the controls completed the study (Padden, 2011). Additionally, the number of students need for the power analysis was not reached (Padden, 2011). Twenty-two of the thirty-three intervention students submitted only two journals over the 14 weeks of the study, the minimum number needed to be considered to have completed the study (Padden, 2011).

Perera et al.'s (2010) quasi-experimental study with control found significant differences in the OSCE scores in their sample of 202 first year medical students. The intervention group students were trained on how to give feedback to peers and evaluate performance with a standardized patient. Students used a self-assessment tool to guide reflection and identify performance gaps of their simulated patient encounter. After review the reflections, peers and then faculty gave written feedback on any additional uncovered gaps in performance. Both the intervention and control groups received immediate feedback from the standardized patient and the facilitator. Intervention group students also improved their interview style, listening and building rapport skills. However, there was no difference between the groups in use of language or interview

structure. Ninety-percent of the intervention group students used self and peer reflective evaluation during their own spontaneous practice sessions. These students shared their new skills with some control group members and confidentiality may not have been maintained about the intervention and diluted the results. Students were assigned to groups based on pre-admission scores and there was no significant difference between the groups in gender or age distribution. Assessors were blindly assigned both intervention and control group students. Overall, the intervention was successful in assisting students in improving their interview skills and 86.4% of the students believed it was a positive process that developed team skills (Perera et al., 2010).

Jarris et al.'s (2012) study has been discussed previously in this chapter. The convenience sample of 190 first year medical students was divided into unequal groups, with 47 students comprising the intervention group (Jarris et al., 2012). There was no randomization and no demographic description of the sample (Jarris et al., 2012). Students in the intervention group viewed recordings of their first clinical skill assessment, completed a self-assessment, received immediate feedback from the standardized patient and faculty and delayed online feedback from faculty, and then wrote a reflective entry (Jarris et al., 2012). The study found no difference between the intervention or control group in pre or post-test clinical skill assessments (Jarris et al., 2012). One reason for the lack of significant findings may have been history or maturation since there was three months between pre and post testing (Jarris et al., 2012). The researchers felt that the lack of significant post-intervention differences between the groups may have been due to the students having not received any guidelines or instruction on how to critically reflect (Jarris et al., 2012). The researchers assumed that

the intervention students would complete all steps of the reflective process (Jarris et al., 2012). The steps of the reflective process used were defined by Sargeant, Mann, van der Vleuten, and Metsemakers (2009) as beginning with assessing performance and providing feedback. However, the last two steps, defining and putting into place an improvement plan based on all the feedback received, were not addressed by the students, perhaps due to a lack of faculty instruction or guidance (Jarris et al., 2012).

Ip et al. (2012) found that when given instruction and faculty support, students were able to improve their level of reflective writing. As discussed earlier, students kept voluntary reflective journals in Ip et al.'s cohort study. Most students were able to significantly improve their level of reflective writing after just two weeks of faculty intervention (Ip et al., 2012). Most students progressed from non-reflective to reflective, with 92.1% of the sample rated as non-reflective in the pre-test and 23.7% at two weeks, and 13.2% at four weeks (Ip et al., 2012). A small percentage (13.2%) of students attainted a critical reflector rating at weeks two and four (Ip et al., 2012). There was high inter-rater reliability on the rating of the students' writing samples (Ip et al., 2012). Limitations of this study are that 76.3% of students who completed the study requirements were regular writers in diaries, and that completers were not compared to non-completers (Ip et al., 2012). In the qualitative portion of Ip et al.'s study, students revealed that they thought the two biggest barriers to improvement in reflective ability were lack of time and the unavailability of the faculty. Other descriptive evidence reported that having time was a critical factor in students being able to successfully complete a reflective writing assignment (Croke, 2004; Donovan, 2007; Dye, 2005; Grant, Kinnersley, Metcalf, Pill, & Houston, 2006; Gwozdek, Klausner, & Kerschbaum,

2009; Harrison & Fopma-Loy, 2010; Hill, Davidson, & Theodoros, 2012; Kok & Chabeli, 2002; Lähteenmäaki, 2005; O'Donovan, 2006; Skovsgaard, 2004). Many descriptive studies also stressed the importance of having had the faculty provide guidelines, instruction, or assistance in critical reflection (Barry, 2008; Beyer, 2012; Boyd, 2002; Croke, 2004; Dye, 2005; Gwozdek et al., 2009; Kautz et al., 2005; Kelly, 2012; Kok & Chabeli, 2002; Ladyshewsky & Gardner, 2008; Padden, 2011; Pee, Woodman, Fry, & Davenport, 2002; Williams, Wessel, Gemus, & Foster-Seargeant, 2002). The body of evidence describes many different methods to collect students' reflective thinking: reflective papers, logs, journals, tape recording, OPT model completion, case studies, transcription of videotape, discussion boards or groups, or Wiki (Barry & O'Callaghan, 2008; Beyer, 2012; Chou et al., 2011; Croke, 2004; Daly, 2010; Dunfee et al., 2008; Durso, 2006; Dye, 2005; Grant et al., 2006; Gwozdek et al., 2009; Ho & Whitehill, 2009; Jarris et al., 2012; Kautz et al., 2005; Ker, 2003; Kuiper, 2005; Kuo, Turton, Cheng, & Lee, 2011; Ladyshewsky & Gardner, 2008; Lai & Hu, 2012; Lindgren & Athlin, 2010; Lutz et al., 2013; Makoul, Zick, Aakhus, Neely, & Roemer, 2010; Manning et al., 2009; Mamede et al., 2012; McMahon et al., 2005; O'Donovan, 2006; Plack, Driscoll, Blissett, McKenna, & Plack, 2005; Plack et al., 2007; Plack et al., 2008; Rowe, 2012; Tsang, 2012).

Assessment

Periodically, faculty should review student progress and assess long term outcomes from simulation activities including: themes of student learning, level and types of reflection, and proxy measures for higher level thinking skills. Epp's (2008) systematic review reported that undergraduate nursing students' reflective writing ability

develops over time, produces shifts in students' perspectives, and changes in practice. Although, undergraduate nursing students reflected primarily at lower levels, students were capable of reflecting at higher levels (Epp, 2008). Wald, Borkan, Taylor, Anthony, and Reis (2012) performed a systematic review of PubMed articles from 1995 to 2008 reviewing evidence of the best way to evaluate medical student reflective writing. A formative analytical rubric should have four steps: reading the narrative in its entirety, finding the criteria to support the analysis, deciding what level of reflection the writing represents, and listing the quotes that support the assessment (Wald et al., 2012). Several descriptive studies also stress the importance of a formative review of a student's reflective work by faculty (Bruce et al., 2001; Daly, 2010; Donovan, 2007; Duggan et al., 2009; Silvia, Valerio, & Lorenza, 2013). The descriptive evidence contains several different ways to evaluate reflection (Bae, 2012; Beyer, 2012; Boyd, 2008; Hulsman et al., 2009; Ip et al., 2012; Pee et al., 2002; Plack, et al., 2005; Plack et al., 2007; Silvia et al., 2013). Since it can be time consuming to gauge the level of reflection in written work, proxy measures have been used to monitor student progress (Dreifuerst, 2012; Lai & Hu, 2012; Mariani et al., 2013; Schwartz & Bohay, 2012).

Summary

The level of evidence concerning reflective thinking in HFPS is primarily at the descriptive level and extrapolated from other types of simulated patient experiences. Without higher levels of evidence focused on testing interventions mentioned in the descriptive literature, promoting reflective thinking in HFPS will be haphazard at best. If reflective practice is a goal of the nursing or simulation program, then reflective thinking must be required of the students, and reviewed and evaluated by faculty (Mann, Gordon,

& Macleod, 2009). Since promoting and assessing the reflective thinking of students is an arduous and expensive process, faculty need to apply the evidence already accrued from other healthcare educational programs.

CHAPTER 4

GUIDELINE

Introduction

In order to organize the guideline, recommendations were divided into eight sections. Evidence was first rated on a scale of 1++ to 4 using the SIGN (2011) that is located in Appendix C. Expert opinion was the lowest rated evidence and was not used. Recommendations were then graded according to the SIGN scale (2011) which is located in Appendix E. A √ which would have indicated an opinion of the author, but this level of evidence was not used. None of the recommendations had a very high level of evidence supported by a number of analytical studies. All recommendations received a grade of D which was based on a body of level 3 and 4 evidence or extrapolated 2++ level evidence. The SIGN scale (2011) continues upward to a grade of A, which was the best supported level of evidence.

Best Practice to Promote Higher Order Thinking Skills in HFPS

1) **Prepare students for simulations including an assessment of what the students already know** - Grade of Recommendation D. In order to make sure that students get the most from a HFPS, the faculty must be sure that the students have been properly prepared and have mastered the fundamental knowledge needed to be successful in a given scenario (Cahalin et al., 2011; Corrigan & Hardham, 2011; Delany & Watkin, 2009; Hatlevik, 2012; Lasater, 2007b; McMahon et al., 2005; Perera et al., 2010;

Thompson et al., 2010; Tofil et al., 2010). Additionally, students' reflective thinking ability will affect their ability to successfully complete scenarios (Decker, 2007).

2) **Carefully design all aspects of the simulation with the goal of maximizing the opportunity for student reflective thinking** - Grade of Recommendation D. HFPS must not only be linked to class material but progressively train students on harder scenarios containing similar concepts that may allow students to showcase their knowledge and abilities (Blatt et al., 2007; Lasater, 2007b). Students should not work in isolation; since other health professionals, friends, and family are all potential sources of assistance with a patient and having these roles in HFPS makes the experience move cognitively similar to real life (Bruce et al., 2001; Cahalin et al., 2011; Corrigan & Hardham, 2011; Daly, 2010; Ertmer et al., 2010; Lasater, 2007b; Lindgren & Athlin, 2010; Perera et al., 2010; Thompson et al., 2010).
3) **Videotape or otherwise record the simulation processes -** Grade of Recommendation D. Use of a videotaped orientation and briefing may make the experience more standardized, so that no points are forgotten and could save faculty time. Review of a videotaped scenario was helpful to students whether or not they were involved in the scenario (Corrigan & Hardham, 2011; Daly, 2010; Hulsman et al., 2009; Hussin, 2013; Kalish et al., 2011; Lasater, 2007b; Maloney et al., 2013; Thompson et al., 2010). Review of the debriefing recording can allow faculty to evaluate which facilitators and activities are most successful.
4) **Faculty need to conduct all simulation activities in a psychologically safe environment** - Grade of Recommendation D. Whether faculty are working in person

with students or asynchronously, students must believe that faculty are accepting and willing to help (Epp, 2008).

5) **Faculty need to provide education, training, and materials; and evaluate facilitators that conduct the scenario, debriefing, and extended reflection activities -** Grade of Recommendation D. Trained facilitators result in higher student satisfaction and self-confidence (Hallmark, 2010). Students' reflective writing level improved after interacting with trained facilitators (Ip et al., 2012). In order to see which methods are working, both the students' results and the facilitators' methods must be evaluated.

6) **Conduct an immediate debriefing, in a different area than the scenario, which should include: affective and cognitive content; simulation anomalies; a summary; and a focus on student learning, gaps in knowledge, learning process, and goals for future improvement -** Grade of Recommendation D. Simulated scenarios have a considerable impact on students' emotions and this emotional reaction must be dealt with before the cognitive aspects can be discussed (Dreifuerst, 2012; Ho & Whitehill, 2009). This emotional reaction is also the reason that it is better to immediately explore the emotional impact of the scenario and to move the debriefing from the bedside (Dreifuerst, 2012; Mariani et al., 2013). Debriefing methods that focus on the needs of the student will find greater acceptance than those based on a standard faculty derived protocol (Dreifuerst, 2012; Mariani et al., 2013).

7) **Give students time and guidelines to undertake one or more extended reflection activities: paper, journal writing, taped log, care mapping, or planning related case studies, transcribed scenario, online or face-to-face discussions, or group**

Wiki - Grade of Recommendation D. Without designated guidelines about what is expected of students, extended reflection activity results will not measure up to the faculty standards (Aronson, Niehaus, et al., 2012; Aronson et al., 2011; Fakude & Bruce, 2003; Ip et al., 2012; Jarris et al., 2012).

8) **Periodically review student progress and assess long term outcomes from simulation activities including: level and types of reflection, themes of student learning, and proxy measures for higher level thinking skills** - Grade of Recommendation D. Reflective abilities develop over years, so it is necessary to begin reflective thinking activities early in the nursing program and repeat frequently (Epp, 2008). To understand the impact of a sequence of reflective thinking activities, monitoring will be necessary (Epp, 2008). Since measuring reflective thinking abilities in writing, actions, and behaviors can be time consuming; proxy measure may be substituted for some assessments.

Summary

The body of evidence for encouraging reflective thinking during HFPS is insubstantial. Most studies are extrapolated from other simulated patient experiences. The body of work supporting this guideline is generally descriptive with a few higher level studies interspersed. The evidence for interventions is idiosyncratic and few studies built on the work of previous findings. Some studies looked for correlations with concepts not directly related to reflective thinking (Blatt et al., 2007; Cook, 2010). Several analytical studies used proxy measures of reflective thinking; some with positive findings and some with negative findings (Dreifuerst, 2012; Hallmark, 2010; Mariani et al., 2013; Tofil et al., 2010). However, correlations tended to be found when studies

directly measured reflective changes in students' writing, actions, and behaviors (Aronson, Niehaus, et al., 2012; Aronson et al., 2011). These recommendations serve to suggest avenues that may yield the best results and highlight methods that have been successful.

CHAPTER 5

CONCLUSION

Discussion of Recommendations

Though the body of evidence for reflective thinking best practices in HFPS is lacking in analytical studies, the descriptive evidence lays groundwork for future research. Educators can begin to apply the recommendations to their HFPS programs. This would include assessing the students' current level of reflective thinking and determining the best way to link reflective thinking practices in HFPS to other simulated patient experiences. Since King and Kitchener (1994) have shown that progress in reflective thinking can continue after graduation, employers of new graduate nurses must consider how they will promote reflective thinking habits of mind and reflective practice (Schon, 1983). Policy makers will have to decide the best ways to ensure that reflective thinking is both taught and reinforced to ensure that nurses are able to reflectively and critically think about their patients. Reflection has long been used by healthcare education programs including nursing. Although thoroughly described in the literature, not enough analytical studies have yet been published that would support highly graded recommendations and create a well-founded guideline for promoting reflection during HFPS. The current outcomes of the available research point the way for changes in the way student nurses are educated. Additional changes in the focus of continuing education for practicing nurses need to be considered. These changes will need to be

evaluated against current practice to see if there is an improvement in critical and reflective thinking as well as any other related outcomes.

Implication of Outcomes for Research

The first task will be to establish the relationship between critical and reflective student in nurses and practicing nurses. Standardized, objective tools that measure changes in the critical thinking of nurses may need to be developed. These tools need to be used at appropriate time intervals during which significant changes in higher order thinking skills develop. Additionally, the benefits of critical and reflective thinking in nurses needs to be tied to standardized objective measures of practice improvements. It is not known whether reflective thinking improves understanding, learning, self-assessment, clinical practice, or patient care (Mann et al., 2009). Also, possibility of harm to the student when forcing reflective thinking during simulated patient activities needs to be investigated. At least one study has reported increased stress as being among the possible negative effects of reflective thinking activities (Corrigan & Hardham, 2011).

The reflective thinking body of literature outside of health pre-professional programs needs to be analyzed to discover what is known about how to best promote and measure reflective thinking. One of the strongest pieces of evidence (Ip et al., 2012) found that students' reflective writing can be rapidly improved with facilitator intervention; suggesting that investigating students' reflective writing ability is a worthwhile research area. The Debriefing for Meaningful Learning method (Dreifuerst, 2012) needs to be compared to less well researched debriefing methods such as the Debriefing with Good Judgment (Rudolph, Simon, Dufresne, & Raemer, 2006; Rudolph,

Simon, Raemer, & Eppich, 2008; Rudolph, Simon, Rivard, Dufresne, & Raemer, 2007). Other questions remain to be answered by additional research. At least one study has reported increased stress among the possible negative effects of reflective thinking activities (Corrigan & Hardham, 2011). Additionally, the occurrence of reflection without learning and "recipe-following" should be investigated and methods found that can limit these outcomes (Mann et al., 2009). Teaching reflection is a nuanced facilitated activity that requires attention to individualized support of the learner. The best methods for reducing "answer grabbing" strategies of students and maximizing mastering of professional reflective ability need to be delineated so that they can be adopted and modified by teachers.

Implications of Outcomes for Education

Making reflective thinking a common thread within the nursing curriculum, beginning with reflections on students' previous experiences, may assist in developing reflective thinking. Mann et al.'s (2009) systematic review of reflective thinking in the health professions reported that improvement may be linked to professional development and other types of learning that take place over several months or years. All faculty, students, and preceptors will need to be trained in reflective thinking. Comprehensive guidelines will need to be developed for both creating and scoring reflective assignments. Summative feedback of the level of reflection will need to be provided to students, along with formative feedback whenever reflective assignments are given. Additional hardware expenses may be incurred by the video-recording and retention of HFPS activities: orientation, pre-briefing, scenario, debriefing, and extended reflection

exercises. The retention of these materials will allow for future analysis and improvement of the program.

Lab faculty will need time to be instructed in debriefing training, and designing and running a HFPS program that promotes reflective practice. Immersive experiences can be designed to allow for reflection before, during, and after action (Levett-Jones et al., 2011). The use of "time out" period during the running of a scenario may provide students with an opportunity to reflect during action (Hill et al., 2012). Specific feedback needs to be provided to students based on their unique experiences (Dreifuerst, 2012; Mariani et al., 2013).

Interdisciplinary education that reduces the silos in healthcare can be accommodated by HFPS. Building teams, improving communication, and understanding the roles of each of the healthcare specialties have been addressed in reflective activities after simulation (Chou et al., 2011; Perera et al., 2010; Smith & Cole, 2009). By inculcating students in the habits of reflective thinking during HFPS, faculty can establish reflective thinking habits of mind that may continue long after graduation.

Implications of Outcomes on Practice

The progression of nurses' ability to think about their patients changes dramatically in the first ten years of practice (Benner, 1982). The reflective thinking of both new graduate nurses and those that have reached competency should be reinforced. Employers of new graduate nurses should consider using a residency or internship to improve professional reflective thinking among their nursing staff. Mentors used by hospitals to train new graduate nurses should be well versed in how to encourage reflective thinking. Reflective thinking exercises that were previously used with

simulated patients could now be transferred to a reflective practice environment. One exercise could be to require reflective journals of new graduate nurses where perplexing cases could be re-examined and discussed with the mentor. These reflective thinking programs could be also be used with more experienced nurses to ensure that they are continuing to develop their reflective practice.

Implications of Outcomes for Policy

The CCNE and NLN need to consider whether reflective thinking should be requirement of a nursing educational program; since reflective thinking is an essential part of critical thinking. Due to the developmental nature of reflective thinking and the progression in the thinking of a practicing nurse (Benner, 1984), policy makers should consider making HFPS programs part of mandated continuing education requirements for all new graduate nurses. Medicine has long had a nearly universal, formal residency program required of all new physicians that has resulted in practice trained professionals who are allowed to develop their practice over additional years of progressive training. Additionally, nurses re-entering practice, and nurses changing their practice focus should train in a simulated environment that develops the habits of mind and reflective practice they will need in their chosen area. HFPS continuing education for nurses might be able obtained in larger facilities Simulation Centers, but might have to be contracted out to schools of nursing. HFPS can be used not only to monitor a nurses' practice but to provide instruction in reflective thinking and measure reflective thinking skills. Educational programs that are leveled to student nurses, new graduates, and nurses who are re-entering practice or changing practice areas can serve to assist nurses in

establishing habits of mind, developing professionally, and maintaining a reflective practice.

Summary

A great change is coming is post-secondary education, where the emphasis will not be upon the delivery of facts to the students, but upon the cultivation of higher level thinking skills. Nursing is at the forefront of this movement and can lead the way in increasing students' reflective and critical thinking abilities. The simulation laboratory is an excellent place to increase the discipline specific thinking skills required of nurses. The controlled environment allows for extensive planning and preparation that can dramatically enhance the experience of the students. The selection of one "perfect" case can take the place of many real patients (Dewey, 1933).

Furthermore, it is not cost effective to spend thousands for dollars on laboratory equipment without investing time and effort into the running of a HFPS program (Lapkin & Levett-Jones, 2011). Faculty members are needed to discover and put into place the latest findings in patient simulation. Without a concerted effort to maximize specific student outcomes, an opportunity to greatly enhance the student's experience will be lost. The best use of resources may be carefully coordinated HFPS programs that encourage students to reflect on their experiences and incorporate learning into practice.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

AACN. (2008). *The essentials of baccalaureate education for professional nursing practice.* Retrieved from American Association of Colleges of Nursing website: www.aacn.nche.edu/education-resources/baccessentials08.pdf

Alinier, G. (2007). A typology of educationally focused medical simulation tools. *Medical Teacher, 29*(8), e243-250. doi: 10.1080/01421590701551185

Aronson L, Kruidering M, Niehaus B, O'Sullivan P. (2012). *UCSF LEaP (Learning from your Experiences as a Professional): Guidelines for Critical Reflection.* Available from: www.mededportal.org/publication/9073

Aronson, L., Niehaus, B., Hill-Sakurai, L., Lai, C., & O'Sullivan, P. S. (2012). A comparison of two methods of teaching reflective ability in Year 3 medical students. *Medical Education, 46*(8), 807-814. doi: 10.1111/j.1365-2923.2012.04299.x

Aronson, L., Niehaus, B., Lindow, J., Robertson, P., & O'Sullivan, P. (2011). Development and pilot testing of a reflective learning guide for medical education. *Medical Teacher, 33*(10), e515-521.

Bae, M.-J. (2012). Student music therapists' differences in their clinical reflections across practicum levels. *Music Therapy Perspectives, 30*(1), 89-93.

Barry, P., & O'Callaghan, C. (2008). Reflexive journal writing: A tool for music therapy student clinical practice development. *Nordic Journal of Music Therapy, 17*(1), 55-66.

Becherer, V. H. (2011). *The impact of a reflective thinking intervention on nursing students in a child and family nursing course.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3473942)

Benner, P. (1982). From novice to expert... the Dreyfus model of skill acquisition. *American Journal of Nursing, 82*, 402-407.

Benner, P. (1984). *From novice to expert: Excellence and power in clinical nursing practice.* Menlo Park, CA: Addison-Wesley Publishing.

Beyer, D. A. (2012). Enhancing critical reflection on simulation through Wikis. *Clinical Simulation in Nursing, 8*(2), e67-70. doi: 10.1016/j.ecns.2010.12.003

Birch, L., Jones, N., Doyle, P. M., Green, P., McLaughlin, A., Champney, C., . . . Taylor, K. (2007). Obstetric skills drills: Evaluation of teaching methods. *Nurse Education Today, 27*(8), 915-922.

Blatt, B., Plack, M., Maring, J., Mintz, M., & Simmens, S. J. (2007). Acting on reflection: The effect of reflection on students' clinical performance on a standardized patient examination. *Journal of General Internal Medicine, 22*(1), 49-54. doi: 10.1007/s11606-007-0110-y

Boud, D. (2001). Using journal writing to enhance reflective practice. *New Directions for Adult and Continuing Education*(90), 9-17.

Boyadjian-Samawi, Z. (2006). *The effect of concept mapping on critical thinking skills and dispositions of junior and senior baccalaureate nursing students.* (Doctoral

Dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. AAI3207567)

Boyd, E. M., & Fales, A. W. (1983). Reflective learning: Key to learning from experience. *Journal of Humanistic Psychology, 23*(2), 99-117. doi: 10.1177/0022167883232011

Boyd, L. D. (2002). Reflections on clinical practice by first-year dental students: A qualitative study. *J Dent Educ, 66*(6), 710-720.

Boyd, L. D. (2008). Development of reflective judgement in the pre-doctoral dental clinical curriculum. *European Journal of Dental Education, 12*(3), 149-158. doi: http://dx.doi.org/10.1111/j.1600-0579.2008.00511.x

Brannan, J. D., White, A., & Bezanson, J. L. (2008). Simulator effects on cognitive skills and confidence levels. *Journal of Nursing Education, 47*(11), 495-500. doi: 10.3928/01484834-20081101-01

Brookfield, S. D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting.* San Francisco: Jossey-Bass.

Brooks, K. L., & Shepherd, J. M. (1990). The relationship between clinical decision-making skills in nursing and general critical thinking abilities of senior nursing students in four types of nursing programs. *Journal of Nursing Education, 29*(9), 391-399.

Brown, D., & Chronister, C. (2009). The effect of simulation learning on critical thinking and self-confidence when incorporated into an electrocardiogram nursing course. *Clinical Simulation in Nursing, 5*(1), e45-52.

Brown, F. S. (2011). *Nursing faculty beliefs and practices regarding debriefing human patient simulation experiences.* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3474352)

Bruce, C., Parker, A., & Herbert, R. (2001). The development of a self-directed and peer-based clinical training programme. *International Journal of Language & Communication Disorders, 36*, 401-405.

Cahalin, L. P., Markowski, A., Hickey, M., & Hayward, L. (2011). A cardiopulmonary instructor's perspective on a standardized patient experience: Implications for cardiopulmonary physical therapy education. *Cardiopulmonary Physical Therapy Journal, 22*(3), 21-30.

Cant, R. P., & Cooper, S. J. (2010). Simulation-based learning in nurse education: Systematic review. *J Adv Nurs, 66*(1), 3-15. doi: 10.1111/j.1365-2648.2009.05240.x

Carr, S., & Carmody, D. (2006). Experiential learning in women's health: Medical student reflections. *Medical Education, 40*(8), 768-774. doi: http://dx.doi.org/10.1111/j.1365-2929.2006.02536.x

CASP. (2010). *Cohort study checklist.* Retrieved from Critical Appraisal Skills Programme International website: http://www.caspinternational.org/index.asp?o=1061

CASP. (2013a). *Qualitative research checklist.* Retrieved from Critical Appraisal Skills Programme International website: http://www.casp-uk.net/wp-content/uploads/2011/11/CASP-Qualitative-Research-Checklist-31.05.13.pdf

CASP. (2013b). *Randomised controlled trials checklist.* Retrieved from Critical Appraisal Skills Programme International website: http://www.caspinternational.org/index.asp?o=1075

CASP. (2013c). *Systematic review checklist.* Retrieved from Critical Appraisal Skills Programme International website: http://www.casp-uk.net/wp-content/uploads/2011/11/CASP-Systematic-Review-Checklist-31.05.13.pdf

Chabeli, M. M. (2007). Facilitating critical thinking within the nursing process framework: A literature review. *Health SA Gesondheid, 12*(4), 69-89.

Chau, J. P., Chang, A. M., Lee, I. F., Ip, W. Y., & Wootton, Y. (2001). Effects of using videotaped vignettes on enhancing students' critical thinking ability in a baccalaureate nursing programme. *J Adv Nurs, 36*(1), 112-119.

Chou, C. L., Johnston, C. B., Singh, B., Garber, J. D., Kaplan, E., Lee, K., & Teherani, A. (2011). A "safe space" for learning and reflection: One school's design for continuity with a peer group across clinical clerkships. *Acad Med, 86*(12), 1560-1565. doi: 10.1097/ACM.0b013e31823595fd

Colucciello, M. L. (1997). Critical Thinking Skills and Dispositions of Baccalaureate Nursing Students--A Conceptual Model for Evaluation. *Journal of Professional Nursing, 13*(4), 236-245.

Cook, J. L. (2010). *Can student reflection predict academic success and clinical performance in a physical therapist education program?* (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3398450)

Corrigan, R., & Hardham, G. (2011). Use of technology to enhance student self evaluation and the value of feedback on teaching. *International Journal of Therapy & Rehabilitation, 18*(10), 579-590.

Croke, E. (2004). The use of structured reflective journal questions to promote fundamental development of clinical decision-making abilities of the first-semester nursing student. *Contemp Nurse, 17*(1-2), 125-136.

Daly, G. (2010). Formative assessment as a clinical supervision tool. *Perspectives on Administration & Supervision, 20*(3), 113-116.

Decker, S. I. L. (2007). *Simulation as an educational strategy in the development of critical and reflective thinking: a qualitative exploration.* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3271413)

Delany, C., & Watkin, D. (2009). A study of critical reflection in health professional education: "Learning where others are coming from". *Advances in Health Sciences Education, 14*(3), 411-429.

Dewey, J. (1933). *How we think.* Boston: D. C. Heath and Company.

Donovan, M. O. (2007). Implementing reflection: Insights from pre-registration mental health students. *Nurse Education Today, 27*(6), 610-616.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health, 52*(6), 377-384. doi: 10.1136/jech.52.6.377

Dreifuerst, K. T. (2009). The essentials of debriefing in simulation learning: A concept analysis. *Nursing Education Perspectives, 30*(2), 109-114.

Dreifuerst, K. T. (2012). Using Debriefing for Meaningful Learning to foster development of clinical reasoning in simulation. *J Nurs Educ, 51*(6), 326-333. doi: 10.3928/01484834-20120409-02 [doi]

Duggan, A., Bradshaw, Y. S., Carroll, S. E., Rattigan, S. H., & Altman, W. (2009). What can I learn from this interaction? A qualitative analysis of medical student self-reflection and learning in a standardized patient exercise about disability. *J Health Commun, 14*(8), 797-811. doi: 10.1080/10810730903295526

Dunfee, H., Rindflesch, A., Driscoll, M., Hollman, J., & Plack, M. M. (2008). Assessing reflection and higher-order thinking in the clinical setting using electronic discussion threads. *Journal of Physical Therapy Education, 22*(2), 60-67.

Durso, S. C. (2006). Reflection in action: Learner-centered teaching. *Johns Hopkins Advanced Studies in Medicine, 6*(5), 243.

Dye, D. (2005). Enhancing critical reflection of students during a clinical internship using the self-S.O.A.P. note. *Internet Journal of Allied Health Sciences & Practice, 3*(4), 10p.

Ekebergh, M. (2007). Lifeworld-based reflection and learning: A contribution to the reflective practice in nursing and nursing education. *Reflective Practice, 8*(3), 331-343. doi: 10.1080/14623940701424835

Ekebergh, M. (2011). A learning model for nursing students during clinical studies. *Nurse Education in Practice, 11*(6), 384-389. doi: 10.1016/j.nepr.2011.03.018

Epp, S. (2008). The value of reflective journaling in undergraduate nursing education: a literature review. *International Journal of Nursing Studies, 45*(9), 1379-1388.

Ertmer, P. A., Strobel, J., Cheng, X., Chen, X., Kim, H., Olesova, L., . . . Tomory, A. (2010). Expressions of critical thinking in role-playing simulations: Comparisons across roles. *Journal of Computing in Higher Education, 22*(2), 73-94.

Facione, N. C., & Facione, P. A. (1994). *The "California Critical Thinking Skills Test" and the National League for Nursing accreditation requirement in critical thinking.* Millbrae, CA: California Academic Press. Retrieved from ERIC database. (ED380509)

Facione, N. C., Facione, P. A., & Sanchez, C. A. (1994). Critical thinking disposition as a measure of competent clinical judgment: the development of the California Critical Thinking Disposition Inventory. *Journal of Nursing Education, 33*(8), 345-350.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.* Newark, DE: American Philosophical Association. Retrieved from ERIC database. (ED380509)

Fakude, L. P., & Bruce, J. C. (2003). Journaling: A quasi-experimental study of student nurses' reflective learning ability. *Curationis, 26*(2), 49-55.

Flanagan, B., Nestel, D., & Joseph, M. (2004). Making patient safety the focus: Crisis resource management in the undergraduate curriculum. *Med Educ, 38*(1), 56-66. doi: 1701 [pii]

Fonteyn, M. E., & Cahill, M. (1998). The use of clinical logs to improve nursing students' metacognition: A pilot study. *Journal of Advanced Nursing, 28*(1), 149-154. doi: 10.1046/j.1365-2648.1998.00777.x

Goettl, B. P., Ashworth, A. R. S., & Chaiken, S. R. (2007). Advanced distributed learning for team training in command and control applications. In S. M. Fiore & E. Salas (Eds.), *Toward a science of distributed learning* (pp. 93-117). Washington, DC: American Psychological Association.

Gordon, J. M. (2000). Congruency in defining critical thinking by nurse educators and non-nurse scholars. *Journal of Nursing Education, 39*(8), 340-351.

Grant, A., Kinnersley, P., Metcalf, E., Pill, R., & Houston, H. (2006). Students' views of reflective learning techniques: An efficacy study at a UK medical school. *Medical Education, 40*(4), 379-388.

Green, S. B., & Salkind, N. J. (2008). *K related-samples tests: The Friedman and the Cochran test. Using the SPSS for Windows and Macintosh: Analyzing and understanding data* (6th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.

Greenwood, J. (1993). Reflective practice: A critique of the work of Argyris and Schon. *Journal of Advanced Nursing, 18*(8), 1183-1187. doi: 10.1046/j.1365-2648.1993.18081183.x

Griggs, R. R. (2003). *The effects of the use of a human patient simulator on the acquisition of nursing knowledge in undergraduate nursing students at a university in Illinois.* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3100754)

Gross, Y. T., Takazawa, E. S., & Rose, C. L. (1987). Critical thinking and nursing education. *Journal of Nursing Education, 26*(8), 317-323.

Grunwald, T., & Corsbie-Massay, C. (2006). Guidelines for cognitively efficient multimedia learning tools: Educational strategies, cognitive load, and interface

design. *Academic Medicine: Journal of the Association of American Medical Colleges, 81*(3), 213-223.

Gwozdek, A. E., Klausner, C. P., & Kerschbaum, W. E. (2009). Online directed journaling in dental hygiene clinical education. *J Dent Hyg, 83*(1), 12-17.

Hallmark, E. F. (2010). *The effect of faculty debriefing training on the achievement and satisfaction of pre-licensure nursing students participating in simulation.* (Doctoral Dissertation). Available from ProQuest Dissertation and Thesis database. (3431934)

Harrison, P. A., & Fopma-Loy, J. L. (2010). Reflective journal prompts: A vehicle for stimulating emotional competence in nursing. *Journal of Nursing Education, 49*(11), 644-652. doi: 10.3928/01484834-20100730-07

Hassan, K. E., & Madhum, G. (2007). Validating the Watson Glaser Critical Thinking Appraisal. *Higher Education: The International Journal of Higher Education and Educational Planning, 54*(3), 361-383.

Hatlevik, I. K. R. (2012). The theory-practice relationship: Reflective skills and theoretical knowledge as key factors in bridging the gap between theory and practice in initial nursing education. *Journal of Advanced Nursing, 68*(4), 868-877. doi: 10.1111/j.1365-2648.2011.05789.x

Henneman, E. A., & Cunningham, H. (2005). Using clinical simulation to teach patient safety in an acute/critical care nursing course. *Nurse Educator, 30*(4), 172-177.

Hill, A. E., Davidson, B. J., & Theodoros, D. G. (2012). Reflections on clinical learning in novice speech-language therapy students. *International Journal of Language &*

Communication Disorders, 47(4), 413-426. doi: 10.1111/j.1460-6984.2012.00154.x

Ho, D. W. L., & Whitehill, T. (2009). Clinical supervision of speech-language pathology students: Comparison of two models of feedback. *International Journal of Speech-Language Pathology, 11*(3), 244-255.

Hoffmann, R. L., O'Donnell, J. M., & Kim, Y. (2007). The effects of human patient simulators on basic knowledge in critical care nursing with undergraduate senior baccalaureate nursing students. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare, 2*(2), 110-114. doi: 10.1097/SIH.0b013e318033abb5

Honey, M., Waterworth, S., Baker, H., & Lenzie-Smith, K. (2006). Reflection in the disability education of undergraduate nurses: An effective learning tool? *Journal of Nursing Education, 45*(11), 449-453.

Howard, V. M. (2007). *A comparison of educational strategies for the acquisition of medical-surgical nursing knowledge and critical thinking skills: Human patient simulator vs. the interactive case study approach.* (Doctoral Dissertation). Available from Proquest Dissertations and Theses database. (UMI No. 3270096)

Hulsman, R. L., Harmsen, A. B., & Fabriek, M. (2009). Reflective teaching of medical communication skills with DiViDU: Assessing the level of student reflection on recorded consultations with simulated patients. *Patient Education and Counseling, 74*(2), 142-149. doi: 10.1016/j.pec.2008.10.009

Hussin, V. (2013). Student and teacher reflections on indirectness as a pragmatic feature of pharmacist-patient simulations. *English for Specific Purposes, 32*(2), 110-121.

Ip, W. Y., Lui, M. H., Chien, W. T., Lee, I. F., Lam, L. W., & Lee, D. T. (2012). Promoting self-reflection in clinical practice among Chinese nursing undergraduates in Hong Kong. *Contemporary Nurse: A Journal for the Australian Nursing Profession, 41*(2), 253-262.

Jarris, Y. S., Saunders, P., Gatti, M., & Weissinger, P. (2012). Critical reflection: Lessons learned from a communication skills assessment. *Medical Education, 46*(5). doi: 10.1111/j.1365-2923.2012.04237.x

Jeffries, P. R. (Ed.). (2007). *Simulation in nursing education: From conceptualization to evaluation.* New York, New York: National League for Nursing.

Jeffries, P. R., & Rizzolo, M. (2006). *Designing and implementing models for the innovative use of simulation to teach nursing care of ill adults and children: A national, multi-Site, multi-method study* [Summary Report]. Retrieved from National League for Nursing website: www.nln.org/researchgrants/LaerdalReport.pdf

Jones, I., & Alinier, G. (2009). Introduction of a new reflective framework to enhance students' simulation learning: A preliminary evaluation. *Blended Learning in Practice, 1*(1), 8-19.

Kalish, R., Dawiskiba, M., Sung, Y. C., & Blanco, M. (2011). Raising medical student awareness of compassionate care through reflection of annotated videotapes of clinical encounters. *Educ Health (Abingdon), 24*(3), 490. doi: 490 [pii]

Kataoka-Yahiro, M., & Saylor, C. (1994). A critical thinking model for nursing judgment. *Journal of Nursing Education, 33*(8), 351-356.

Kautz, D. D., Kuiper, R., Pesut, D. J., Knight-Brown, P., & Daneker, D. (2005). Promoting clinical reasoning in undergraduate nursing students: Application and evaluation of the Outcome Present State Test (OPT) model of clinical reasoning. *International Journal of Nursing Education Scholarship, 2*(1), 21p.

Kelly, P. J. (2012). Using popular movies and reflections to icrease the cultural sensitivity of students. *Journal of Physician Assistant Education, 23*(4), 39-42.

Ker, J. S. (2003). Developing professional clinical skills for practice - the results of a feasibility study using a reflective approach to intimate examination. *Med Educ, 37* Suppl 1, 34-41.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults.* San Francisco, CA: Jossey-Bass.

Kok, J., & Chabeli, M. M. (2002). Reflective journal writing: How it promotes reflective thinking in clinical nursing education: A students' perspective. *Curationis, 25*(3), 35-42.

Kolb, D. A. (1984). Learning and development in higher education. *Experiential Learning*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Kuiper, R. (2005). Self-regulated learning during a clinical preceptorship: The reflections of senior baccalaureate nursing students. *Nursing Education Perspectives, 26*(6), 351-356.

Kuiper, R. A., Heinrich, C., Matthias, A., Graham, M. J., & Bell-Kotwall, L. (2008). Debriefing with the OPT Model of Clinical Reasoning during high fidelity patient simulation. *International Journal of Nursing Education Scholarship, 5*(1), 1-13.

Kuiper, R. A., & Pesut, D. J. (2004). Promoting cognitive and metacognitive reflective reasoning skills in nursing practice: Self-regulated learning theory. *Journal of Advanced Nursing, 45*(4), 381-391. doi: 10.1046/j.1365-2648.2003.02921.x

Kuo, C. L., Turton, M., Cheng, S., & Lee, H. (2011). Using clinical caring journaling: Nursing student and instructor experiences. *Journal of Nursing Research, 19*(2), 141-149. doi: 10.1097/JNR.0b013e31821aa1a7

Ladyshewsky, R. K., & Gardner, P. (2008). Peer assisted learning and blogging: A strategy to promote reflective practice during clinical fieldwork. *Australasian Journal of Educational Technology, 24*(3), 241-257.

Lähteenmäaki, M. (2005). Reflectivity in supervised practice: Conventional and transformative approaches to physiotherapy. *Learning in Health & Social Care, 4*(1), 18-28.

Lai, C., & Hu, C. (2012). Supporting nursing students' critical thinking with a mobile web learning environment. *Nurse Educator, 37*(6), 235-236. doi: 10.1097/NNE.0b013e31826f27ee

Landeen, J., Byrne, C., & Brown, B. (1995). Exploring the lived experiences of psychiatric nursing students through self-reflective journals. *Journal of Advanced Nursing, 21*(5), 878-885. doi: 10.1046/j.1365-2648.1995.21050878.x

Lapkin, S., & Levett-Jones, T. (2011). A cost-utility analysis of medium vs. high-fidelity human patient simulation manikins in nursing education. *Journal of Clinical Nursing, 20*(23/24), 3543-3552. doi: 10.1111/j.1365-2702.2011.03843.x

Lasater, K. (2007a). Clinical judgment development: Using simulation to create an assessment rubric. *Journal of Nursing Education, 46*(11), 496-503.

Lasater, K. (2007b). High-fidelity simulation and the development of clinical judgment: Students' experiences. *Journal of Nursing Education, 46*(6), 269-276.

Lederman, L. C. (1992). Debriefing: Toward a Systematic Assessment of Theory and Practice. *Simulation & Gaming, 23*(2), 145-160.

Leigh, G., & Hurst, H. (2008). We have a high-fidelity simulator, now what? Making the most of simulators. *International Journal of Nursing Education Scholarship, 5*(1), Article 33.

Levett-Jones, T., Gersbach, J., Arthur, C., & Roche, J. (2011). Implementing a clinical competency assessment model that promotes critical reflection and ensures nursing graduates' readiness for professional practice. *Nurse Education in Practice, 11*(1), 64-69. doi: 10.1016/j.nepr.2010.07.004

Linden, L. L. (2008). *The effect of clinical simulation and traditional teaching versus traditional teaching alone on critical thinking of nursing students.* (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3310890)

Lindgren, B., & Athlin, E. (2010). Nurse lecturers' perceptions of what baccalaureate nursing students could gain from clinical group supervision. *Nurse Education Today, 30*(4), 360-364. doi: 10.1016/j.nedt.2009.09.008

Ling-Na, K., Bo, Q., Ying-qing, Z., Shao-yu, M., & Hui-Ming, G. (2014). The effectiveness of problem-based learning on development of nursing students' critical thinking: A systematic review and meta-analysis. *International Journal of Nursing Studies, 51*(3), 458-469. doi: 10.1016/j.ijnurstu.2013.06.009

Long, A. (2005). *Evaluative Tool for Mixed Method Studies.* Retrieved from http://usir.salford.ac.uk/13070/1/Evaluative_Tool_for_Mixed_Method_Studies.pdf

Lutz, G., Scheffer, C., Edelhaeuser, F., Tauschel, D., & Neumann, M. (2013). A reflective practice intervention for professional development, reduced stress and improved patient care-A qualitative developmental evaluation. *Patient Educ Couns.* doi: 10.1016/j.pec.2013.03.020

Magee, M. (2006). *State of the field review: Simulation in education.* Retrieved from Canadian Council on Learning website: http://www.ccl-cca.ca/pdfs/StateOfField/SFRSimulationinEducationJul06REV.pdf

Magnussen, L., Ishida, D., & Itano, J. (2000). The impact of the use of inquiry-based learning as a teaching methodology on the development of critical thinking. *Journal of Nursing Education, 39*(8), 360-364.

Makoul, G., Zick, A. B., Aakhus, M., Neely, K. J., & Roemer, P. E. (2010). Using an online forum to encourage reflection about difficult conversations in medicine. *Patient Educ Couns, 79*(1), 83-86. doi: 10.1016/j.pec.2009.07.027

Maloney, S., Storr, M., Morgan, P., & Ilic, D. (2013). The effect of student self-video of performance on clinical skill competency: A randomised controlled trial. *Advances in Health Sciences Education, 18*(1), 81-89.

Mamede, S., van Gog, T., Moura, A. S., de Faria, R. M. D., Peixoto, J. M., Rikers, R. M. J. P., & Schmidt, H. G. (2012). Reflection as a strategy to foster medical students' acquisition of diagnostic competence. *Medical Education, 46*(5), 464-472. doi: http://dx.doi.org/10.1111/j.1365-2923.2012.04217.x

Mann, K., Gordon, J., & Macleod, A. (2009). Reflection and reflective practice in health professions education: a systematic review. *Advances in Health Sciences Education, 14*(4), 595-621. doi: 10.1007/s10459-007-9090-2

Manning, A., Cronin, P., Monaghan, A., & Rawlings-Anderson, K. (2009). Supporting students in practice: An exploration of reflective groups as a means of support. *Nurse Education in Practice, 9*(3), 176-183. doi: 10.1016/j.nepr.2008.07.001

Mariani, B., Cantrell, M. A., Meakim, C., Prieto, P., & Dreifuerst, K. T. (2013). Structured debriefing and students' clinical judgment abilities in simulation. *Clinical Simulation in Nursing, 9*(5), e147-155. doi: 10.1016/j.ecns.2011.11.009

McCarthy, P., Schuster, P., Zehr, P., & McDougal, D. (1999). Evaluation of critical thinking in a baccalaureate nursing program... are there differences between sophomore and senior students. *Journal of Nursing Education, 38*(3), 142-144.

McGinty, S. M. Y. (2001). *Levels of reflection of physical therapy students using structured journals during a 9-week clinical education experience.* (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3035639)

McMahon, G. T., Monaghan, C., Falchuk, K., Gordon, J. A., & Alexander, E. K. (2005). A simulator-based curriculum to promote comparative and reflective analysis in an internal medicine clerkship. *Acad Med, 80*(1), 84-89. doi: 80/1/84 [pii]

Medley, C. F., & Horne, C. (2005). Using simulation technology for undergraduate nursing education. *Journal of Nursing Education, 44*(1), 31-34.

Murphy, J. I. (2004). Using focused reflection and articulation to promote clinical reasoning: An evidence-based teaching strategy. *Nursing Education Perspectives, 25*(5), 226-231.

National Council of the State Boards of Nursing. (2012). *NCLEX-RN examination: Test plan for the national council licensure examination for registered nurses.* Retrieved from National Council of the State Boards of Nursing website: https://www.ncsbn.org/2013_NCLEX_RN_Test_Plan.pdf

National League for Nursing Accrediting Commission. (2006). *Accreditation manual with interpretive guidelines by program type for postsecondary and higher degree programs in nursing.* New York: NLNAC. Retrieved from National League for Nursing Accrediting Commission website: http://www.nlnac.org/manuals/NLNACManual2006.pdf

Navedo, D. D. (2006). *A descriptive study of nursing judgment in senior nursing students: And the relationship with reflective judgment.* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3221048)

Nehring, W. M. (2008). U.S. Boards of Nursing and the use of high-fidelity patient simulators in nursing education. *Journal of Professional Nursing, 24*(2), 109-117.

Nishigori, H., Otani, T., Plint, S., Uchino, M., & Ban, N. (2009). I came, I saw, I reflected: A qualitative study into learning outcomes of international electives for Japanese and British medical students. *Med Teach, 31*(5), e196-201. doi: 10.1080/01421590802516764

Notarianni, M. A. (1991*). An investigation of the critical thinking ability of associate and baccalaureate degree nursing students.* (Doctoral Dissertation). ProQuest Dissertations and Theses database. (UMI No. 303993029)

O'Sullivan, P., Aronson, L., Chittenden, E., Niehaus, B., & Learman, L. (2010). *Reflective ability rubric and user guide.* Retrieved from: www.mededportal.org/publication/8133

O'Donovan, M. (2006). Reflecting during clinical placement – Discovering factors that influence pre-registration psychiatric nursing students. *Nurse Education in Practice, 6*(3), 134-140.

Padden, M. L. (2011). The effects of guided reflective journaling on nursing students' level of reflection, self-awareness, and perceived clinical decision making skills. (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3500773)

Pee, B., Woodman, T., Fry, H., & Davenport, E. S. (2002). Appraising and assessing reflection in students' writing on a structured worksheet. *Medical Education, 36*(6), 575-585.

Perera, J., Mohamadou, G., & Kaur, S. (2010). The use of objective structured self-assessment and peer-feedback (OSSP) for learning communication skills: Evaluation using a controlled trial. *Adv Health Sci Educ Theory Pract, 15*(2), 185-193. doi: 10.1007/s10459-009-9191-1 [doi]

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years.* Orlando: Holt, Rinehart and Winston, Inc.

Plack, M. M., Driscoll, M., Blissett, S., McKenna, R., & Plack, T. P. (2005). A method for assessing reflective journal writing. *Journal of Allied Health, 34*(4), 199-208.

Plack, M. M., Driscoll, M., Marquez, M., Cuppernull, L., Maring, J., & Greenberg, L. (2007). Assessing reflective writing on a pediatric clerkship by using a modified Bloom's taxonomy. *Ambulatory Pediatrics, 7*(4), 285-291.

Plack, M. M., Dunfee, H., Rindflesch, A., & Driscoll, M. (2008). Virtual action learning sets: A model for facilitating reflection in the clinical setting. *Journal of Physical Therapy Education, 22*(3), 33-42.

Profetto-McGrath, J. (2003). The relationship of critical thinking skills and critical thinking dispositions of baccalaureate nursing students. *Journal of Advanced Nursing, 43*(6), 569-577. doi: 10.1046/j.1365-2648.2003.02755.x

Rapps, J. A. (1998). *Testing a theoretical model of critical thinking and cognitive development.* (Doctoral Dissertation.). Available from ProQuest Dissertations & Theses database. (UMI No. AAI9830331)

Ravert, P. K. M. (2004). *Use of a human patient simulator with undergraduate nursing students: A prototype evaluation of critical thinking and self-efficacy.* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database database. (UMI No. 305134378)

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

Rhodes, M., & Curran, C. (2005). Use of the human patient simulator to teach clinical judgment skills in a baccalaureate nursing program. *CIN: Computers, Informatics, Nursing, 23*(5), 256-264.

Rowe, M. (2012). The use of assisted performance within an online social network to develop reflective reasoning in undergraduate physiotherapy students. *Medical Teacher, 34*(7), e469-475. doi: 10.3109/0142159x.2012.668634

Rudolph, J. W., Simon, R., Dufresne, R. L., & Raemer, D. B. (2006). There's no such thing as "nonjudgmental" debriefing: a theory and method for debriefing with good judgment. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare, 1*(1), 49-55.

Rudolph, J. W., Simon, R., Raemer, D. B., & Eppich, W. J. (2008). Debriefing as formative assessment: Closing performance gaps in medical education. *Academic Emergency Medicine, 15*(11), 1010-1016.

Rudolph, J. W., Simon, R., Rivard, P., Dufresne, R. L., & Raemer, D. B. (2007). Debriefing with Good Judgment: Combining rigorous feedback with genuine inquiry. *Anesthesiology Clinics, 25*(2), 361-376. doi: 10.1016/j.anclin.2007.03.007

Ruggenberg, S. (2008). *The effect of simulated clinical experience on knowledge, near transfer, and far transfer in nursing education.* (Doctoral Dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3305962)

Sargeant, J. M., Mann, K. V., van der Vleuten, C. P., & Metsemakers, J. F. (2009). Reflection: A link between receiving and using assessment feedback. *Advances in*

Health Sciences Education: Theory and Practice, 14(3), 399-410. doi: 10.1007/s10459-008-9124-4

Scherer, Y. K., Bruce, S. A., & Runkawatt, V. (2007). A comparison of clinical simulation and case study presentation on nurse practitioner students' knowledge and confidence in managing a cardiac event. *International Journal of Nursing Education Scholarship, 4*(1), Article 22.

Schon, D. A. (1983). *The reflective practitioner*. New York: Basic Books.

Schon, D. A. (1987). *Educating the reflective practitioner. Toward a new design for teaching and learning in the professions*. San Franciso: Jossey-Bass.

Schwartz, B., & Bohay, R. (2012). Can patients help teach professionalism and empathy to dental students? Adding patient videos to a lecture course. *J Dent Educ, 76*(2), 174-184.

Sedlak, C. A. (1997). Critical thinking of beginning baccalaureate nursing students during the first clinical nursing course. *Journal of Nursing Education, 36*(1), 11-18.

Shepherd, I. A., Kelly, C. M., Skene, F. M., & White, K. T. (2007). Enhancing graduate nurses' health assessment knowledge and skills using low-fidelity adult human simulation. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare, 2*(1), 16-24. doi: 10.1097/SIH.0b013e318030c8dd

SIGN. (2011). *SIGN 50: A guideline developer's handbook.* Retrieved from Scottish Intercollegiate Guidelines Network, Healthcare Improvement Scotland website: http://www.sign.ac.uk/guidelines/fulltext/50/index.html

Silvia, B., Valerio, D., & Lorenza, G. (2013). The reflective journal: A tool for enhancing experience-based learning in nursing students in clinical practice. *Journal of Nursing Education & Practice, 3*(3), 102-111. doi: 10.5430/jnep.v3n3p102

Skovsgaard, A. (2004). Dialogue and reflection between student nurses and their instructor in clinical practice. *Nordic Journal of Nursing Research & Clinical Studies / Vård i Norden, 24*(1), 45-46.

Smith, J. R., & Cole, F. S. (2009). Patient Safety: Effective Interdisciplinary Teamwork Through Simulation and Debriefing in the Neonatal ICU. *Critical Care Nursing Clinics of North America, 21*(2), 163-179. doi: 10.1016/j.ccell.2009.01.006

Stirling, K., Smith, G., & Hogg, G. (2012). The benefits of a ward simulation exercise as a learning experience. *British Journal of Nursing (BJN), 21*(2), 116-122.

Stone, C. A., Davidson, L. J., Evans, J. L., & Hansen, M. A. (2001). Validity evidence for using a general critical thinking test to measure nursing students' critical thinking. *Holistic Nursing Practice, 15*(4), 65-74.

Thompson, B. M., Teal, C. R., Scott, S. M., Manning, S. N., Greenfield, E., Shada, R., & Haidet, P. (2010). Following the clues: Teaching medical students to explore patients' contexts. *Patient Educ Couns, 80*(3), 345-350. doi: 10.1016/j.pec.2010.06.035

Tofil, N. M., Benner, K. W., Worthington, M. A., Zinkan, L., & White, M. L. (2010). Use of simulation to enhance learning in a pediatric elective. *American Journal of Pharmaceutical Education, 74*(2), 21. doi: 10.5688/aj740221

Tsang, A. K. L. (2012). Oral health students as reflective practitioners: Changing patterns of student clinical reflections over a period of 12 months. *Journal of Dental Hygiene, 86*(2), 120-129.

Vaughan-Wrobel, B. C., O'Sullivan, P., & Smith, L. (1997). Evaluating critical thinking skills of baccalaureate nursing students. *Journal of Nursing Education, 36*(10), 485-488.

Wald, H. S., Borkan, J., Taylor, J. S., Anthony, D., & Reis, S. P. (2012). Fostering and evaluating reflective capacity in medical education: Developing the REFLECT rubric for assessing reflective writing. *Academic Medicine, 87*(1), 41-50. doi: 10.1097/ACM.0b013e31823b55fa

Wallace, D. (1996). Experiential learning and critical thinking in nursing. *Nursing Standard, 10*(31), 43-47.

Walsh, C. M., & Seldomridge, L. A. (2006). Critical thinking: Back to square two. *Journal of Nursing Education, 45*(6), 212-219.

Williams, R. M., Wessel, J., Gemus, M., & Foster-Seargeant, E. (2002). Journal writing to promote reflection by physical therapy students during clinical placements. *Physiotherapy Theory & Practice, 18*(1), 5-15.

APPENDIX A

Skills and Sub-Skills of Critical Thinking
APA Consensus Definition

Interpretation

Categorization

- Decoding sentences
- Clarifying meaning

Analysis

Examining ideas

- Identifying arguments
- Analyzing arguments

Evaluation

Assessing claims

- Assessing arguments

Inference

Querying evidence

- Conjecturing alternatives
- Drawing conclusions

Explanation

Stating results

- Justifying procedures
- Presenting arguments

Self-regulation

Self-examination

- Self-correction

Note. Adapted from "Critical thinking: A statement of expert consensus for the purposes of educational assessment and instruction," by P.A. Facione. 1990.

APPENDIX B

Affective Disposition Related to Critical Thinking
APA Consensus Definition

General Approach to Life

- Inquisitive
- Desires to be well-informed
- Aware of opportunities for critical thinking
- Belief in the process of reasoning
- Self-confident in own ability to reason
- Open-minded
- Flexible in consideration of different points of view
- Seeks to understand points of view of others
- Uses a balanced approach when evaluating reasoned arguments
- Aware of own biases
- Able to suspend or alter judgments and uses consideration in forming judgments
- Able to change beliefs when warranted by further reflection

Approach to Specific Dilemmas

- Clarifies question or problem
- Organizes complicated information
- Diligently seeks all related information
- Prudent in selection and application of criteria
- Focuses attention on current concern
- Persists through difficulties
- Appropriately chooses degree of precision required

Note. Adapted from "Critical thinking: A statement of expert consensus for the purposes of educational assessment and instruction," by P.A. Facione. 1990.

APPENDIX C

Scottish Intercollegiate Guidelines Network

Levels of Evidence

1++ High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias

1+ Well-conducted meta-analyses, systematic reviews, or RCTs with a low risk of bias

1- Meta-analyses, systematic reviews, or RCTs with a high risk of bias

2++ High quality systematic reviews of case control or cohort studies
High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal

2+ Well-conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal

2- Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal

3 Non-analytic studies, e.g. case reports, case series

4 Expert opinion

Note. Adapted from " SIGN 50: A guideline developer's handbook," by SIGN, 2011.

APPENDIX D

Evidence Tables

Table D.1 Quantitative and Mixed Methods

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Aronson, L., Niehaus, B., Hill-Sakurai, L., Lai, C., & O'Sullivan, P. S. (2012) 2++	Cohort quasi-experimental pre-test post-test, 4 groups testing 2 variables	Convenience sample of all 149/167 third year medicals student. Random assignment.	Previously validated scoring rubric; protocols for feedback	Uneven arms of study, 18 students were excluded d/t only taking part of class	Blinding, 4 different groups allowed comparison of effect of both variables and possible interaction.	Students were divided into 2 groups that either received LEaP critical reflection guidelines or just a definition of critical reflection. All students received feedback on content but half of each group also received feedback on their reflective ability. 1st & 3rd reflections were scored. When students were provided critical reflection guidelines, their reflective ability was greater than when they received the definition only. Feedback improved reflective ability but only when both aspects: content & ability were addressed.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Aronson, L., Niehaus, B., Lindow, J., Robertson, P., & O'Sullivan, P. (2011) 2-	Cohort quasi-experimental with control group post-test only	Convenience sample of 115/122 third year medical students. Assigned to group based rotation. Rotations 1 & 2 were controls and 3,4,5, & 6 were intervention	Previously validated scoring rubric; 0.89 inter-rater reliability in previous study using 2 raters.	5 essays were not able to be used. 2 students did not complete course. Non-random assignment. Scores from different rotations were combined.	Guide was used by students to write and instructors to grade. No maturation effect detected.	Developed a reflective learning guide based on a SOAP note format that improved the level of critical reflection in students' written work. Guide was given to students before reflection & used by faculty to direct the feedback. Before writing a one page reflection students received either a short prompt or the guideline. Intervention groups writing was rated sig, higher in critical reflection that intervention groups. Student comments & discussion were used to revamp the guide.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Bae, M-J. (2012) 3	Cohort study	Convenience sample of 23 music therapy students in 3 levels of practicum.	Inter-rater reliability was 73.84%	Dichotomous scale may not reveal small improvements in writing level.	Rater was blinded to level of student	Students completed a reflective assignment after practicum sessions. The data was coded on 4 dichotomous scales: constructiveness (emotional vs. objective), focus of attention (self vs. others), reactive vs. proactive, & ambiguous vs. specific. Over the course of 3 semesters, students did not change in the areas of constructiveness or focus of attention. Students' comments were more proactive & specific when writing about levels II & III. Being proactive & specific was felt to be more of a skill. The lack of change in constructiveness & focus of attention was felt to be related to developmental level & therefore unlikely to undergo any significant change in 3 semesters.
Barry, P., & O'Callaghan, C. (2008) 3	Case Study	1 music therapy student	N/A	Sample of 1	Followed progress of 40 days of clinical practice over 5 months.	Student's reflective journal includes: descriptive journal writing, self-critiquing, integration of new insights, & evaluation. Benefits of journal writing: understanding influence of context, reframing clinical problems with theory, self-evaluation & redirection from clinical supervision, develop insight, self-awareness, & analytical thinking, & clarifying utility of music therapy.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Beyer, D. A. (2012) 3	Case Study	Faculty of 16 nursing students in a med-surg course.	N/A	No grading guidelines given. No examples of prompts that students had to answer.	Describes in detail how to set up this type of assignment.	Simulation groups of 4 students participated in successive unfolding simulations with a brief debriefing after all had been completed. The last part of one group's scenario was the giving of report to class members. Once all groups had completed the simulation, this was followed by a class debriefing emphasizing the progression of the patient's symptoms & care. Each group collaboratively created a Wiki. The Wiki assignment was based on the perceived needs of the class & all groups responded to the same questions. The history function of the Wiki allowed for the identification of individual content & editing, and assessments could be made on each student's contribution to the completed project. During the week long creation of the wiki, students & faculty made comments on the work in progress. Evaluation of the wiki allowed for identification of areas needed further clarification or additional instruction.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Blatt, B., Plack, M., Maring, J., Mintz, M., & Simmens, S. J. (2007) 2-	Cohort study using mixed methods	Convenience sample of 149 third year medical students	Medical skills and pt satisfaction behavior checklists had face validity.	Not all students revisited the cases, self-selection bias. Changes in median revisit scores were much smaller than the standard deviation.	Many researchers involved in the review of the coding of the data.	Students rotated through 6 different standardized pt cases in rotating order & could revisit their last 3 pts for an additional 5 minutes. After each pt students completed a 5 min. reflection. The pt gave feedback from checklists but did not reveal scoring. After last 3 visits, students were surveyed, and either completed a Likert scale or explained why they decided not to revisit that pt. Sig. improvements were found in the medical skills revisit scores for all cases. Overall statistical sig. was achieved for pt satisfaction scores. Inverse relationship between first score & revisit score. 63% of the revisit opportunities were taken; 12% of the students never revisited a pt. Themes from students that elected not to revisit a pt included: sufficient information gathered to make decision, & all issues have been addressed with the pt. Positive themes were that the intervention: improved clinical decision making, pt education, clinical realism; & student & pt satisfaction. 16% of the revisits generated negative themes: decreased student satisfaction, neg. impact on the pt, or that the intervention was unnecessary.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Boyd, L. D. (2008) 3	Mixed Methods: Case study with qualitative methods	Convenience sample of 16 third year dental student volunteers during their first year of clinicals.	Validation of coding scheme by 3 faculty using random transcripts. Cronbach's alpha was 0.76 for coding scheme of King and Kitchener's scale	Self-selection bias, 23.2% volunteered. Most students failed to maintain journal writing.	Only small differences between study group and national pop. Interviews tape-recorded and transcribed. Field notes taken. 5 member committee developed coding scheme. Coding protocol used.	Students kept reflective journals and participated 3 spaced in interviews to provide material for estimating their reflective judgment. Average growth in reflective judgment from Stage 4.89 to 5.59 on King and Kitchener's Reflective Judgment Scale (1-7). Students were given guiding questions for journal and samples. Reflective judgment in treatment planning grew more than on other aspects that the students were not as exposed to. Reflective thinking that occurred was thought to be caused by "Trigger events" and disequilibrium and this was thought to be the main reason students' reflective judgment grew in such a short time.
Brown, F. S. (2011) 3	Multiple case study with demographic survey, observation of faculty conducting simulation and debriefing, and interview	9 nurse educators teaching in ADN or BSN programs who had been using HFPS routinely for over one year.	N/A	5 interviews took place immediately after observation. 4 took place up to one week later.	Each faculty member had conducted between 50 and 750 simulation and debriefings Triangulated data.	Instructors were observed for use of reflection techniques during debriefing and evaluated using extensive criteria (pp. 69-74). Debriefings that had greater student than facilitator talk time were more reflective in nature. Use of video to evaluate debriefings was underutilized.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Bruce, C., Parker, A., & Herbert, R. (2001) 3	Case study	Convenience sample of final year speech and language therapy students	N/A	No mention of size of sample	Planned progressive withdrawal of direct supervision and higher expectations for student reflections.	Students were paired & alternated taking the role of learning or learning facilitator and were overseen by a clinician. The learning facilitator takes notes of the pt session & completes a feedback form. The learner answers several reflection questions. Reflection after action & before action is encouraged in Stage 1. The clinician gives feedback & assists the learner in evaluating the session. In Stage 2, the learner is supposed to reflect during action as well as the previous activities. During Stage 3, the learner is expected to gain an overview & insights. The clinician does not view the pt session but offers feedback & guidance. All students are asked to review the entire experience. Students generally felt that the experience was positive & that they gained a greater understanding of themselves & their clinical practice. They also felt that they become better at giving peer feedback. Some felt the process was time consuming & inflexible.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Cahalin, L. P., Markowski, A., Hickey, M., & Hayward, L. (2011) 3	Case study	Random sampling of 14 fifth year Doctor of Physiotherapy students. Random selection of group member to examine standardized pt.	N/A	No evaluation of reflection by students. Small sample size.	Random sampling and selection. Triangulation of sources	3-5 students were grouped into a virtual community of practice that first prepared for the simulation via online problem solving of a case. After the simulation, the students' work was evaluated by: debriefing, video, a reflective paper, and instructor, standardized pt. & peer assessment. Working with standardized pt's allowed the assessment of both professional behavior & clinical decision making skills (rubrics p. 8). Instructors also provided feedback to all students on their participation in their online discussion group, assessment of the standardized patient interactions, & group decision trees. Peers & the pt completed a professional behavior rubric. Peers & faculty used a clinical decision making rubric to evaluate the diagnosis, prognosis, & plan of care. Students felt that the exercise promoted critical thinking & improved their communication skills. All of the students wanted more standardized patients. Allowed instructors to find gaps in the curriculum where the students either needed more practice or were not applying theory to the patient exam.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Carr, S., & Carmody, D. (2006) 3	Mixed Methods: Descriptive study of 2 successive cohorts with qualitative theme identification of the content of the writing.	Fifth year medical student volunteers in a yearlong women's health course. 87/115 in first cohort and 62/72 volunteered.	Not addressed	1 coder, no theoretical basis for coding, possible history or maturation effect on researcher	Successive samples	Students wrote a reflective case summary that included a reflection before action component. The summary is discussed at mid-term with a facilitator. Another summary is turned in for a grade. 1 of 4 levels of reflection was assigned to the summative writing: listing, describing, applying, and integrating. Most students reflected at the level of application (46%), 28% at describing, 16% at integration, and 10% at listing. Reflection allowed students to see the positives of a situation, helped students discover the way in which they wished to grow, and exposed students to different perspectives.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Cook, J. L. (2010) 2-	Retrospective cohort study	75 physical therapy students who had matriculated from 2003-2009 and had National Physical Therapy Exam and Clinical Performance Instrument scores, and journals from first and last clinical course.	Pilot study found initial inter-rater reliability of .823 and after refinement 0.940. Inter-rater reliability was .849 for this study.	Journal entries had no stipulations on content. Confounding d/t not knowing if lack of effect is d/t reflection level not being a predictor or lack of guidance in reflective thinking.	3 coders reviewed 900 journals entries	Used 3 level reflection rating: non-reflection, reflection, and critical reflection. Student reflection level was not a predictor of National Physical Therapy Exam or Clinical Performance Instrument scores. There was no difference in student reflection level between their first clinical course and their last. Student received little guidance on what to write about in their journal.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Corrigan, R., & Hardham, G. (2011) 3	Non-analytical report Pre-experience & simulation, post-experience survey. Designated how to received feedback (individually, with simulation group, in class, or not at all). Students evaluated feedback received. Online surveys contained both open and closed response items.	60/61 physiotherapy students in their 3rd year of undergraduate work	N/A	25/60 students completed the pre-experience survey, 25/60 completed the post-experience survey, and 20/60 attempted the feedback evaluation. Small sample size and low response rate.	Voluntary participation; Anonymous responses; 61 students volunteered and 60 completed the simulation and feedback sessions.	Pre-experience themes: anticipated technical problems, inadequate knowledge or preparation, lack of time to complete the simulation, fear of being judged, lack of direct supervision, & no immediate feedback. Only 3/25 had no concerns. Students' roles: pt, video recorder, physiotherapist. Post-experience survey revealed positive themes: additional practice opportunity, time limit on simulation made for realistic practice session for exam, preparation requirement, getting to choose a particularly difficult case for the student, & use of video & discuss their performance. Negative themes from the post-experience survey centered on technical issues. Some students felt that they had to rush, since other students ran over time. Stress because they felt unprepared & from seeing themselves on camera. Helped students gauge how much more work they needed to prepare for exams. 16 students felt they were better able to evaluate their performance. 13 asked to receive feedback with group. Preferred verbal feedback and 17 found the feedback helpful.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Daly, G. (2010) 3	Descriptive study with survey after completing 1st year.	13 speech-language therapy student volunteers	N/A	No description of sample pool, self-selection bias	5 point Likert scale used to rate interventions.	Students complete a variety of reflective assignments & rated them out of 5: video (4.33) & transcript (4.15) review of client sessions, setting clinical goals (3.76), dyad observations (3.38), & evaluation of their clinical effectiveness (4.30). When reviewing the video, students were given specific tasks that focused on the student's behavior & then create a plan for changing their behavior. Students participate in a team session of 4-5 students & a facilitator to identify & discuss clinical concerns. Students also transcribed a session & evaluated behaviors which were then discussed in the team sessions. In order to complete the assignment, the students had to reflect after action & before action.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Delany, C., & Watkin, D. (2009) 3	Mixed methods Interpretive and constructionist methods used.	Convenience sample of 14 third year physiotherapy students during 1st clinical placement	N/A	Small sample size	Informed consent. Triangulation of data. 2 coders. Facilitator not faculty for students.	6 weeks of 3 hours of critical reflection intervention. Ground rules for participating established. Students were encouraged to deconstruct critical incidents in an appropriate place and time. Emotions addressed. Sessions were assigned objectives relating to narrative discussion, deconstruction, identifying values, examining assumptions, putting new theories into practice, and continuing to evaluate clinical practice. Feedback from students via email, field notes, and transcribed sessions. Reflection sessions seen as sharing & bonding experience in safe environment, with peer learning, time to sort things out and consider holistic elements. Benefits of program still present after 6 weeks but diminished. 2 didn't care for program. Facilitator not working with students in clinical or class.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Dreifuerst, K. T. (2012) 2+	Quasi-experimental pre-test post-test	Convenience sample 238 senior BSN student volunteers in 3 successive classes	HSRT has established reliability & 3 subscales have high internal consistency(Evaluation, inductive, deductive reasoning) Strong reliability for tool & subscales. Content & construct validity. DASH-SV established content and criterion validity Cronbach's alpha of 0.82.	Self-selection bias. No criterion validity of the HSRT. The HSRT is not specific to nursing. No reliability data for the DASH.	Random assignment by clinical group. Less than 1% LTF.	DASH-SV asks students to assess debriefer's ability to: create engaging learning experience, organize debriefing, stimulate discussions, & assist student in identifying performance gaps. DML begins with the affective response & moves to analysis. Uses guided reflection to improve a student's ability to reason clinically. HSRT given 3 weeks before and 3 weeks after simulation. Student roles in the simulation were: primary nurse, secondary nurse, family member, 2 recorders, observers or health professional. Students debriefed using the DML method. DASH-SV & DML supplement questions were given after the simulation. DML questions: student worksheet, reflective thinking, treating patients with similar conditions, & time spent debriefing. DML was significant for improvement in HSRT scores. DASH-SV scores were higher for the DML group except for on pre-briefing, which was the same for both groups. Significant positive relationship between all items on the DML supplement questions, the DASH-SV & the HSRT, except for student worksheet & pre-briefing items. Students that highly rated the debriefing scored higher on post-test clinical reasoning.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Dunfee, H., Rindflesch, A., Driscoll, M., Hollman, J., & Plack, M. M. (2008) 3	Descriptive case study	Convenience sample of 2 groups of 3 or 4 students in their final clinical physical therapy course	Agreement for reflection elements range was 72.9% to 95.9% with kappa coefficients from 0.11 to 0.45 and PABAK coefficients from 0.46 to 0.92. Level of cognitive processing agreement ranged from 68.8% to 95.1%, with a kappa of 0.35 to 0.45 and PABAK of 0.49 to 0.57.	Small sample size. Limited variability in data and high prevalence made kappa coefficients deflated and hampered interpretation.	3 raters. Rating was evaluated with the kappa and the prevalence-adjusted bias-adjusted kappa (PABAK) coefficient to account for high agreement and low disagreement in the data.	Action learning sets are small groups that work through problems together and seek to learn from the experience through reflection with or without a facilitator. Over 4 weeks, students used an online discussion board to reflect on critical clinical incidents, provide commentary, and pose questions to their group members to assist in developing a solution. All students received a class on reflective practice and orientation to the discussion board. 122 entries were coded and the percentage for the raters was averaged. The comments were assessed for reflection during (4.3%), after (91.0%), and before action (29.8%). The entries were also coded as to the level of cognitive processing: data gathering (non-reflective, 97.5%), data analysis (reflective, 84.2%)), and conclusion drawing (critically reflective, 58.8%). More explicit criteria for coding may improve rater agreement.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Durso, S. C. (2006) 3	Case Study	1 first year medical student	N/A	Does not appear to be comprehensive report	Quotes used from student's reflective writing	Describes process of student's experience with a written reflective log kept while shadowing a clinician to be used to guide weekly discussion with the clinician. Included is the student's report of lessons learned. Reflective log helped student fit experiences into a pattern. Issues drawn from the experience include: awareness that the clinician has to work at communication; building relationships with patients relies on the development of skills; reflection led to evaluation and recognition of the considerable effort needed to create a successful relationship with pts; and realization that the student would need to master these communication skills. The reflective experience transformed the student's view of communicating with patients and the work that she would need to do to acquire the communication skills.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Dye, D. (2005) 3	Semi-structured focus group interview	Random sample of 4/15 physical therapy students	N/A	Small sample size,	Random sampling and the students all chose to complete group interview.	Students wrote weekly self-SOAP notes: subjective feelings, summary of skills performed, assessment of student's own performance, and plan for improvement. A previous self-SOAP note was provided as a guide. Notes were submitted by email or fax. The group was positive about the intervention and liked: ease of use of self-SOAP note; instructions that covered topics to be written about, having a guide; immediate feedback from clinical instructor that encouraged further reflection; self-improvement was highlighted, kept track of and encouraged ; and guided future learning. Time consuming for students and faculty.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Epp, S. (2008) 1+	Systematic Review	150 abstracts reviewed from OVID, EBSCO and Blackwell synergy from articles published from 1992 to 2007 covering reflective journaling by undergraduate nursing students.	Focus was on undergraduate educational process.	One reviewer.	Each article described in detail.	Undergraduate nursing students primarily reflect at the lower levels but are capable of higher level reflection. Reflective writing develops over time and has produced shifts in students' perspectives and changes in their practice. An environment of trust is needed to support reflective writing. Undergraduates may not have experienced and learned from reflective writing in the same way as graduate nurses.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Ertmer, P. A., Strobel, J., Cheng, X., Chen, X., Kim, H., Olesova, L., . . . Tomory, A., (2010) 3	Comparative Case Study	Convenience sample of 17 out of 164 students in a junior level adult nursing care course participated in a simulation, video review, and group debriefing. 14 students took part in individual interviews.	N/A	Small sample size. Selection bias.	Written reflection while video of simulation was paused. Taped & transcribed collaborative debriefings and individual interviews. 3 coders worked both individually and cooperatively; inductively and deductively with the data.	Identified the critical thinking and habits of mind used by students in different roles of a simulation. Individual interviews took place one week after simulation. Student's felt that the role they played and lack of experience with simulation limited their ability to actively participate and to learn. 3 habits of mind of critical thinkers were used by the students: reflection, contextual perspective, and confidence. Two skills were demonstrated by the students: applying standards and logical reasoning. 15/17 participants exhibited reflection. Self-evaluation, a subcategory of reflection was identified in 14/18. Contextual perspective was identified in 13/17. Only 4 students mentioned 3 or more perspectives. 11/17 students demonstrated applying standards. Logical reasoning was demonstrated by 15/17 students, a total of 33 times.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Fakude, L. P., & Bruce, J. C. (2003) 2-	Quasi-experimental post-test only with control group	Convenience sample of 43/53 first year nursing student volunteers	Content validity of evaluation tool based on Gibbs' Reflective Cycle established by peer review.	Non-randomized sample, possible ceiling effect. Evaluation tool may not be sophisticated enough to detect graduations of ability to reflect.	Cross contamination was avoided by basing groups on which campus students attended.	Students in intervention group wrote weekly journal entries on clinical experience using guidelines. Then all students were asked to write a reflective paper. All work was evaluated as to whether the questions posed by the guideline were answered. Intervention group performed better on the 2 the highest levels of reflection: exploring alternatives of action and formulating responses in similar future situations. A ceiling effect may have affected the lack of a sig. difference in the most categories: description of event (100%, 100%), exploring thoughts and feelings (100%, 100%), evaluation of good/bad aspects (100%, 100%), and analysis for interpretation/meaning (85%, 91.3%). There was improvement in all categories for the intervention group from journal to paper, but it was not sig.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Flanagan, B., Nestel, D., & Joseph, M. (2004) 3	Non-analytical descriptive case study reporting evaluation data from a simulation	Convenience sample of 132 4th year medical students, 30 interns, and 137 practitioners. Only data from students will be used for this table.	N/A	No reporting of demographics	Actual comments reported as well as themes	Reflective debriefing was used to collect evaluative comments from the participants. Interns were able to identify leadership and communication issues. Immediate feedback after simulation was perceived as an extremely helpful learning method. Participants were able to identify gaps in knowledge but had difficulties implementing what they knew. Simulation was able to test whether a student was able to translate knowledge into practice. Cues used during the scenario were: phone calls, pagers, and other unspecified distractions.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Grant, A., Kinnersley, P., Metcalf, E., Pill, R., & Houston, H. (2006) 3	Mixed Methods – Grounded Theory & Case Study	65/232 Third year medical student volunteers.	N/A	167 students were excluded because they did not attend introductory class. Small sample size. Selection bias.	Participants, drop outs, and non-participants each had focus group interviews focusing on issues pertinent to their group. 2 coders & software. Student quotes included	2 reflective seminars. Students kept a journal based on critical incidents. Templates shared with students. Discussion groups run by 10 different instructors. As students dropped out, groups were consolidated in 4 groups. Semi-structured interviews recorded and transcribed. Saturation reached on: prior learning & context, reasons for non-participation and dropping out. All participants were interviewed. Reasons for dropping out were logistics and time problems. Non-participation views were: that reflection wouldn't be helpful, not useful to student, and logistics and time problems. The learning context was: a culture of not discussing work, & large volume of work. Participants: valued peer's reflections, gained confidence, felt emotionally supported, discovered norms of peers, and appreciated feedback from instructors. No sig. dif. in exam grades for participants, drop outs, non-participants, or non-attenders.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Hallmark, E. F. (2010) 2-	Mixed Methods – Qualitative and Post-test only Quasi-experimental on the relationship between student variables and faculty training on debriefing.	84/157 third year nursing student volunteers	HESI is a valid and reliable tool. Reflective Learning Continuum was adapted for nursing students.	Varied backgrounds of the faculty debriefers. HESI may not be an appropriate measure.	Random assignment; 2 coders; student quotes included	Prior patho grades & reflective thinking inventory, 2 different simulation scenarios (heparin & blood), trained or untrained faculty debriefing, Post-test HESI and satisfaction survey. Faculty were trained via a NLN course. No difference in HESI scores for trained or untrained faculty. After controlling for age, gender, grades, and educational level, faculty training was a sig. factor in student satisfaction. Students believed that simulation and debriefing improved critical thinking scores and enhanced learning. Reflective Learning Continuum Likert scale survey levels revealed a sig. difference in students led by trained faculty.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Harrison, P. A., & Fopma-Loy, J. L. (2010) 3	Case study	Convenience sample of 16 associate degree nursing students in an psychiatric course	N/A	No standard method of evaluation of logs.	Each week entry was examined for all students before moving on. 2 coders. Student quotes.	Students were given progressively more in depth reflective writing prompts to respond to each week moving through self-awareness, social awareness, self-management, to a reflection on the patterns in previous journal entries. Entries allowed faculty to assess student strengths and weakness. Prompts were judged as needing revisions and additions. Students and faculty found the intervention time consuming and emotionally draining. Prompts were effective in getting students to expand their emotional intelligence. Clinical instructors need to be explicit in their learning goals and assist students in making connections between their journal writing and clinical problems. A psychological safe space is essential. Trust must be generated. Faculty need to share and develop with the students.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Hatlevik, I. K. R. (2012) 3	Secondary analysis of a cross-sectional correlational study. Data retrieved from national pre-professional questionnaire	446 third year nursing students	Single scale measures for most variables. Face and discrimina-tion validity.	Response rate 71%. Single item measurement of variables meant that measurement error unable to be estimated.	Comprehensiv e sample of Norwegian students	Students' ability to see the connections between theory and practice was related to reflective ability and knowledge of the underlying theory. Students' subjective Likert rating of their knowledge of theory, skills, reflective ability, and coherence.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Hill, A. E., Davidson, B. J., & Theodoros, D. G. (2012) 3	Descriptive cohort study	52 undergraduate speech-language therapy students	Inter-rater reliability for reflective elements ranged from 81.48% to 98.77%. Overall assessment of student's depth of reflection had a mean of 96% (range 33.33% to 100%). Face validity for checklist of reflective elements.	94% of writings rated as reflectors by both raters. Either the coding criteria was not refined enough or the sample was too homogenous.	All reflections were coded by 2 raters. Ten students' reflections were used to refine the coding system and the other 42 students' reflections were used for analysis. Substantial to almost perfect agreement was established. 9 were re-rated to establish inter-rater reliability.	Students interviewed 3 different standardized pts either with 1 or 2 partners. Instructor called time out periods, used to provide feedback and prompt student reflection. Instructors and the standardized pt provided feedback. Reflective journals evaluated & coded according to Plack et al.'s (2005) non-reflective, reflective, or critically reflective and nine criteria. Reflective questions given to the students immediately after the standardized patient interview. 94% were reflectors and their writings primarily contained content and process reflection, and reflection after action and for action. 3% were non-reflectors and 3% critical reflectors. Few writings contained reflection during action elements or premise reflections; which comprise critically reflective writing. Researchers postulated that the use of specific prompts for the writing assignment may have caused the students' writing to be more similar in content and level than other studies which did not use prompts. Students did not received instruction on reflective writing or receive feedback on their writing prior to the next interview. Writing immediately after the interview may have affected the lack of depth.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Ho, D. W. L., & Whitehill, T. (2009) 2-	Mixed Methods Quasi-experimental. Motivated Strategies for Learning Questionnaire given after first session. Subjective comments collected.	Convenience sample of 19 third year speech language pathology students	Motivated Strategies for Learning Question-naire is a reliable tool; but was modified to reflect clinical learning.	3 sets of variables examined without separation into groups (immediate, verbal, group). Small sample size with homogenous results a possible confounder	Random assignment to control group. 100%. volunteered. No sig. dif. between control and intervention group.	Intervention group received immediate verbal feedback. Feedback to controls was individualized, written, and delayed. Both groups were asked to write a reflection on their performance using a guideline. All students' assessment scores improved from the mid-semester evaluation to the end. The intervention group had sig. higher ratings on subscale of clinical skills. The MSLQ self-evaluation ratings went down over the semester but the intervention group was sig. higher than the control. Intervention group felt they learned from other students and the students' clients but that it was time consuming. Control group felt they were better able to reflect given the delay and that it was more time efficient. All intervention group and most of control group preferred verbal feedback.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Hulsman, R. L., Harmsen, A. B., & Fabriek, M. (2009) 2+	Cohort study	331 2nd year med students.	Observed behaviors inter-rater reliability was 76.5%. No internal reliability on behavior checklist. Face validity.	The role of the student rotated through the 3 cycles, so that no student was the care provider more than once. Different questions used for each trial.	Over 90% of students filled out a questionnaire and were evaluated. Two evaluators categorized reflections and created rating manual based on 30% of the responses.	3 cycles of simulation followed by video review for reflective activities. Groups of 15 students presented and reviewed key events, reflections, and feedback. Students rotated thru 3 roles of care provider/reflector, feedback provider/presenter, and feedback provider. 3 different scenarios were used. Assignments became progressively harder. In cycles 2 & 3 the reflection questions were categorized as: observations, describing motives or effects, asking for feedback, and indicating a goal or effect. 93% of the students found solutions. ~39% made observations of their behavior, 16% motive or effect, 7% direct question, ~10% indicated a desired goal. Only 26% believed that their medical knowledge was sufficient for exercise. Students had greater difficulty reflecting and enjoyed it less than giving feedback to peers. Both activities were not as highly valued as observing themselves and peers' recordings. Self-reflection made the students more aware of weaknesses while peer feedback revealed strengths. Simulation was the most helpful (95.4%), receiving instructor feedback (93.8%), standardized patient feedback (92.4%) and peer feedback (90.8%).

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Hussin, V. (2013) 3	Case Study	20 third year pharmacy students all were non-native English speakers.	N/A	Limited generalizability since all students were non-native English speakers.	Long term follow-up.	Simulations of patient encounters involving the student pharmacist giving advice or the staff/patient voicing concerns were videotaped and reviewed. Problem areas were identified. Both the staff/patients and the student pharmacists reviewed the simulation tape and then were interviewed. The interviews were analyzed for the staff/patients and student pharmacists' awareness of and explanations for problem areas. The interviews of the staff/patients and the matching student pharmacists were compared for contrasts and similarities. One year later, 2 focus groups of the participants were prompted to reflect. Audio tapes of the focus groups were analyzed for increased competence and professional maturity. The students liked receiving individual feedback and found the simulation and reflection helped them focus on communication areas that needed improvement. Students desired more individualized feedback but staff felt they did not have enough time. Students felt that reviewing the video was helpful because it showed both verbal and nonverbal communication.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Ip, W. Y., Lui M. H., Chien W. T., Lee I. F., Lam L. W., & Lee D.T. (2012) 2-	Quasi-experimental study pre-test, post-test with no control group	Convenience sample of 62/178 sophomore nursing students volunteers; only 38 completed all aspects of study	Cronbach's alpha was 0.82 for the Student Opinion Scale. Inter-rater reliability of 95%. 90% agreement on main themes. Friedman test & Wilcoxon signed-ranks test used to prove statistical sig.	24/62 dropped out and not compared to completers. Most completers kept a diary. Significant results may have been because those students that were good at reflective writing completed study.	2 coders. Student quotes included.	Students went to a 3 hour workshop on reflective skills and received 4 weeks of coaching from their clinical instructor on how to integrate reflective skills into practice. Student Opinion Scale was used to collect survey data. Reflective logs were collected before intervention, at the end of the 2nd week, and at the end of the 4th week. Logs were coded as non-reflective, reflective, or critically reflective. Role of faculty was considered very important but 3 students did not establish a trusting relationship. Benefits of reflective learning were an increased understanding about nursing practice. Barriers to self-reflection were: lack of time, and unavailability of faculty. Suggestions for improvement were to provide more time for reflection: lengthen the clinical placement and ease teaching load of faculty. Students' level of reflective writing was sig. different from pre-test to post-test. There was no sig. difference between the post-test measurements.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Jarris, Y. S., Saunders, P., Gatti, M., & Weissinger, P. (2012) 2-	Quasi-experimental pre-test post-test with control group	Convenience sample of 190 first year medical students	N/A	No randomization. No description of sample.	Negative report.	All students had 2 clinical skills assessment on a standardized pt. 3 months apart. 47 students in intervention group viewed recordings, completed self-assessment, and received feedback from pt. and faculty. Online feedback given on specific behaviors after review of tape. Students reflected on their self-assessment and faculty comments. 12 weeks later all students went through another simulation. No sig. diff. between groups pre-test or post-test was thought to be due to lack of practice and no guidelines for reflection. Students were more critical of themselves than either the pt or faculty. No instruction was given on how to critically reflect. Later parts of feedback model, refining and implementing an improved plan, did not occur.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Kalish, R., Dawiskiba, M., Sung, Y. C., & Blanco, M. (2011) 3	Mixed methods study	11/12 third year medical student volunteers	Cronbach's alpha .75 for question-naire. Paired t- test to examine difference in students' ratings.	Small sample size	Multiple coders. Student quotes included.	Students had to read chapter, article, watch video, and review compassionate care questionnaire before pt exam. Student presents synopsis, receives preceptor feedback, re-examines pt with preceptor, pt gives feedback, student & pt complete questionnaires, student submits diagnosis, receives feedback, videotape is reviewed & tagged by student, preceptor, & 4th year student, all 3 complete questionnaire, and student is debriefed by preceptor. All students participated in taped focus group. Students' self-assessment of compassionate care sig. dropped after video review. Students tagged 21 missed opportunities for compassionate care. Video allowed students to observe themselves more objectively, but felt that being taped took away from the encounter.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Kautz, D. D., Kuiper, R., Pesut, D. J., Knight-Brown, P., & Daneker, D. (2005) 2+	Quasi-experimental design	Purposive sample of all 23 junior nursing students enrolled in a med-surg course	Random sampling of journals used to establish inter-rater reliability	Dichotomous variables	3 coders, prompts and evaluation tools included; all students enrolled in course volunteered	2 weeks of class used to train students how to use the self-regulation prompts (on p. 19) and the OPT model. Clinical faculty worked closely with the students to frame their work, provide guidance, and rate OPT model. Students kept reflective journals for 10 weeks on using the OPT model guided by the self-regulation prompts. Verbal protocol analysis revealed that the students addressed all 3 concepts the OPT model: behavioral (52-54%), thinking through problems (13-16%), and metacognitive (31-34%). Students used primarily connotative statements (62-74%), followed by causal (6-21%), and indicative (4-18%), and comparative (8-10%). Journals were collected each week, but no feedback was given in order to encourage free expression of thoughts. Over ten weeks the students' writing in a reflective log showed evidence of being better able to frame situations, and choose interventions.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Kelly, P. J. (2012) 3	Case Study	Convenience sample of 45 physician assistant students in a medical communication course	N/A	No student quotes used	Very detailed response themes	Students answered a set of reflective questions about the characters in 4 movies, their feelings, and application of lessons learned. Students did not always answer the question in the way it was intended; focusing instead on their emotions, response, and beliefs. The reflective writings revealed how students were internalizing the material.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Ker, J. S. (2003) 3	Descriptive case study	Convenience sample of 6/150 junior medical students	N/A	Students volunteered for this component of the class. Daily written reflections not reviewed. Small sample size. Selection bias.	Well-structured learning plan with appropriate development of facilitator leadership in students.	Students began module with study guides, participation in clinical, analyzing their own learning needs, and skill training. Structured one hour reflection groups were scheduled for 4 weeks. Facilitator gradually reduced role as leader. In session 1: strengths, weaknesses, and professional concerns were discussed. Discussed in session 2 were: technical skills needed for clinical, study guide, and preparation of scripts for simulation. The third session was after a training session with the standardized pt. During this session students: re-evaluated their communication skills and discussed professional concerns. Last session was after the simulation and concerned progress in clinical. An open ended questionnaire was filled out by the students at the end of the 4 sessions. Students wrote a reflective paper based on their reflections during each of the reflection groups as well as integration of skills into practice, and how the intervention helped them. The module was highly rated by students. 2 students who did not do very well in the simulation wrote descriptive (non-reflective) but not evaluative (reflective) reports.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Kok, J., & Chabeli, M. M. (2002) 3	Case Study	Convenience sample of 6 senior nursing student volunteers	N/A	Only 6/17 volunteered, self-selection bias	Saturation of data, triangulation of sources for codes, 2 coders, student quotes included	Focus group after course to discuss reflective journals. Ground rules established. Positive themes were that reflection involved: integration of theory and practice through problem solving, self-evaluation, intellectual growth, and self-awareness. Subthemes of problem solving were that reflection was carried out through: critical and analytical thinking skills, evaluation, and synthesis. Negative themes were: journaling was time consuming, trust was not established, there was a lack of clear expectations, and writing was recounting of the events.
Kuiper, R. (2005) 3	Case study with comparisons to previous study	Convenience sample of 40 senior BSN students in 2 semesters	Percent agreement between coders and researcher's examples was 90%.	Completion rate 78%, 10/40 excluded for not completing weekly entries	Coding by 2 independent faculty	Weekly audiotaped reflective journal to remain confidential and ungraded. Clinical faculty received instruction. Tapes were make immediately following clinical experiences using the "think-aloud" method. Longer entries were produced than when written journals were used in previous study. Higher order thinking was expressed and the pattern of thinking did not tend to change over semester. Verbal protocol analysis of the entries. Connotative (62-72%), indicative (16-23%), comparative (6-8%) and causal (7-10%). Every major critical thinking skill was represented.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Kuiper, R., Heinrich, C., Matthias, A., Graham, M. J., & Bell-Kotwell, L. (2008) 3	Descriptive Design	Purposive sample of 44 senior nursing students	OPT Model tool is a reliable and valid instrument. Inter-rater reliability of 87%.	Small sample size. How sample was chosen was not explained.	Maturation was controlled for.	Students completed a simulation & OPT model worksheet. There was no sig. diff. between OPT scores for simulation and clinical.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Ladyshewsky, R. K., & Gardner, P. (2008) 3	Case Study	Convenience sample of 38 senior undergrad physiotherapy students	N/A	Issues that arose at midterm mentioned but not identified.	Random selection and assignment to groups.	8 discussion groups of 4-5 members & moderator. 1 hour class on reflection. Moderators reduced support. Focus groups at the end. Moderators thought: students were more engaged, guidelines should be developed & introduced earlier in curriculum. Students thought participating: was easy to do, quick, & convenient, allowed editing & work throughout the semester. Students liked: writing informally, peer learning, social connections, building trust, & quick feedback. Some liked to provide support & coaching. Forced students to pause & reflect, and process & structure their thinking. Felt it was a safe place. Students did not like: having technical issues, discussing difficulties, few members, not getting feedback on the final entries, not having a guide, topic assignments, delay in responses, time it took to get process working, too much moderator participation, not being able to access discussion at practice site, not being permitted to view other groups, & having a moderator who was also in clinical. Changes suggested by students: introduce earlier in curriculum, deadlines for posting, issue based boards, & larger groups.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Lai, C. & Hu, C. (2012) 3	Case Study	8 nursing students in a psychiatric clinical	LCJR is a valid and reliable tool.	Small sample size, lack of detail in findings	Used established criteria and tool.	Students were provided with a computer notebook to access web. 3 reflective activities based on: John's, Tanner's, and the OPT model were put online for students to complete and share. LCJR showed gains in student learning from the developmental level to the accomplished level. Survey indicated that students thought the activities helped them learn reflection and nursing skills. The instructor said it helped with early identification of student problems and their critical thinking skills.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Lasater, K. (2007b) 3	Descriptive study	Convenience sample of 8/48 junior nursing students in Nursing Care of the Acutely Ill Adult course participated in focus group. 15 non-traditional students volunteered but only 8 were able to participate.	N/A	Small sample size. Focus group volunteers were all non-traditional students. Self-selection bias.	All 12 person simulation teams were represented	Students felt it would be a more useful learning experience with improved reflection in the debriefing process, more time debriefing, structured observation roles, and definitive & straightforward feedback. Wanted group video review with analysis and facilitator feedback on what the students were thinking as well as doing. Wanted a "follow-up" scenario with a similar pt to show improved performance. A pre-briefing was valued but did need not to cover every detail. Collaborating with other students was helpful. Learning was transferrable to clinical. Simulation was anxiety provoking although a valuable learning experience. Students learned from hearing peers debriefed.
Makoul, G., Zick, A. B., Aakhus, M., Neely, K. J., & Roemer, P. E. (2010) 3	Mixed methods cohort study	315 third year medical students over 2 academic years	N/A	Only 5 students elected to post a 2nd time. No F/U on if students valued the board.	Comprehensive guideline for postings.	An anonymous online discussion board was used to collect guided reflections one or more difficult conversations. A guide for posting and responding was given to students. 93 students requested a faculty member respond to a post. Students identified lessons learned from the experience.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Maloney, S., Storr, M., Morgan, P., Ilic, D. (2013) 1-	RCT, double blind; post-test only	100% of 60 third year physiotherapy students in one setting	Behavioral Checklist for OSCE exam. Face validity	6.7% of students all from intervention group lost to attrition, lost students not compared to others, possible contamination of intervention and control groups	OSCE examiner blinding	5 min. student produced video of assessment of a clinical situation. Online tutors reviewed the videos & provided group feedback on strengths & weaknesses. Students compared & contrasted their performance to a peer's video. At week 8, students were randomized into an intervention group that had to produce a video of a cervical spine assessment (skill A) or a control group that filmed a related assessment. Students preformed 2 OSCEs (skill A & one other) in random order. The teacher gave quantitative & qualitative feedback. Students were given a questionnaire to rate the utility of the self-videos. All students found reflection on the video helpful for identifying areas for improvement in mannerisms & communication. They also found teacher comments & comparing videos with peers helpful. Sequential viewing of simulation videos allowed the students to reflect and monitor their progress. The intervention group had sig. higher scores on skill A than the control group.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Mamede,S., van Gog, T., Moura, A. S., de Faria, R. M. D., Peixoto, J. M., Rikers, R. M. J. P., et al. (2012), 2+	Quasi-experimental with 3 intervention groups with post-test immediately after and again 1 week later diagnosing four different cases. Random assignment.	46 fourth year medical student volunteers diagnosed six clinical cases as part of learning experience.	Inter-rater reliability of 92%.	Only 46/120 volunteered	2 evaluators of answers to cases. Blinding	Initially, test scores in the reflection group were sig. lower than the other 2 groups. 1 week later, the test scores in the reflection group were sig. higher than the other 2 groups. Scores in the reflection group sig. improved between testing, but scores fell in the other 2 groups, sig. in the immediate diagnosis group. Previous clinical exposure to the conditions did not differ among the groups.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Mariani, B., Cantrell, M. A., Meakim, C., Prieto, P., & Dreifuerst, K. T. (2013) 2+	Mixed method quasi-experimental with control group, random assignment to clinical groups	Convenience sample of 86/90 junior nursing students. Very homogenous sample mean age 20.5 years.	LCJR is a valid and reliable tool. Inter-rater reliability was high (r =.92; p<.01).	Self-selection in focus groups. LCJR scores completed by faculty member for first simulation and by researcher for second.	Blinding attempted in LCJR rating. Neg. Report	Intervention was the DML. Researchers completed LCJR after both simulations, faculty member after 1st. All students received DML after 2nd simulation. Audio-taped 2 Focus group interviews contained 7 volunteers and were transcribed and coded for themes. No sig. diff. in LCJR scores. DML was seen as: improving student learning, being learner focused, a holistic approach, and promoting figuring out problems and helping students make connections. The standard debriefing was seen as: more instructor focused, concentrating on right vs. wrong, not giving the whole picture, and not as helpful for learning. History and maturation effect since students were in clinical 4-5 weeks between simulations.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
McGinty, S. M. Y. (2001) 3	Cohort study	Convenience sample of 27/30 second year physical therapy students. Journals were a requirement of the course, but 3 students choose not to be a part of the study.	Inter-rater reliability .72 for both reflective thinking levels, not established for critical thinking skills. Intra-rater reliability 87% for reflection and 83% for critical thinking skills. Clinical Performance Instrument is a valid and reliable instrument.	4 students failed to turn in all 9 journals. Possible ceiling effect: 100% of students reached 4 of the levels of reflection, and 89% reached the other 2 categories; 100% had 3 of the critical thinking skills, 96%, 89%, and 93% for others.	Blinding. 2 coders. Triangulation of data by student interviews, journals, and clinical performance instrument comments. Student quotes.	Clinical Performance Instrument instructor narrative comments had a 72% agreement for levels of reflection, and 80% for critical thinking. Levels of reflection and critical thinking had an r =.87 that was statistically sig. Students had kept unstructured, ungraded reflective journals during program. Were given specific guidelines to write weekly entries focusing on reflection before, during, and after action. 1 on 1 interviews with 5 randomly selected students. Evaluated for 6 levels of reflection: Descriptive, Affective, Evaluative, Value Judgment, Conceptual, and Theoretical Reflectivity. Evaluated for 6 critical thinking skills.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
McMahon, G. T., Monaghan, C., Falchuk, K., Gordon, J. A., & Alexander, E. K. (2005) 3	Descriptive study	Convenience sample of 90 third year medical students	N/A	No description of sample.	Previous pilot of learning module and protocol.	Module imbedded a group reflective session after 3 cases & 2 teaching sessions. During reflective session students were able to integrate all sources of knowledge. 72% of the students felt the reflective analysis was the most critical component of the module. Learning goals established up front. Use of multiple cases showed students progressing in their ability to care for the simulated pt. Instructors observed students reflecting after action, reviewing case details, finding errors, & identifying solutions. In the reflection session, students were able to formulate the underlying general principles, & compare & contrast the cases. Student comments included: "very supportive environment - tolerant of mistakes & therefore conducive to learning" (p. 88). Trained faculty were needed to conduct the reflective session. A group size of 3 was optimal for reflective discussion.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Murphy, J. I. (2004) 2-	Quasi-experimental post-test only with control group	33 Nursing student volunteers from four different cohorts of first semester students, random assignment	Internal consistency was acceptable (Cronbach's alpha (0.90)	Self-selection bias, but volunteers were compared to volunteers. Researcher developed instruments not fully described.	Contamination prevented by having groups at different campuses. Student quotes included.	Intervention students and instructors received training and reinforcement on focused reflection and articulation to connect theory to practice. Assessment and Analysis Instrument, based on Gordon's functional patterns, used to rate student write ups of pts during weeks 7 & 15. Clinical reasoning ability was defined as the number of correct items on test plus the instrument score. Interviews of 6 high and 6 low scorers on clinical reasoning measures. No difference in clinical reasoning score. Sig. dif. in Assessment and Analysis Instrument scores. The 6 students with the highest clinical scores were in the intervention group but so were the 2 students with the lowest scores. Students with a high clinical reasoning score: had a more positive attitude toward reflection, were intrinsically motivated and enthusiastic, described clinical events more fully, and connected reflective writing with learning.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Padden, M. L. (2011) 2-	Quasi-experimental, pre-test, post-test design with control group.	Convenience sample of 112/157 ADN student volunteers enrolled in third clinical course over 14 weeks at 4 different schools. Intervention group at one school.	Inter-rater reliability at .80. Content validity of Level of Reflection on Action Assessment. The Self Reflection and Insight Scale is a valid and reliable tool, for this study (.87 pre-test & .91 post-test). Clinical Decision Making skills in Nursing Scale is valid and reliable (.72,.79)	Maturation, Instrumentation (researcher's skill in rating journals may have improved). Self-selection. No random selection or assignment. 33/60 (55%) students in control group completed compared to 79/93 (85%) # needed to meet power analysis of intervention group not reached (33/51).	10% of journals Randomly selected & rated by a 2nd rater. 3rd rater was to be used if agreement could not be reached but was not needed. All students asked to participate volunteered.	Intervention was instruction on and researcher guided reflective journaling, and students were given The Guide to Reflection. The researcher provided feedback, suggestions, and strategies for improvement. The Level of Reflection on Action Assessment was used to rate reflection is 1 of 6 levels. The intervention did not have a sig. effect on level of reflection, self-awareness, or perceived clinical decision making skills. There was a sig. positive relationship between level of reflection and self-awareness and a sig. neg. relationship between self-awareness and clinical decision making skills, age, and hours worked. Self Reflection and Insight Scale, and Clinical Decision Making in Nursing Scale were used as pre and post-test measures. Students were to post their journal entries online but due to technical difficulties some chose to turn in print copies.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Perera, J., Mohamadou, G., & Kaur, S. (2010) 2-	Quasi-experimental with control group	Convenience sample of 202 first year medical students; only 190 completed class.	Face validity of question-naire.	No blinding. No questionnaire given to control group about their simulation experience. Confidentiality may not have been maintained about the intervention	Students divided based on pre-admission scores. No sig. diff. in gender, or age distribution.	Experimental groups trained to give feedback to peers & evaluate performance. Standardized pts gave feedback to students. Self-assessment tool used to guide reflection & identify performance gaps. Peers gave feedback on uncovered gaps using reflection guide. Facilitators addressed any other uncovered gaps. Interview skills learning sessions conducted by the pt. Control group had only feedback from the pt & facilitator. Interview skills assessed at the end of semester by a 3 station OSCE with experimental subjects mixed with controls. Sig. diff. in OSCE total score, interview style, listening, & building rapport. No sig. diff. in language or interview structure. 88.7% of experimental group completed questionnaire. Less than half had formally self or peer assessed. 70% thought they identified gaps in pt feedback. 90.4% used self & peer evaluation during practice sessions. Areas needing improvement were: interview style, addressing pt concerns, empathy, pt understanding, non-verbal communication, & paraphrasing. 86.4% of the students felt the intervention was a positive process & developed skills needed for team learning.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Plack, M. M., Driscoll, M., Blissett, S., McKenna, R., & Plack, T. P. (2005) 3	Descriptive cohort study	Convenience sample of 27 physical therapy students who submitted a total of 48 journals	Inter-rater reliability ranged from 65.1% to 93.0% for the 9 elements and from 67.4% to 85.7% for the 3 types of reflective writing ability (γ=0.88 to 0.98, ICC of 0.74).	Lack of variability in writing samples, led to some low Φ and ICC values for the 9 elements. 3rd rater was not as theoretically accurate as the other 2 raters in the areas of "returns to experience" and "attends to feelings" which led to low inter-rater reliability and no Φ value. Further refinement of these definitions needed.	3 coders. 5 journal entries were used to refine coding.	Reflective elements coded in 1 of 9 categories: reflection during action, after action, before action, content (uses different perspectives), process, premise (identifies assumptions), returns to experience, attends to feelings, or reevaluates by comparing to past experiences. Axis I was time dependent: reflection during action, reflection after action, and reflection before action. Axis II was content dependent: content, process, and premise. Axis III was stage dependent: returns to experience, attends to feelings, and reevaluates. Writing samples were then classified as either non-reflective (14.7%), reflective (43.4%), or critically reflective (41.9%). The non-reflective writing sample simply describes the experiences, and rejects learning from new experience. Premise reflection was typically a characteristic of critical reflection.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Plack, M. M., Driscoll, M., Marquez, M., Cuppernull, L., Maring, J., & Greenberg, L. (2007) 3	Descriptive study	Convenience sample of 21 third year med student volunteers during their pediatric clerkship submitted 308 journal entries	Inter-rater reliability ranged from 78.2% to 100% with a kappa statistic of 0.57.	81 students in class and 21 volunteered; self-selection bias.	All 3 coders rated each entry.	3 levels of reflective writing were compared to a modified Bloom's Taxonomy definition. 5 unrelated writing samples were used to refine coding. 93.5% of the entries contained level I elements, 68.9% contained level II elements, and 48.4% contained level III elements. Level III writing elements received the lowest inter-rater agreement, indicating a need for further refinement of the definition.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Plack, M. M., Dunfee, H., Rindflesch, A., & Driscoll, M. (2008) 2-	Mixed method Case Control	Convenience sample of 7 physical therapy students completing their final clinical internships.	Inter-rater reliability was 87% with a kappa statistic of 0.82 for coding of the reflective essays.	Small sample size. Groups were from different semesters.	3 researchers coded discussion board data. Two researchers analyzed & coded essay data. Triangulation of data.	Web based discussion board to record participant comments. Students received instruction on reflective practice and a set of reflective questions to use. The experimental group, had a faculty facilitator, received a 30 min. introduction to action learning. Students presented & discussed critical incidents. After the discussion, each student wrote a reflective paper. Comments were evaluated as containing reflection during action (5.2%), after action (92.4%), or before action (29.6%) and noted for data gathering (93.5%), data analysis (83.2%), and conclusion drawing (62.9%). No sig. differences were found between groups on reflection during, after, or before action. The experimental group had more entries that contained data gathering. Essays contained 3 themes: collaborative learning was enhanced; and reflective practice is a conscious, active, analytical method; and facilitates broader and deeper thinking that offers insight into clinical problems.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Schwartz, B., & Bohay, R. (2012) 3	Cohort study with pre and post intervention surveys	Convenience sample of 224 pre-doctoral and 24 certification students	Jefferson Scale of Empathy for students a validated and reliable instrument.	Less than half completed pre-intervention survey. 59.3% of the 2nd years and 79.7% of 3rd years completed post survey. No control group	Easy to administer	Watched 11 videos of pts talking about dental experiences. 2nd year pre-clinical students wrote a 1,200 word reflective essay. 3rd year clinical students wrote 1,000 words. One month later, the 2nd & 3rd year students were asked to rate the intervention. Students completed a 20 question empathy survey. Reflective essay was thought to significantly raise empathy for pt by 71.9% of 2nd year & 43.7% of 3rd year students. Students commented that writing turned a passive experience into an active one; and forced reflection; but that maybe a discussion would have been better. 3rd year scores on empathy were sig. lower than 2nd year scores which was a normal finding. 100% of 2nd & 95% of 3rd year thought the video time was just right or could be increased. 97% of 2nd year & 82% of 3rd year students thought the intervention improved their educational experience. 100% of the 2nd year & 91% of 3rd year thought the videos made the learning more memorable. 84% of the 2nd, & 67% of the 3rd year said it made them more committed to being a professional.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Thompson, B. M., Teal, C. R., Scott, S. M., Manning, S. N., Greenfield, E., Shada, R., et al. (2010) 2-	Cohort study with pre-test & post-test on attitude and confidence concerning pt contextual clues.	171 First year medical students (166 had complete data sets)	N/A	So many variables within between groups introduced confounding	Many different opportunities for the students to learn the process.	An online identification and reflection activity was used to prepare students, along with a video vignette based large group activity. Individually students preformed a history on a standardized pt, reviewed the video tape, and completed a reflective assignment. In facilitated small groups, students show a snippet of the video and had a discussion. Students highly rated the facilitator and felt the small group discussion was effective. Overall, students found the activities effective in promoting reflection. The only variable with a sig. change was students' confidence in their ability to effectively identify pt contextual concerns. Facilitators felt the students had been poorly prepared and this was confusing for the students.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Tofil, N. M., Benner, K. W., Worthington, M. A., Zinkan, L., & Lee White, M. (2010) 2-	Mixed Methods Quasi-experimental pre-test post-test without control group	Convenience sample of 42/45 pharmacy students over 2 years enrolled in course.	Not calculated.	Self-selection d/t elective course; no control group. History. Maturation	Sig. change in small sample.	Content covered 1st, case based learning, and 2 simulations. Sig. change in exam score from pre-test to post-test. Application knowledge improved the most from a Bloom's taxonomy perspective. 95% of students improved scores. Students liked reflecting on the experience & instructors believed students benefitted from reflecting. Realism of the simulation felt to allow students to suspend belief. During 2nd year pre-briefing included an introduction on what to expect & how to do things in simulation. Pre-briefing was added in response to student concerns.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Tsang, A. K. L. (2012) 3	Cohort study	17 Senior Bachelor of oral health students enrolled in a year-long course	N/A	Small sample size	1 year study	Trained in 2 seminars about reflective learning and writing. Students were given a guide to reflection. Reflective journal entries discussed critical incidents that happened. Faculty feedback was given in emails and one-to-one interviews. Journals were grade as pass/fail. Students' ability to reflect improved. A software program was used to evaluate the reflective writings. Students' first entries were mostly descriptive (revisit & react; two lowest categories), but by week 12 the entries were 35% relational and 15% respond. By week 12 of the second semester, 32% of the entries were relational, and 26% were responsive. Students who sought feedback and guidance on their reflective writings tended to have higher levels of reflection in their journal entries. Reflection is a learned experience. To have transformational reflection, students must have experiences that are out of their comfort zone.

Author Ev Lev	Methodology	Sample & Setting	Validity & Reliability	Limitations	Strengths	Synopsis
Wald, H. S., Borkan, J. M., Taylor, J. S., Anthony, D., & Reis, S. P. (2012) 2+	Systematic review with bibliography search to create rubric. Iterative development of rubric with successive trials. Random selection of narratives	Reviewed PubMed for articles written from 1995-2008. 5 samples of medical students. 5 iterations: first 4 samples were 10 narratives apiece and last was 60.	Final ICC was 0.632 and Cronbach's alpha was 0.774	ICC may have improved over iterations d/t researcher training, and increasing familiarity with rubric.	Three raters	4 types of reflection assessment found: scales, thematic coding, qualitative analysis for model formation, analytical instructional rubrics. Formative analytical instructional rubrics were found to be the best for the faculty's assessment of reflective levels. Process for rubric starts with reading the entire narrative, zooming in to find criteria, zooming out to decide what level the writing sample represents, and listing quotes that support the level assigned. The rubric criteria are: writing spectrum, presence of writer, descriptive level, attention to emotions, analysis, answers the assignment question. Critically reflective writing is also classified as either transformative or confirmatory (p. 48).

Table D.2 Qualitative Evidence

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Becherer, V. H. (2011) 3	To identify student perceptions of facilitator led group reflective review of material and subsequent reflective journaling before tests.	Reflection assisted students in: learning, developing emotional intelligence & professional practice, helped them recognize what they did & didn't know, prompted them to study sooner, think at a deeper level, consider perspectives, & deciding whether information was germane. Time needed to consider the problem & think about context. Solving problems was made easier by being asked questions, discussion, & thinking out loud.	Grounded Theory	Purposive sampling of 65 nursing students in 2 sections of a Child and Family Course. 45 students made journal entries. 7 students were in a focus group interview. 3 students had one - on-one interviews.	A learning activity then reflective thinking reviews held prior to each of the 5 tests. Then a reflective journal entry. 10 Students who participated in all reviews were interviewed in a focus group. Survey when course was over.	Possible researcher bias, subjectivity of information No comparison of students who participated in reflective exercises with students who did not.	Random selection, Triangula-tion of data, blinding of journal data	Grounded theory used to verify that reflective thinking is key to the learning process, emotional intelligence, and professional development of student nurses.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Boyd, L. D. (2002) 3	To explore the development of critical thinking through reflection	In the process of connecting lecture material to clinical experiences students progressing from: questioning what it is they see, to looking at things in a new way, to recognizing the need to care for the patient. Considerable affective component to the reflections.	Thematic analysis	Convenienc e sample of the cohort of 69 first year dental students. 3 Interview and clinical observation subjects chosen from 10 volunteers.	Data collected via: reflection papers, audiotaped semi-structured interviews, and clinical observation. Guidelines given for reflective paper. Field notes taken during observation.	Non-random selection of students interviewed Saturation not achieved. Portion of study reviewed here small part of larger pilot study and not the focus of the paper.	Representa-tive sample. Student quotes. Triangula-tion of sources.	Compre-hensive data collection. Thematic analysis appropriate for the identification of feelings, beliefs, attitudes, and values.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Chou, C. L., Johnston, C. B., Singh, B., Garber, J. D., Kaplan, E., Lee, K., Teherani, A. (2011) 3	To describe student perceptions of a peer support group in the VALOR program	Students felt the peer group experience was best part of program. Enjoyed working with the same group through 3 rotations. Felt this led to a supportive environment, facilitated reflection, & communication. Students felt that the group enhanced sharing, caring, & peer assistance. Peer groups were a "safe place" where emotional venting was permitted. Long-term impact of the program was that students built relationships, & learned skills for team building. Many students found it useful for reflecting on and the processing of stressful experiences.	None identified	42 medical students in a voluntary 6 month long program in peer groups of 6 students	Post experience surveys immediately, at 5, and 27 months.	Students self-selected into program.	2 coders, data was identical and so was aggregated. Long term follow-up.	No methodology specified. No mention of how themes were identified and organized.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Croke, E. (2004) 3	To find out if the process of reflection after action would improve the clinical decision making abilities of nursing students	Students wrote about how they used critical thinking skills and what critical thinking dispositions to make clinical decisions. Initial improvements were reported in assessment, diagnosis, and evaluation. Later on, students noted progress in planning and implementing a plan of care. Practice was felt to be key the students' progress.	Participatory action research	34 first semester nursing students	Extensive guidelines and instructions on reflective journal writing were given to students. Feedback and clarification provided by instructor. 10 weekly journals took 1 hour to write apiece.	No structure in themes. 1 coder, who was also teacher, researcher. Did not describe how process would change as a result of this study.	Student quotes included.	Application of participatory action research to students' learning process; as students explore their approach to old solutions they become better at solving future problems.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Decker, S. (2007) 3	Thoughtful practice combines critical and reflective thinking. Can simulation be used as a tool to enhance both?	Groups were either in task oriented (21.4%), situation specific (39.3%) or critical thinking stage (39.3%). Reflective thinking was divided into levels: Non-Reflectors, Reflectors, & Critical reflectors. Types of reflection: during action & conscious review to discover new understandings with the intent of applying the new knowledge to practice. Ability of the facilitator to support students' reflections assists them in reflecting after action. Reflective and critical thinking positively correlated. Level of reflective thinking of student affected their ability to successfully complete the scenario. Socratic questioning and cues used.	Grounded Theory – Mixed method	Purposeful sampling of 114/154 seniors who were exposed to a previous pilot study.	Demographic survey. Self-selected groups of 4-5 then had observation during 20 minute simulation exercise and followed by 20 min. group interview. Responses were coded and assigned to categories and sub-categories.	Only one school of nursing used. Possible self-selection bias on the part of students selected to participate in faculty's research. One coder.	Taped interviews. One-way windows used for observation. Eight student volunteers checked the merged data and agreed that it was on the mark.	Appropriate use of grounded theory to verify theory of thoughtful practice and investigate whether simulation can assist thoughtful practice.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Donovan, M. O. (2007) 3	To find out nursing students perceptions of the reflective process	Subcategories of understanding the process of reflection: looking back & thinking about what happened; tearing the experience apart, sitting down & thinking; discussion as reflection; sharing experiences; improving practice. Subcategories of using reflection: developing self-awareness; affective component, becoming aware of limitation; climate of trust, & client care focus. Subcategories of needing support & guidance: guidance needed to learn to reflect, need to start early in program since reflective ability improves over time; assessment possible barrier, needed time to reflect, preceptor is key to process.	Grounded theory	5 third year diploma nursing students	Interviewed one-on-one	Small sample size, saturation not achieved in all categories, one coder	Participants given pseudonyms, constant comparative method of data analysis	In-depth interview process but all themes probably not revealed. Appropriate use of grounded theory to construct students' understanding of the reflective process.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Duggan, A., Bradshaw, Y. S., Carroll, S. E., Rattigan, S. H., & Altman, W. (2009) 3	To identify areas of learning, and reflection during debriefing	3 areas for potential learning were identified: how a disability affects the treatment plan, using reflection to identify attitudes about people with disability, & the practice of medicine. 23 different categories of student learning during the debriefing were identified. Students were able to reflect after action & to articulate strategies for overcoming difficulties in interviewing the pt with a disability. Students appreciated feedback that identified areas of strength & gaps in performance. The need for balance between the patient's desires & the practitioner's expertise was recognized.	None specified.	Convenienc e sample of 138 3rd and 4th year med student volunteers in successive classes. 16 students were excluded due to poor tape quality.	Videotaped exam with standardized pt and debriefing with feedback from the facilitator, pt, and a peer.15 transcripts of debriefing were reviewed by researchers to identify areas of student learning. 12 practice transcripts were used to refine coding. Kappa of .89 was achieved on practice coding.	Non-verbal communica-tion not analyzed. 16 students' transcripts were not analyzed due to technical difficulties.	154 out of students 186 asked participated in videotaping Researchers were blinded as to who gave consent. 2 coders consensus was achieved.	No methodology specified. All data was reanalyzed after all subthemes had been created.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Ekebergh, M. (2007) 3	To find out how the weaving of the students' life-world, and theory and practice knowledge affect the learning process.	Reflection on and analysis of the pt moves understanding from piecemeal to holistic. Learning requires an open approach to the students' understanding of the world. Students' learning needs should have priority. Supervisors feel that: course was useful for teaching how to conduct this method, mutual respect is required; openness leads to co-operation; must undertake reflection also; meet students where they are; joyful environment that promotes interest in students learning; and they must remain in student experience.	Phenomen-ological epistemeology	25 nursing students, 8 of their clinical instructors, 8 nurses who worked with the students	5 focus group interviews with clinical groups on last day of clinical; separate group interviews of teachers and nurses. Selection of reflective individuals: 8 student, 2 instructors, and 2 nurses for individual interviews 2 weeks later	All meaning felt to have been contextual, which limits generaliz-ability. No mention of saturation.	Interviews tape - recorded and transcribed.	Phenomen-ological a good fit for uncovering the precursors to a good reflective learning experience.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Gwozdek, A. E., Klausner, C. P., & Kerschbaum, W. E. (2009) 3	To report on the content of online student journal entries as a reflection and sharing strategy.	29% of the journal entries related didactic material to clinical, and 32% mentioned student collaboration. 77% of the students agreed that the reflection journaling was helpful. 87% found reading other students' post helpful, and 58% found commenting helpful. A sense of community was developed through the sharing of entries. They found it allowed them to individualize their learning, but was time consuming. Students preferred online to in person discussion because they could spend time on content they needed.	None specified	28 first semester dental hygiene students	Online directed reflective journaling for 8 weeks. Students wrote 6 entries and commented on 2 peer postings.	Saturation not reached.	2 coders, student quotes included.	No methodology specified. No structure to the categorization of themes.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Honey, M., Waterworth, S., Baker, H., & Lenzie-Smith, K. (2006) 3	To evaluate the usefulness of formal reflection in undergraduat e nursing disability module	Coping with clinical practice subthemes were: fear and anxiety, feeling alone, feeling unprepared, and coping strategies. Coping strategies identified by students were: setting boundaries, reflecting on previous knowledge and experience, and seeking understanding through knowledge. Students' reflections focused more on overall learning and clinical practice than the disability placement. Clearer guidelines were felt to be needed. Researchers felt that students needed an opportunity to reflect before action prior to beginning. Students identified gaps in their knowledge, and took steps to bridge that gap.	Qualitative approach	Convenienc e sample of 12 second year nursing student volunteers who had been enrolled in the Nursing in Mental Health and Disability course the previous year	Guide provided to students, to assist them in writing a 1,000 word paper. 12 reflective assignments were analyzed	Small sample size, self-selection bias	Anonymous submission, 2 coders with separate reviewer	No specified methodology

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Kuo, C. L., Turton, M., Cheng, S., & Lee, H. (2011) 3	To explore the experience of a clinical caring journal by students and instructors.	Six themes: journal guided caring behavior; enabling students' reflective caring abilities; provides a sense of accomplishment and self-awareness; increasing and deepening interactions between student and instructors; improving the students' learning and self-development, and improved writing skills. Students felt that the journal should be used throughout the program. Students wanted more guidelines and examples.	Qualitative. Constant comparative method to create categories and generate themes.	16/880 senior students and 7/90 clinical instructors volunteered for focus group.	Students wrote 2 reflective entries for 8 rotation month long rotation. Instructors provided written feedback. Audio taped & transcribed semi-structured focus group interview.	Small sample size. Self-selection bias. Saturation not reached.	Student quotes included. Multiple coders.	No methodology specified

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Lähteenmäaki, M. (2005) 3	To discover how learning takes place in physiotherapy clinicals	Traditional method of teaching where student is shown how to do a skill and then replicates the skill was seen as an obstacle to thinking. Reviewing clinical sessions helped them to reflectively think. Observational experiences in clinical assisted students in attending to details of the procedure & identifying areas for future learning. More experienced students valued clinical questioning. Writing reports seemed to make the experience clearer to some. Negative emotions got in the way of learning and positive ones helped the students to focus. Writing out plans for future pts was seen as burdensome.	Ethno-method-ology	Convenience sample of 32 physio-therapy students in 5 groups; 4 students lost for various reasons	5 Group discussions over 2.5 years; video tape and field notes used to help students recall events that happened in clinical. Discussions were tape and video recorded and transcribed. Researcher moderator seen as an obstacle to process at first. Students anxious to learn from one another.	Not generaliz-able. Only one cohort of students at one school	Student quotes included	Ethnography was useful in finding out how the learning needs of students changed over the course of their education.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Lindgren, B., & Athlin, E. (2010) 3	To describe the value of clinical group supervision.	Satisfaction with being in a group sub-themes were: sharing and recognition; and support and challenges. Personal and professional development was the other main categories with sub-categories. Becoming aware of feelings, attitudes, strengths, and weaknesses. Understanding of others, ethics, and cultural issues. Preparing for coming events: new situations, encounters with pt and family, and being a nurse in the future. Gaining strength: being honest and plain, and taking risks. Being inspired in further learning: searching for knowledge, and asking for judgment.	Qualitative descriptive	8 nurse instructors who led clinical supervision groups for 8-9 meetings over the semester	Instructors took field notes during each session as to what the students had gained from the session.	Possible recall bias on part of instructors and the students.	2 coders	No specific methodology used.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Lutz, G., Scheffer, C., Edelhaeuser, F., Tauschel, D., & Neumann, M. (2013) 3	To gain an understand-ing of how reflection training is perceived by students	Students liked: having a trained and supportive facilitator, a safe place to talk, a supportive group, and focusing on real clinical problems. Students felt that reflective training: reduced stress, improved quality of pt care, helped them deal with adversity, improved the learning process, helped them identify stressors, and enhanced personal and professional development. Students recommended: more reflection training throughout the program, individual coaching, use of a neutral facilitator, and more direct feedback.	Develop-mental evaluation	18/30 fourth year medical student volunteers	Students taught about reflective practice. 90 min. reflection training group every 2 weeks. Audio-taped & transcribed semi-structured individual interviews.	Data saturation. Selection bias. Researcher conducted the reflection training sessions.	2 coders and software, with a 3 researcher acting as reviewer. Student quotes included.	Developmenta l evaluation technique does not seem to be an appropriate choice of methodology, since interviews were conducted after training was completed by an author not involved in the training. Interview questions and follow-ups are appropriate for eliciting data.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Manning, A., Cronin, P., Monaghan, A., & Rawlings-Anderson, K. (2009) 3	To discover the utility of optional reflective groups connected to a clinical	7 major categories with sub themes: Needs (settling in, unmet reflection needs, sharing, expectations, time, changing priority, and differing objectives); Confidentiality (process, fear of disclosure, free to disclose, disclosing); Facilitator (skills, supportive environment), Group Processes (content of reflection, sharing, being together, interconnectedness); Value of Sessions (time out, deal with being a student, relating); Perceived value (resource, coping, learning, sharing, developing) Outcomes (altered perspectives, options, interpersonal skills, feeling valued, application, support).	Phenomen-onological	Purposive sample of 2 cohorts, first and third year, of nursing students	Audiotaped focus group interviews, transcribed verbatim. Follow up focus group based on the transcripts from first interview.	Small sample size. No number of students given, just 4 small focus groups	Students quotes included, coding done by 2 researchers in stages, first separating the data from the different years and then combining	Questionable combining of data from 2 different groups that had different outlooks on reflection. Phenomenon may not be the same for 1st and 3rd year students who are at different stages in their learning.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Nishigori, H., Otani, T., Plint, S., Uchino, M., & Ban, N. (2009) 3	To classify what students learned from inter-national electives.	9 learning outcomes were identified. Students were found to learn about most items and especially professional issues by reflecting on how practice was different between the 2 countries. Reflection was identified as the most important process affecting the learning that took place during the exchange.	Semi-structured individual interviews were analyzed by the thematic synthesis method.	Convenienc e sample of 6 British and 15 Japanese medical students who participated in an international exchange.	Tape-recorded and transcribed immediately	British students were interviewed 10 months after experience, 1 Japanese student not interviewed	1st author reviewed all transcripts. 2nd author reviewed Japanese transcripts. 3rd author reviewed British transcripts. Triangulatio n of themes.	Appropriate use of thematic synthesis. Text was 1st coded, organized by descriptive themes and then analytical themes were developed.
O'Donovan, M. (2006) 3	To explore perceptions of reflection as a learning strategy during clinical placement.	Sub-themes for needing support and guidance in reflective process: faculty have key roles, additional time, and more preparation, guidance, and support needed. Need to introduce reflection training early in curriculum. Reflective journals should be required.	Grounded theory constructivis t approach. Constant comparative method of data analysis.	Purposive sample of 5 third year diploma nursing students	Audio-taped interview and transcribed. Field notes taken. Initial categories verified by 2 participants.	Small sample size, saturation of data not reached. One coder	Student quotes included. Triangulatio n of methods and sources.	Grounded theory appropriately used to identify themes that influence students' use of reflection in clinical setting.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Pee, B., Woodman, T., Fry, H., & Davenport, E. (2002) 3	To describe student reflective entries at each level and compare peer ratings with 2 different researcher methods.	Students were reflecting at different levels and had evidence of different aspects of reflection in their entries. Explicit questions in the tool were more frequently addressed. Questions that are asking for descriptions or the students' perspective are more likely to be addressed. Questions that are analytical in nature are less frequently addressed. Improvements considered were asking for: reasons, factors influencing events, and pts' and students' feelings. Peers' ratings were consistent with researcher ratings.	Mixed methods. Qualitative and case study.	14/26 dental therapy student volunteers wrote entries. 20 students from other schools were peer judges. 18/26 returned survey of tool.	Guideline developed to assist students in writing reflective entries on critical incidents. Students rated peers' worksheets for evidence of reflection. Researchers used established criteria to evaluate writing samples. Students completed survey.	Turning in reflective entries was voluntary, selection bias.	Protocols for inter-rater agreement using 2 different methods. Student example and quotes included. Both methods had acceptable (.74 & .86) inter-rater agreement.	No methodology specified.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Rowe, M. (2012) 3	To determine if an online social network could be used to reveal students' understanding of clinical practice issues	Modeling of desired, Contingency management, Providing feedback to students, Teaching the learning and reflective processes, Stimulate thoughtful responses, Create the framework for cognitive development	Assisted performance through the zone of proximal development reported on qualitatively	Convenience sample of 70 third and fourth year physio-therapy students	Facilitated blog assignments linked to module outcomes. Seniors wrote on clinical experiences, juniors wrote on ethical dilemmas experience during clinical. Students were to read, comment, add links, and media to each other's' posts.	Contingency management not well connected to quotes. No definition of what this terms means. No saturation of data.	2 coders using pre-determined themes according to the Theory of Assisted Performance, student quotes included	No methodology named. Fitting data into a pre-selected framework, rather than letting the data be organized into its own logical structure.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Silvia, B., Valerio, D., & Lorenza, G. (2013) 3	To describe the level of reflection that student journal entries attain in a 15 day period and their perception of the experience.	Themes related to journal writing were: uneasiness about someone reading their writing, anonymity would be preferable; evaluation should not be based on journal entries; helpfulness of journal writing. 459 reflective levels were assigned to portions of text. The majority of the ratings were Level 1 Descriptively reflective (51.63%). Only 4.36% of the ratings were Level 7, Theoretical reflectivity. Level 2, Affective reflectivity (17.43%), Level 3, Discriminant (20.94%), Level 4, Judgmental (1.96%), Level 5, Conceptual (4.36%), Level 6, Psychic (0.22%).	Qualitatively based on Mezirow's 7 levels of reflectivity	12/13 2nd and 3rd year student volunteers' journals. 1 student only drew in journal. Focus group was 6/13.	Students wrote in a reflective journal for 15 days. Journal entries were analyzed. Tape-recorded & transcribed semi-structured focus group interview.	Small sample size, saturation, selection bias.	2 coders, Student quotes included.	No specific methodology named. Rating scale is nominal not ordinal in nature.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Skovsgaard, A. (2004) 3	To describe the use of dialogue and reflection between students and their clinical instructors.	Most dialogues and reflections focus on tasks and/or how to share the responsibility for tasks. Students believe they learn in 3 steps: observing the instructor do the task and provide explanations, practice tasks with instructor evaluating, and dialogue and reflect with instructor. Dialogue and reflection is at odds with the need to perform tasks, consuming both time and attention. Students tend not to initiate dialogue or reflection with their instructors. The conscious use of dialogue to develop knowledge and reflection to problem solve is not commonly used by clinical instructors.	None described	4 first year student nurses and their clinical instructors on 4 different units.	Observation, field notes and tape-recorded semi-structured interviews with students and their instructors.	No quotes. No data saturation.	Ties together what instructors and students do with how students believe they learn.	No guiding methodology.

Author Ev Lev	Aims	Sub-Concept Analysis & Findings	Method	Sample/ Setting	Procedure	Limitations	Strengths	Rigor
Williams, R. M., Wessel, J., Gemus, M., & Foster-Seargeant, E. (2002) 3	To describe perceptions of clinical learning and to promote reflective thinking	Reflective themes were: process of making clinical decisions; complexity and richness of interactions with pts; effects of clinical environment on learning and pt care; acquisition of skills; value of clinical experiences in integrating & adapting theory; different learning methods. 22 students achieved the highest level (reflection before action), 20 the next (gains a new understanding), 13 the next (verifies learning), 1 the next (analyzes learning), and all students moved beyond the lowest level (describes learning).	Mixed methods. Qualitative and Case Study	56 physical therapy students all with previous baccalaureate degrees.	Used a 5 level reflective thinking rubric, reliability .68. 10 randomly selected journals used to establish coding and themes.	Very high levels of reflection may be d/t instructions given defining the highest level of reflection as application to future practice.	2/4 coders graded each entry. Extensive student quotes.	Mixture of methods without thorough description of criteria for either type.

DASH - SV = Debriefing Assessment for Simulation in Healthcare-Student Version	DML = Debriefing for Meaningful Learning	HSRT = Health Sciences Reasoning Test
LCJR = Lasater Clinical Judgment Rubric	OPT model = Outcome Present state Test model	OSCE= Objective Structured Clinical Examination

APPENDIX E

Scottish Intercollegiate Guidelines Network

Grades of Recommendations

A At least one meta-analysis, systematic review, or RCT rated as 1++, and directly applicable to the target population; *or*

A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results

B A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; *or*

Extrapolated evidence from studies rated as 1++ or 1+

C A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; *or*

Extrapolated evidence from studies rated as 2++

D Evidence level 3 or 4; *or*

Extrapolated evidence from studies rated as 2+

Good practice points

√ Recommended best practice based on the clinical experience of the guideline development group

Note. Adapted from " SIGN 50: A guideline developer's handbook," by SIGN, 2011.