

Concepts of Evidence Used in Science-Based Occupations:

Acute-Care Nursing

RESEARCH REPORT

Glen S. Aikenhead

College of Education

University of Saskatchewan

28 Campus Drive

Saskatoon, SK, S7N 0X1

glen.aikenhead@usask.ca

October 28, 2003

A research project funded by the University of Saskatchewan President's SSHRC Research Fund.

Table of Contents

Background to the Study

Rationale and Purpose of the Study

Research Methodology

 Research Question

 Participants

 Procedures

 Limitations

Research Context

 Purpose of the Setting

 Accountability in the Setting

 Knowledge-in-Use Enacted in the Setting

 Evidence

Evaluating Data

 Becoming Evidence

 Taking It to the Next Level

 Initiating an Intervention

 Summary

Concepts of Evidence

 Reliability

 Accuracy

 Other Concepts of Evidence

Scientific Knowledge-in-Use

Summary

Acknowledgement

References

Figures

Background to the Study

Studies have shown a poor match between the scientific content generally taught in high school science courses and the type of scientific understanding required for success in science-based occupations in which knowledge of the practice of science and technology is either *critical* to the job or *enhances* occupational competence (Chin, Munby, Hutchinson, Taylor & Clark, in press; Coles, 1997; Lottero-Perdue & Brickhouse, 2002). Duggan and Gott (2002) investigated in some detail the role of science for employees in five science-based industries. They discovered that most of the scientific conceptual understanding used by employees is learned on the job (not in high school or university courses), but that “a secure knowledge of procedural understanding appeared to be critical” (p. 674). *Procedural understanding*, the thinking behind the doing of science, draws upon a wealth of ideas about evidence itself, for example, concepts of validity and reliability of evidence. Duggan and Gott called this cluster of ideas “*concepts of evidence.*” Thus, procedural understanding is informed by, in part, concepts of evidence.

Interestingly, Duggan and Gott (2002) also discovered that this core procedural understanding was also critical to a non-science public who were involved with a science-related social issue in an advocacy role. Duggan and Gott’s tentative recommendations for educators suggested that a more relevant approach to high school science teaching would improve the occupational preparation of most science-career bound students, while at the same time improve the scientific functional literacy of the general public. “There should be a greater emphasis on the explicit teaching of procedural understanding and a reduced emphasis on the teaching of conceptual content” (p. 675).

Rationale and Purpose of the Study

The present research study extended Duggan and Gott’s work to the health professions, specifically in this case, acute-care nurses working in one hospital unit. Nursing represents a large sector of science-based jobs in Canada. Nurses must draw upon a constellation of declarative and procedural knowledge to interpret evidence in a particular context. The constellation of knowledge-in-use is of interest to this research project. By learning more about what might constitute effective prerequisite science instruction (in part based on empirical data that describe nurses’ knowledge-in-use on hospital wards), science educators can develop more effective science courses for science-based occupations, such as health professions. For example, it would be helpful to discover what conceptual content in physics has a role in nursing, given the abundance of instruments utilized by nurses. Moreover, science educators can develop a more appropriate science curriculum to improve the

general public's scientific literacy with which the public can better understand and communicate with health professionals (Eijkelfhof, 1990, 1994; Layton, 1991).

It would also be helpful to know if there is a common core of concepts of evidence used by nurses as they engage in critical thinking, problem solving, and decision making when faced with multiple demands, such as: responding to the orders from a doctor, following an appropriate protocol, gathering observational data on a patient, and responding to the patient's physical and emotional needs. An example of problem solving would be resolving conflicting observations of a patient. The topics of critical thinking and problem solving (e.g., novice versus expert) and of decision making (e.g. taking professional action) are beyond the scope of this research. These processes, however, form the context in which evidence is acquired and used; and therefore, the processes form an important context for the study.

Research Methodology

The research was a *preliminary* study carried out on a modest scale. No generalizations were sought, only: (1) a description of the concepts of evidence apparent in nurses' knowledge-in-use, (2) comparisons to other science-based occupations, and (3) implications for the content of school science. A large-scale future study may provide the type of data generalizable to other wards and hospitals, and transferable to health clinics.

In no sense of the term were nurses evaluated in this study. Although critical thinking or problem solving usually leads to decision making and then to action taken by a nurse, this study was limited to the science-related knowledge-in-use in critical thinking, problem solving, and decision making. Excluded from this study were: a nurse's skills at critical thinking, problem solving, and decision making; a nurse's capability with technology; a nurse's personal judgements about clients, doctors, and other staff; and a nurse's self-identity. No observations of patients were made.

Research Question

The study investigated the knowledge-in-use of acute-care hospital nurses, giving specific emphasis to concepts of evidence. The study investigated the following question:

While taking note of the specific declarative knowledge used by acute-care nurses in a hospital (knowledge-in-use associated with the technical field of nursing and the abstract field of science), is there a core set of concepts of evidence that can be identified?

Participants

The process of inviting nurses to volunteer to participate in the study is described here. With the approval of the University of Saskatchewan Behavioural Research Ethics Board, of the research office of the Saskatoon Health Region (SHR), and of the research office of a Saskatoon hospital, the researcher contacted three Unit Managers (administrative heads of three different types of wards, chosen by the SHR research office) and the researcher personally met with them to request their involvement in the study. Three different consequences occurred: one Manager discussed the project with her nurses at a weekly meeting and they decided they were already too busy with other initiatives to take on one more; another Manager delegated responsibility to an assistant administrator who could not locate nurses interested in participating; and lastly, the third Manager contacted nurses she thought might be interested, and when they expressed tentative interest she forwarded their name and telephone number to the researcher so he could contact them personally. This process occurred over a four-week period. The unit involved was a surgical ward.

The researcher met with each potential participant individually to give a short oral description of the study, to answer any questions, and to provide written documents (i.e. a summary of the study and the ethics contract to be used). Several days later, the researcher telephoned a potential participant at home to ask if they had further questions and to ask if they wished to volunteer to participate. All potential participants accepted; six nurses in total, four women and two men. (The surgical unit was comprised of 50 nurses in total, 43 women and 7 men.) For each nurse, a five-minute meeting was held at the hospital to sign the ethics contract, to give the nurse a miniature tape recorder (see below for an explanation), and to set a time for their first interview.

The originally proposed target to involve 12 nurses in the study was reduced to 6 nurses for two reasons: only one of the three units became involved, and it became readily apparent from the first set of interviews that each interview would produce extensive and rich data to answer the research question effectively.

The participants and the hospital were anonymous. Nurses chose the following pseudonyms for themselves: Chloe, Gia, Jamie, Joan, Sarah, and Terry.

The Unit Manager of the surgical ward was involved in the study to help direct the research to ensure minimal disruption and optimum data collection, and to interact with a draft version of the research report.

Procedures

The study clearly falls within the qualitative research paradigm. The task of the researcher was to interpret the words of the participants, in order to identify their science-related knowledge-in-use, expressed during conversations about a personal science-related problem-solving or decision-making event on the ward. Because expert performers are seldom explicitly aware of the knowledge they use at any one moment, the usual type of semi-structured interviewing is rarely successful (Duggan & Gott, 2002). Therefore, unstructured interviews were conducted and they focused on the participants' cognitive engagement in practice.

By talking into a personal miniature tape recorder during a shift, nurses identified on-going events (both normal and discrepant events) related to their evidence-based practice, and then later they were interviewed about these events, usually one per interview. Some nurses chose to use written notes instead of a miniature tape recorder. To ensure professional confidentiality between a patient and nurse, the interviewer did not observe or have any contact with patients.

Each nurse was interviewed four times during a four- to six-week period. The interviews took place in a private seminar room near the surgical unit, at a time convenient to the nurse (usually around noon for day shifts, and 10 pm for night shifts). Each interview required between 10 to 20 minutes, with most lasting 20 minutes. All interviews were audio taped. The project accumulated over 7 hours of focused discussions. Relevant portions of each tape were transcribed. Before any portion of a transcription became public data, it was cleared by the participant in terms of its accuracy in portraying the participant's meaning and in terms of it safeguarding the participant's anonymity. Each nurse scrutinized a draft of a transcript, made appropriate changes if they wished, and then signed a release statement.

The data (approximately 88 pages of transcriptions) were analyzed to tease out concepts of evidence specifically, and declarative and procedural knowledge in general, that contributed to the critical thinking, problem solving, or decision making in which a nurse had been engaged.

A draft version of the research results was written, including a description of the context of the study. The draft version was read by the Unit Manager who checked it for accuracy and for anonymity of the nurses and the hospital. She was interviewed once to discuss her reaction to the research results, and this information was included in the final draft of the Research Report. For this purpose, this interview was audio taped, relevant portions transcribed, and the final transcription signed off by the Manager.

Limitations

It is important to note potential limitations to the interview data. The interviews took the form of a conversation between a nurse and myself, an outsider to nursing. Being an outsider gave me an advantage because I could “make the familiar strange” in order to discover implicit concepts of evidence used by nurses. Making the familiar strange is a conventional process in qualitative research. However, being an outsider might have had disadvantages as well. Even though the nurses were aware of my science background as a science educator, they may have simplified their descriptions by using a non-scientific genre of communication in much the same way as they would with a patient or a patient’s relatives. Because I did not observe nurses speaking among themselves or to other hospital professionals, I have no data with which to compare those conversations with my interview conversations with the nurses.

This issue of simplified descriptions never arose during the interviews with nurses, but it was discussed in the interview with the Unit Manager. The six nurses had consistently referred to a blood pressure instrument on a portable trolley as a “Dynamap.” The Unit Manager pointed out that there were actually two types of machines that measure blood pressure: one produced by the Critikon company and one by the Welch-Allyn company. Only the former has the brand name “Dynamap,” which happens to have a poor track record for accuracy (the Unit Manager, October 14, 66-71). The nurses referred to both machines as “Dynamaps.” It is not possible to conclude whether this simplification was part of their normal professional discourse (much like using “Kleenex” to represent different brands of tissues even though “Kleenex” refers to only one brand name), or whether the simplification was for my benefit as an outsider.

My sense of my conversations with each nurse, however, was that the nurses spoke to me much as they spoke to each other professionally, because I continually had to ask them to translate abbreviations they automatically used (e.g. “BP,” “sat,” and “DC” – not “direct current” but “discontinue”), and because their description of a sequence of events relied on tacit knowledge of nursing, a situation that required my constant probing to sort out the proper sequence of events in my mind.

The first potential limitation in the data is, therefore, that one cannot be certain whether the nurses spoke in a lay genre to me as an outsider, or spoke in a professional genre to me as a science person. (Terry no doubt spoke to me in a science genre, as discussed later in this Report.)

A second potential limitation in the data concerns the scope of the study. During the 24 interviews conducted, there were about 30 events discussed, some of which overlapped between nurses. This represents a

limited number of events. Hence, some key events on the surgical ward are most likely missing from this preliminary research project.

Research Context

The knowledge one uses and the way one uses it depend on the function of the setting in which the knowledge is used. The way scientific knowledge is used in any particular setting often depends on the setting itself (Chin, Munby, Hutchinson, Taylor & Clark, in press; Layton, 1991; Ryder, 2001). Accordingly, Chin and colleagues (in press) proposed three features of any setting that involves science-related knowledge: purpose, accountability, and the substance (knowledge-in-use) found in that context. These three contextual features organize the description of the context of this research study, a hospital's surgical unit.

Quotations from participants are referenced citing their pseudonym, the interview date, and lines in the interview transcript from which the quotation was taken.

The description of the research context is followed by the study's research results partitioned into three sections: "Evaluating Data," "Concepts of Evidence," and "Scientific Knowledge-in-Use."

Purpose of the Setting

In a number of science-rich workplaces studied by Duggan and Gott (2002), the purpose of the workplace was quality control of a product or process, a purpose that affords the luxury of repeated measurements and the creation of new methods to defend claims made in the workplace. However, the main *purpose* of nursing on a surgical ward is to ameliorate the health of patients and to reduce their pain ("to improve the condition and comfort of the patient;" Chloe, May 26, 31). Given the constraints of time, resources, and the immediate consequences to a patient, empirical evidence serves a much different purpose for acute-care nurses than for workers in most other science-related occupations. One indication of time constraints was the fact that nurses were unable to participate in a research interview for about 20% of the prearranged visits to the unit, due to workload duties.

The nurses in the research study perceived their primary role as *advocates* for their patients' physical and emotional healing, in a milieu of resources (e.g. medication, tests, and procedures) and of people (e.g. doctors, fellow professionals, technicians, and visiting family and friends of the patient).

Terry: It's more my responsibility to advocate for that patient, to make the surgeon aware of what my findings are, and you say, "Well, you know, these [chest tubes] have been in for so long, and this is what's draining and there's this bubbling or tidalling" [fluctuating]. (June 15, 8-11)

Thus, the *purpose* of knowledge-in-use for acute-care nurses encompassed three domains: healing of patients, proper use of resources, and effective interaction with people. The last two domains always depended on the healing of patients – the primary purpose of a surgical ward.

Accountability in the Setting

For most science-related careers in business, industry, and government laboratories, for example, accountability is assessed with respect to the quality and efficiency of the product or with respect to the correctness and appropriate use of a procedure.

The nurses in this study talked about everyday events related to gathering and evaluating evidence in the context of clinical reasoning. Based on these focused discussions, I inferred the following outsider’s perspective on accountability in the surgical unit: the nurses were held accountable for the patient’s physical and emotional well being, for the appropriate use of resources (e.g. calling doctors/residents to perform a function), and for maintaining the hospital’s cultural standards of physical and emotional safety and comfort (e.g. managing a patient’s family and friends). Because the formal administrative hierarchy of accountability was never discussed in the interviews, an administrative perspective on accountability cannot be extrapolated from the data.

Knowledge-in-Use Enacted in the Setting

Although nurses do not stop to reflect on the various types of knowledge they happen to use during a problem-solving or decision-making event, it is convenient for a researcher to describe these types of knowledge in terms of categories found in the research literature. Categories help to articulate nurses’ knowledge-in-use that comprises an important aspect to clinical reasoning (Higgs & Jones, 2002). The categories used in this research report are summarized in Table 1 and are described here.

Figure 1 fits here.

The first distinction to be made within the category of a nurse’s knowledge-in-use is between *declarative* knowledge (propositional knowledge, “knowing that”) and *procedural* knowledge (non-propositional knowledge, “knowing how”) (Chin et al., in press; Higgs & Jones, 2002). Declarative knowledge

possessed by a nurse in a surgical ward, “declarative understanding,” can be divided into two further categories: *scientific* knowledge, abstract canonical content found in high school and university science courses – facts, concepts, and values; and *professional* knowledge of nursing, abstract and technical content found in nursing courses and apprenticeships – facts, concepts, and values.

Scientific knowledge is comprised of mechanistic explanations and classification schemes universally applicable (i.e. context independent). Its cognitive purpose is to explore the applicability of currently held paradigms and to create new knowledge by either resolving discrepancies that arise or by exploring new phenomena made accessible by advances in technology (Kuhn, 1970). Of potential interest to a surgical ward is “core science” (Cole, 1992), the kind of scientific knowledge (e.g. air pressure) that has been validated by such a strong consensus of scientists that it is not considered open to change. (Not of interest to a surgical ward is “frontier science,” which is highly tentative or speculative scientific knowledge [e.g. the link between high-voltage power lines and childhood leukemia] that lacks a strong validating consensus at the time.) Validation through consensus making is usually manifested by published articles accepted by recognized paradigm practitioners.

Declarative professional knowledge of nursing, according to Higgs and Jones (2002), is comprised of multi-paradigmatic facts, concepts, and values that give emphasis to research-based empirical information directly related to nurses’ problem solving and decision making, contextualized in clinical reasoning.

For effective clinical reasoning, we consider that health professionals rely upon the scientific knowledge of human behaviour and body responses in health and illness, the aesthetic perception of significant human experiences, a personal understanding of the uniqueness of the self and others, and the ability to make decisions within concrete situations involving particular moral judgements. (p. 27)

According to Higgs and Jones (2002, p. 28), declarative knowledge has a clear purpose: to inform wise intuitive clinical reasoning. (Figure 1 does not represent clinical reasoning, but only the knowledge-in-use involved in clinical reasoning.) Accountability in a surgical unit is tied to this purpose in the context of a unique individual patient’s well being. Professional knowledge of nursing encompasses mechanistic explanations for a particular event (context dependent), and empirical relationships and correlations (also context dependent). The following excerpt expresses a mechanistic explanation contextualized in a nursing event:

Terry: What happened was this: he was accumulating a lot of fluid in his lungs, so the membrane was getting thicker. So when you have a larger barrier between the respiratory and circulatory systems, you’re going to get poorer oxygen exchange. (June 25, 62-64)

Empirical relationships within professional knowledge of nursing are exemplified by the following excerpts:

Terry: Even a temperature of 38 degrees: we get them all the time at 38 degrees up here. And the problem is that people are not breathing deeply enough. *Deep breathing and coughing bring the temperature down.* (June 20, 167-169, emphasis added)

Sarah: Then I remembered from the day before, he had a lower potassium level, it was 3.1. So they were infusing him with some boluses to get it up. The normal is 3.5 to 5.5. *Sometimes when it's low it can cause confusion.* (June 23, 8-10, emphasis added)

Sarah: Males and females are different. Males have more [haemoglobin]. (June 16, 48-49)

Chloe: One of the comments the CCA [Critical Care Associate] made when he arrived was that *if the heart rate is greater than a rate of 140 minus the patient's age, it's not sustainable.* This is a "ventricular rate," and he [the patient] certainly fell into that category. (June 7, 134-136, emphasis added)

Further examples are cited below in the section "Evaluating Data."

The distinction between scientific knowledge and professional knowledge of nursing can sometimes be vague, but a distinction has pragmatic value. A critical feature of the context of nursing is the uniqueness of each patient. Knowledge-in-use is relevant only to the extent that it acknowledges this unique individuality. Thus, decontextualized ideas (i.e. scientific knowledge) by their nature may be out of harmony with the contingencies of a unique patient. In comparison, chemical industries do not treat molecules as unique entities; quite the contrary, all molecules of carbon dioxide, for instance, are assumed to be identical (except for their statistically inscribed thermodynamic properties). Thus, the individual uniqueness of a patient would usually demand contextualized professional knowledge of nursing rather than decontextualized scientific knowledge. The patient's uniqueness constitutes a particular context for knowledge.

The category "procedural knowledge" (Figure 1) comprises a host of facts, concepts, skills, and values, functioning at various levels of concreteness and abstraction. Procedural knowledge informs clinical reasoning (e.g. problem solving and decision making). A nurse's problem solving and decision making served as a context for the study, not its focus; therefore, these processes are not represented in Figure 1. Problem solving and decision making involve an interaction between declarative understanding and procedural understanding, according to Duggan and Gott (2002), an interaction acknowledged in Figure 1 by a simple two-way arrow.

Procedural understanding (Figure 1) is underpinned by (1) the *thinking* associated with the collection of data and the judgment of the data's significance as evidence (using concepts of evidence to do so); and (2) the *action* of nursing ("procedural capability"), that is, knowing what to do, how to do it, and how to communicate

this with fellow nurses and doctors. Action produces data that are processed using *concepts of evidence* to help judge the data's credibility. Credible data inform a nurse's problem solving or decision making. One example would be nurses using a finger probe to determine a patient's blood oxygenation saturation (the patient's "sats"); another example would be nurses providing a patient with a greater flow of oxygen and observing a change in the patient's lip colour. This evidence is then used in problem solving, the result of which is often a decision and action. The relationship is somewhat repetitive: action → data → thinking → decision making → thinking → action → data → thinking → decision making, etc.

The nature of evidence in clinical reasoning is the general focus of this research study; but in particular, the study investigated the implicit or explicit concepts of evidence used by acute-care nurses during specific daily events (problem solving or decision making) on the hospital ward.

Evidence

Evidence is normally thought of as data that have been scrutinized by various methods or validation criteria, such as comparisons with other data, or consistency with accepted knowledge (Gott, Duggan & Roberts, 1999). Scrutiny affords a degree of credibility in the data.

Different science-related workplaces have varying degrees of data richness. Cases of high complexity in some industries led Gott and colleagues (1999) to stipulate the following definitions: several *readings* produce a measurement; several *measurements* establish a *datum*; and a datum repeated over time accumulates into *data*. For simple situations, however, one reading or measurement could establish a datum, defined by Gott and colleagues (1999, p. 1) as "the measurement of a parameter (e.g. the volume of a gas)," and when repeated in concert with a variable, more than one datum becomes data (e.g. the volume of gas measured at various temperatures). A datum can be either quantitative or qualitative. An example of a qualitative datum on a surgical ward is "type of oxygen equipment" (i.e. prong or mask, along with several mask sizes).

Gott and colleagues (1999) devised a model for how a measurement develops into evidence during a science-related event, evidence which in turn is evaluated with respect to a possible outcome, such as making a decision based on the evidence. This outcome is always embedded in a social context of the science-related event (e.g. Does the product meet quality-assurance standards?). The evaluation of evidence is influenced by features of the social context (e.g. cost, practicality, and time). Figure 2 depicts this model (Gott et al., 1999). The model also frames *concepts of evidence*, that is, concepts about data and the credibility of those data (e.g. repeatability, calibration, instrument error, sampling, reliability, validity, and accuracy). Concepts of evidence are usually applied unconsciously as tacit knowledge (Higgs & Jones, 2002) to determine how credible the data

are, and then in turn how credible and important the evidence is, given the social context in which action may occur on the basis of that evidence (Duggan & Gott, 2002).

Figure 2 fits here.

In summary, concepts about data and about the evaluation of data together comprise concepts of evidence; and concepts of evidence plus the evaluation of that evidence in a particular setting, are all embraced by the model (Figure 2) proposed by Gott and colleagues (1999).

Evaluating Data

Gott and colleagues' (1999) model is applicable to the science-related work of nurses on a surgical ward. A typical datum, for example, is a patient's blood oxygenation saturation (the "sats"). It can be a quantitative measure (e.g. taken by a finger probe; 82%) or a qualitative measure (e.g. taken by observing the degree of purple-bluishness in a patient's lips). In terms of the measurement complexity found in other science-rich workplaces (e.g. chemical plants and environmental analysis labs), nursing appears to be at the non-complex end of the spectrum. Thus, the following terms were generally used interchangeably by the nurses in the study: measurement, reading, symptom, and observation. In the context of the surgical unit, these terms were synonymous with model's term *datum* and will be applied interchangeably in this report.

Becoming Evidence

Surgical nurses appeared to assess data in three different ways. Data (readings, symptoms, measurements, or observations) became evidence when: (1) a datum was collaborated by other data, (2) trends in the data were perceived, and (3) there was a consistency or inconsistency between a datum and its context. In some instances, these three different ways worked in various combinations to produce evidence.

The first category (collaborated by other data) can be illustrated by the quantitative and qualitative blood oxygenation examples just above. The two examples are directly related to each other because each datum tends to collaborate the other: patients with very low blood oxygen saturation (i.e. 82%) tend to have purple-bluish lips – a condition called "cyanosis." (This relationship is an instance of "professional knowledge of nursing – empirical relationship," Figure 1.) The two measures taken together (each a datum) produce credible data; hence, together they likely constitute *evidence* upon which to make a nursing decision. (The example of 82% is

a dangerously low reading normally, low enough perhaps for a nurse to ignore other potentially relevant data [e.g. patient's past history] that might otherwise have been integrated into a set of data sufficient enough to constitute evidence.) In other words, several measurements can become evidence if they collaborate an inference about what is happening with a patient; in this case, oxygen deprivation. This type of data collaboration draws upon different ways of measuring the same variable and is known as “validation triangulation” (a concept of evidence).

A second category for describing how observations become evidence is a trend in data, which deals with repeated measurements over time (e.g. sats of 94%, 92%, and 90%). This occurrence produces data that form a trend or pattern recognizable to a nurse, and the data become evidence; in the present example, 90% is not usually considered dangerously low, but the trend itself indicates an oxygen deprivation problem. Terry talked about a post-operative patient's haemoglobin count:

Terry: All of a sudden you're watching the red cell count go 120, 109, 98, over a period of time, and you stop to ask, “It [the haemoglobin] is going somewhere, where is it going?” (June 25, 32-34)

Chloe described the protocol for measuring blood pressure when attending a patient whose heart beat was dangerously increasing.

Chloe: ...So every three minutes it would pump up the cuff and then give you a reading.

Glen: Good.

Chloe: So then we had a sheet of graph paper and as soon as the three minutes past and the data came on the screen, we would chart them on the graph paper and we could see a trend. (June 7, 147-151)

A third way evidence emerges from measurements is when a datum, in concert with its context, becomes evidence. Jamie happened to notice a discrepant event concerning a patient in the Special Observation section of the surgical unit who had recently come from the ICU (the hospital's intensive care unit) after his operation. The patient was sitting comfortably upright happily eating a meal. These two observations (data) – upright and eating – taken out of context would have had no particular meaning. However, given the context that Jamie knew – recent surgery with a spinal anaesthetic – the data had a highly significant, potentially harmful meaning, and therefore they became evidence:

Jamie: Usually when they come back after a spinal anaesthetic, the protocol is to keep the patient relatively flat for 6 hours post-operatively, because they [the anaesthetists] drain off some of the spinal fluid.

So if one sits erect, there is not so much spinal fluid surrounding the brain. Patients can get what we call a “spinal headache.” It’s hard to treat. (May 25, 6-9)

A ward aid had helped the patient into an upright position to make him more comfortable to eat. A nurse would have inclined the patient a maximum of 60-degrees, in spite of it being less comfortable for the patient (an instance of “procedural capability,” Figure 1). Jamie’s *reason* for this protocol, captured by the quotation above, is an instance of “professional knowledge of nursing – mechanistic explanations.”

Another case in which context affected a nurse’s thinking was one of Terry’s patients who was recovering from surgery that produced a colostomy. The patient had experienced a sudden stoppage in his colon output. When the vitals were taken, the data showed higher than normal blood pressure and heart rate (100 to 110 beats per minute rather than 60 to 80). Were these data credible enough to be evidence in formulating an inference? No, the idea of “normal” needed to be contextualized due to the patient’s pain, as Terry explained (drawing upon his “professional knowledge of nursing – empirical relationships”):

Terry: If someone is in pain, you expect a slight rise in blood pressure. You expect a rise in the heart rate. If somebody is having signs of infection then you are expecting those and an increase in temperature. (June 20, 95-97)

Thus, a heart rate of 110 could be normal if someone were in pain due to a blockage of the colon. But temperature needs to be contextualized in another way as well, because:

Terry: On this unit in the afternoon in the summer time, you come in at three o’clock in the afternoon and you can see everyone is running a low-grade temperature because it’s hot outside. Your environment is hot, therefore you’re going to be warm. (June 20, 136-138)

Terry looked for a trend in the temperature change in his patient (category 2, trends in data) in light of the context (time of day and season) before reaching a tentative conclusion. In Terry’s words:

Terry: So it is not only looking at the blood pressure and thinking, “Well, the blood pressure is up.” You have to take it in concert with all of the other things [triangulation and context]. It’s only one little test and you have to take it and you have to synthesize all of the information together before you can actually even form a hypothesis. (June 20, 100-103)

Thus, the context to be considered in the evaluation of a datum can be, for instance: type of surgery, a patient’s immediate circumstances (e.g. pain), a patient’s past history, and time of day/season.

In summary, applying the model (Figure 2) devised by Gott and colleagues (1999) to clinical reasoning on a surgical ward, one detects three pathways for moving from a measurement (reading, symptom, observation, or datum) to evidence. The pathways, either singular or in combination, lead to the next stage in

Gott and colleague's model: How credible and important is the evidence? (i.e. evidence evaluation). Two main functional purposes for evaluating evidence became apparent in the nurses' interviews: (1) to move "to the next level" in attending to a patient's well being, and (2) to initiate a procedure or intervention. Each is examined separately, although in reality they naturally occur simultaneously, as indicated by events mentioned in the following two subsections.

Taking It to the Next Level

Measurements seemed to form a hierarchical pattern on the surgical ward: (1) symptoms (detected by a nurse's senses), (2) vitals (blood pressure, heart rate, temperature, respirations, and blood oxygenation saturation), and then (3) targeted tests to gather further data. Each represents a different level of data gathering, and therefore a different function for evidence at each level. To move from one level to another, nurses made a clinical decision that required credible evidence. (These transitions from one level to another are illustrated by events on the surgical ward reported just below.) Once a nurse reached level 3 (targeted tests), other levels and decisions became apparent: (3.a) tests that a nurse can carry out, (3.b) tests that require hospital specialists, and (3.c) tests that require decisions by residents or doctors. See Figure 3. To carry out these tests, instruments are used, of course, but some are simple/inexpensive while others are complex/expensive. The decision on what instrument to employ lies in predetermined protocols ("procedural capabilities," Figure 1) or in the hands of hospital specialists and doctors (e.g. whether to obtain data with an x-ray or NMI). Several events described by surgical unit nurses clarify the phrase "taking it to the next level."

Figure 3 fits here.

Often patients will report a discomfort or pain to a nurse, or alternatively, a nurse will spontaneously notice something about a patient upon approaching a bed, especially if the observation is unexpected. The following example illustrates moving from the symptom level directly to the targeted tests level.

Joan: A patient rang for me the other day and they were exhibiting symptoms; they said they felt "low". It's what the patient said specifically.

Glen: Now, when they said "low", that's a verbal message. What was the body language? By just the way they said it? What did *you* perceive?

Joan: They looked tired. They were sweating, a little shaky, felt sick to the stomach; all those kinds of things. (June 11, 4-10)

The expression “all those kinds of things” suggests Joan has tentatively recognized a pattern she associates with, in this case, diabetes (“professional knowledge of nursing – correlations,” Figure 1). She continued:

Joan: So those are the things that I see. ... Then we automatically go and do a blood sugar testing. This patient was low. They read 2.3, which is low. A normal Glucometer reading would be 4 to 6, somewhere in there. (June 11, 10-13)

In this illustration, the decision to move to a higher level of data collection was a straightforward protocol (“procedural capability,” Figure 1). The datum “2.3” had sufficient credibility in the context of the surgical units’ familiarity with diabetes to warrant a different type of clinical decision by Joan, a decision to initiate a procedure or intervention – a second function of evidence (discussed below). In this case, she gave the patient sugared apple juice, one intervention among several, each justified by evidence:

Joan: We gave apple juice with sugar in it. And, if their sugars drop too low – it’s a very individual thing. For some people, 2.3 would be low enough that they would be so tired that they couldn’t talk to you. And in that case, we would need to give them either glucose (a syrup under the tongue that would absorb) or we would have to give them IV with dextrose medication to bring the sugar up quickly. So it would just depend on what we saw that patient going through in that state. That would determine what action we would take to solve it. (June 11, 39-44)

For renal patients with high blood sugar counts, for instance 10 or higher, a different intervention would be required (Unit Manager, October 14, 111-115).

Sometimes moving to a higher level of data collection does not bring evidence that a nurse seeks, and one needs to move on to the next level, as this next example demonstrates.

Gia: I walked into a room (that was about half an hour to 45 minutes after giving an Indocid suppository to a woman patient) and she had reacted to the Indocid suppository.

Glen: What did you notice?

Gia: Something was wrong because she was very confused, her eyes were really twitching, and she said she felt very heavy and her whole body felt heavy. ... She said she felt kind of paralysed, she couldn’t move. So, I took a set of vitals, but her vitals were fine. (May 22, 4-17)

The baseline data (vitals) caused a discrepancy for Gia. Her professional knowledge of nursing did not help her in this event other than to tell her something was wrong. But she did not know what targeted measurement she should take next. She went to a higher level of *personnel* by consulting a doctor (level 3.c in Figure 3).

Gia: Indocid had affected her central nervous system. I had talked to Dr. [X] about it because I knew that something was off. He happened to be up here, anyway. He said there was a higher incident of

reaction in women than in men. And that it was not uncommon, and so he discontinued the Indocid and put her on a different pain reliever called “Naprosyn.” (May 22, 22-26)

In the future, Gia will remember this event when she gives a patient Indocid: “I had no idea that it could do that. I will put that in the vault for future reference” (May 22, 44-45). “Vault” seems to be her expression for “professional knowledge of nursing,” and in this case, “empirical relationships.”

The next example takes us through multiple transitions among levels of data gathering (symptoms, vitals, and targeted tests), people, and instruments. Joan attended a patient recovering from an esophagectomy, whose nasal gastric tube had come out on its own during the previous shift. At report, Joan learned the following symptoms of the patient: coughing out stomach content, restless, and the patient’s comment that he was “uncomfortable.” The symptom “restless” was actually an inference (not an observation) that Joan was unsure about, so she checked on the patient.

Joan: I walk into the room and I first just look at him and I can see what she means by “being restless.” He is sitting up, he’s lying down, he’s sitting up, he’s lying down. He cannot keep comfortable. He’s rocking back and forth. I could just tell something is wrong. So, then I ... took his vital signs. When we say vital signs we mean: blood pressure, pulse, temperature, respirations, and oxygen level. We did all those things. By doing that, it gives you a baseline and you can see where everything is at. His oxygen saturations were down. They were 88% on 5 litres of oxygen by nasal prong. He had little nasal prongs in his nose, giving him oxygen at that time. (May 29, 72-80)

Before going to the third level, targeted tests, Joan made a decision concerning an appropriate intervention:

Joan: Knowing that his oxygen sats were only 88% on 5 litres, obviously I need to do something different here. The next step is to put him to a simple mask. I changed him to a simple mask at 10 litres, but that still made no difference. (May 29, 88-90)

This datum of “made no difference” in the context of the patient’s immediate history required more data before it could become evidence. This Joan acquired from a stethoscope reading, and then she made a clinical decision about going to the next level of data collection, one that involved another professional in the hospital:

Joan: So then I listened to his lungs using the stethoscope. His air entry was fairly good, but it was quiet to the bottom of his lungs. He was not getting a lot of oxygen into the bottom of his lungs. But they did not sound very wet or congested. There are lots of different sounds you might hear when you listen to lung sounds. But he was short of breath. He was breathing quite rapidly. His breaths were more shallow. So with these things in mind, I called a respiratory therapist. (May 29, 92-97)

The respiratory therapist had the authority to call for other tests, such as an arterial blood gas test:

Joan: That checks the level of pH, carbon dioxide, oxygen, and all these kinds of things in the blood. And that will help us determine if this person is having a respiratory response or his oxygen is failing because a different system in the body is causing this problem. (May 29, 108-111)

Meanwhile, Joan was wondering, “Why is this man not getting the oxygen that he needs?” (May 29, line 105), as the patient continued to fail, indicated by his slipping into unconsciousness and his lowering sats. The respiratory therapist also ordered a chest x-ray. Now there was sufficient data to become credible evidence for the following inference proposed by the doctor who analyzed the x-ray (Joan, May 29, 124-126): the gastric juices coughed up by the patient were aspirating into his lungs, inhibiting the oxygen from reaching his circulatory system. The final decision was to move the patient into the ICU.

Joan’s event illustrates how clinical decisions are made on the basis of evidence, decisions concerning: going to the next level of data gathering (symptoms, vitals, targeted tests); getting other people involved (e.g. a respiratory therapist); and choosing what instruments to use next (e.g. from the stethoscope to x-rays, a decision dependent upon one’s authority within a hospital). Given the constraints of time and resources of a hospital, these decisions are based on the evaluation of the evidence that might warrant the decision to go to the next level. Joan succinctly summarized this conclusion herself: “To make the choice, evidence is necessary to go to the next point” (May 29, 131).

Not all choices are so straightforward, however. Sarah found herself in an awkward yet not unusual position of deciding whether or not to carry out a doctor-ordered intervention. This decision was directly connected to another decision: whether or not risk going to the next level of involving a doctor. Both decisions focus on the evaluation of evidence (including the lack of evidence in this case – Had the doctor recently seen the patient?), evidence relating to a patient’s well being, and to the social context of the surgical unit and hospital.

Sarah: The other day I saw an order to discontinue a Foley catheter from one of our patients. I saw that the order was written around 10:30 [a.m.] and so I got to it around 11. When I went to the patient, I was going to DC [discontinue] the Foley. But he had this Foley catheter in for quite a while [6 days] and he was quite edematous with his penis and his scrotum. (June 21, 4-7)

Other data included: edematous in the legs, pitting edema of the feet, looked bigger than usual for a small man, and the fact he had no past history of problems voiding.

Sarah: I thought to myself, “Well, should I take the Foley out? Has the doctor seen this or does he presume since it’s been in for so many days, you might as well discontinue it?” (June 21, 11-12)

...

He was an older man, probably around 72. So, I was concerned about discontinuing the Foley because we usually wait about six hours and if he doesn’t void, then we will call the resident or call the doctor. And it was around 11 o’clock. (14-16)

Time became an issue, along with the possibility of a doctor making a special trip to the hospital just for a relatively minor procedure.

Sarah: So at that time, around 5 or 6 p.m., usually the doctors are not around so then we would have to call them in and they would have to put another Foley in if he was unable to void. Considering: the time frame, not having communicated with the doctor about seeing it recently, and seeing what's been going on, I didn't know if I should discontinue it or not. So I decided I would leave it in until the doctor either came up or was notified. (June 21, 18-22)

Complications could arise if the patient was unable to void during the six-hour waiting period: "His bladder could have become full and it could have blocked into his kidneys" (line 34). Possible harm to the patient became part of the context in which the data (edematous in the lower body and no prior history of problems) reached the status of credible evidence to warrant not discontinuing the Foley catheter.

Sarah's sensitivity to calling doctors into the hospital may have stemmed from earlier experiences.

The next event sheds more light on this one small aspect of the social context of the evaluation of evidence (Was the situation serious enough to warrant calling a doctor?).

Sarah: The other day I was just getting some blood work back from the lab. There was a low haemoglobin of – I think it was 77 [a drop of 6 points in two days]. So, when you first look at that (and it was during the day), you have to wonder who to call. First of all you would call the resident. But since they are in the O.R., they don't like to be paged in the O.R. So I wondered, "Should I wait before I call them or should I just call the O.R.? How important is it?" All the time we have the problem of whether or not to call the doctor or the resident. (June 16, 4-9)

The normal intervention for a low haemoglobin count is a blood transfusion, but a doctor or resident can only order it (given the patient's permission). Again, the datum "77" by itself was not credible enough (i.e. not low enough) to become strong enough evidence for interrupting a resident in the O.R. However, the "drastic" drop (from 83 to 77 in two days) was a crucial difference. The context? – The patient was recovering from an amputation. Other data accumulated: the patient had difficulty getting out of bed (i.e. he was lethargic) and he looked pale. On the other hand, all his vitals were normal and his complete blood count (CBC) taken at admission had also been low (i.e. the patient was originally slightly anaemic), which means that his normalcy is not the normal 130-170 range of most people. An important feature of this event – its context – was clarified by Terry when he discussed haemoglobin data:

Terry: Now, depending on the hospital unit, we [in the surgical unit] don't get concerned until it gets to be about 80, when we think seriously about transfusing someone. If it's at 85, 89, then it's something to mention, and once again, it's trending [upward or downward trends]. If someone is chronically anaemic, then a low count is going to be normal for them, and their body has adapted to it; which is very different from someone who has a gastro intestinal bleed, someone who is bleeding heavily from an ulcer or something. (June 25, 27-32)

Sarah, by the way, solved her problem by leaving a message for the resident at the O.R. to inform them of the low haemoglobin count.

Initiating an Intervention

We have seen that Joan gave sugared apple juice to a patient who had a low Glucometer reading, and Sarah did not discontinue a Foley catheter from an edematous patient. Both events show how evidence is evaluated in procedural understanding to warrant initiating (or not initiating) an intervention. A more extensive example will further clarify the evaluation of evidence when it is used to act on a patient.

Under Chloe's care, a patient recovering from vascular leg surgery showed the following symptoms (data): increased pain in the calf when the patient flexed his foot, redness of skin, hot to the touch, and the patient was reluctant to get up. The patient's pain in this context had special meaning to a nurse ("Professional Knowledge of Nursing – Empirical Relationships"):

Chloe: The other significant thing was with the pain; it wasn't the fact that he had pain in the calf, but the fact that when he flexed his foot the pain in the back of his calf got worse. It is a positive Homan's sign, so it's a specific pain that worsens with a specific movement. (May 26, 66-69)

Another concept in the professional knowledge of nursing was the correlation between pain and swelling in this context (71). These data became evidence to warrant going to the next level of a targeted test, in this case measuring the degree of swelling over time (data that showed a 2 cm increase in the leg's circumference over 3 hours). Now, the data reached the status of evidence to support a concern that the patient may have a deep vein thrombosis (DVT, also known as a blood clot). Was the circumference increase of 2 cm the evidence by itself? No.

Chloe: Not in isolation. But if there was significant pain when he flexed his foot, and redness that was hot to touch, all of those things together. So, it is not necessarily any one factor in isolation, but all of them together, you'd want to be sure there was no clot, and that the symptoms were caused by something else. (May 26, 88-91)

Chloe did have evidence for an immediate intervention (i.e. applying anti-emboli – anti-clotting – stockings to the patient's leg) and for going to a higher level of targeted tests (level 3.c) by talking with a resident who authorized a Doppler ultra sound, the result of which ruled out DVT. The resident was then able to tentatively account for the patient's pain by focusing on the muscle damage caused by the surgery. The anti-emboli stockings (support hose) resolved the patient's swelling and pain within a day.

In Chloe's scenario, a cluster of symptoms became evidence for moving to the next level (level 3.a), which was a targeted test (leg circumference measurement) and which in turn yielded a validity triangulation datum (2 cm increase over 3 hours). Both the cluster of symptoms and the triangulation datum suggested the

possibility of DVT (i.e. a blood clot) and led to the decision to initiate an intervention (i.e. applying anti-emboli stockings). Validity triangulation is a concept of evidence explored in the subsection “Accuracy,” below.

Summary

The research results reported in this section, “Evaluating Data,” illustrate various circumstances by which measurements (observations) became evidence on a surgical ward: data collaboration (recognizable patterns or triangulation), data trends, and consistency/inconsistency with the context. The results also indicate two functional purposes for which nurses evaluated evidence: moving to the next level, and initiating an intervention. These results form the context for examining the central issue of this project: concepts of evidence used implicitly or explicitly by acute-care nurses.

Concepts of Evidence

In the process of gathering and evaluating data to determine if the data warrant the status of evidence, and in the process of evaluating evidence to decide what to do next, people use conceptions (or misconceptions) concerning data and evidence (Duggan & Gott, 2002). Gott et al. (2003) provide an encyclopaedia of “concepts of evidence” derived from research in the UK into events experienced by people in science-related careers working in science-rich workplaces such as: a chemical plant specializing in colourants for foods, cosmetics and pharmaceuticals; a biotechnology firm specializing in medical diagnostic kits; an environmental analysis lab; an engineering company manufacturing pumps for the petrochemical industry; and an arable farm. In addition, their research was conducted with parents who had no particular science background but who had to decide whether or not to have their infant child immunized (a science-related decision). Each group of people shared some common concepts of evidence, though each set of concepts differed somewhat due to differences in workplaces or decision-making situations (Duggan & Gott, 2002). Thus, we should expect the set of concepts of evidence employed by acute-care nurses to differ somewhat from chemists, biotechnologists, and pressure physicists.

The present study did indeed find that the science-rich surgical unit differed noticeably from the workplaces studied by Gott and colleagues (2003). The comparison between the nurses’ concepts of evidence and the compendium of concepts of evidence published by Gott and colleagues serves as a method for describing the concepts of evidence used by nurses, and is not in anyway an evaluation of nurses. Comparing and contrasting is a reporting strategy, nothing more.

Reliability

According to Gott et al. (2003), the scientific meaning of reliability usually refers to the *consistency* of readings when multiple readings are gathered. Reliability generally is enhanced by: (1) repetitive readings from the same instrument (e.g. measurement of blood alcohol level can be assessed with a breathalyser, but at least three independent readings are made before the measure is considered legally reliable evidence in Canada); (2) multiple instrument readings using similar types of instruments, a procedure often called “measurement triangulation;” and (3) multiple observers (e.g. spot checks of measurement techniques by co-workers are sometimes built into routine procedures) to minimize human error in the use of an instrument.

A fundamental concept of evidence that underscores reliability is the concept “non-repeatability:” repeated measurements of the same quantity with the same instrument seldom give exactly the same value. The sensitivity of an instrument is a measure of the amount of error inherent in the instrument itself (i.e. measurement error). Sensitive instruments produce less fluctuation in their readings (i.e. they have low measurement error). One way to express sensitivity or measurement error is with a \pm value.

Reliability decreases as an instrument’s measurement error (its \pm value) increases. Thus, a datum is weighed as evidence by considering the instrument’s measurement error and by considering the measurement procedures that have been ascertained (e.g. the reliability of a measurement of a blood alcohol level should be assessed in terms of the measurement error associated with the breathalyser [e.g. ± 0.01] and in terms of how the measurement was taken [e.g. superficial breathing versus deep breathing by a subject]). To investigate a patient’s source of pain, for instance, reliability of the investigation’s design would include an assessment of each measurement and every datum. Factors associated with the choice of measuring instruments must also be considered, for instance, the measurement error associated with each instrument. These concepts of evidence are generally associated with reliability in science-rich workplaces.

Some of these concepts of evidence apply to a surgical nurse’s knowledge-in-use, but some do not. The first concept of evidence listed above was *repetitive readings from the same instrument* (after which an average datum is calculated), that is, “repeated measures.” When Chloe discussed her measurement of the circumference of a patient’s leg, the following exchange occurred:

Glen: When you explain carefully how you want things to be measured, there is an old problem of, “How do you know that if you measured it directly afterwards, you’d get a slightly different measure only because of the tightness that you held it [the measuring tape]?” Do you re-measure or do you just take one reading?

Chloe: Usually just one. (May 26, 63-65)

Nurses seldom had time or the need to take several measures and calculate an average value because the purpose and accountability in the surgical unit militated against it. Precious time could better be spent acquiring triangulation data (the second concept of evidence, stated above) that produce more credible evidence for a nurse to decide what to do next. My questions about taking repeated measurements were often met with either polite incredulity or a diversion of the conversation to a topic that made sense. Several excerpts from interviews (below) illustrate the low status afforded “repeated measures” concept of evidence.

When asked about taking an immediate second reading from a Dynamap machine, Terry described how he would compare the original datum to its context rather than take a second reading, which is one of the ways a datum acquires the status of evidence. If Terry detected a discrepancy in a Dynamap reading of a patient’s blood pressure, he would use a different instrument to measure the blood pressure (e.g. a manual reading), thus demonstrating the use of the concept of evidence “triangulation.” He would not double check the Dynamap reading. Terry also used another concept of evidence about *how* the measurement was taken (i.e. instrument use: in this case using a proper cuff size).

Glen: So what I was focussing on was when you take a reading, how do you know if you need to take another reading for just –

Terry: It gets to be intuition.

Glen: You told me you take it and look at the chart and if it’s that much different than the chart –

Terry: Then I immediately go to the manual reading, because I want to know exactly what I am dealing with.

Glen: So it’s more consistent with –

Terry: If I get a big change, first thing I do is check to see if I have the right cuff size. If I have a larger cuff on the machine, I’m going to get a lower reading. If I took your blood pressure with a paediatric cuff right now on your arm, you would have an outrageously high blood pressure. (June 8, 161-170)

With time constraints and pressure to go to the next level (if necessary), the first and only measurement (datum) is often assessed in terms of its consistency with the context (using one’s “intuition,” as Terry stated above). Gia expressed the idea slightly differently:

Gia: I think that around here our gut judgement is everything. And just because it’s a machine, doesn’t mean that it’s always right. (June 13, 136-137)

Chloe described a typical context for a patient’s symptoms (data) when she talked about a patient whose heart rate had climbed to 140.

Glen: Were there some visual signs you were automatically looking for? The colour and things like that?

Chloe: Yes, he was pale. He didn't start to go blue at all. He was grimacing in a way that is quite typical of someone having a heart attack, in that he was clutching his fists in front of his sternum and frowning. So, he was clearly having that sort of expression of cardiac pain. (June 7, 124-128)

If a measurement (in Chloe's case, a heart rate of 140) is not consistent with a context of symptoms or with a nurse's practical knowledge (i.e. intuition or gut judgement), then a nurse will usually go to the next level (i.e. they apply the concept of evidence "triangulation"), as Gia did:

Gia: I think you always have to go with your gut feeling and if it's not what you expect, then find something more accurate.

Glen: Right, instead of just measuring it again.

Gia: Yes. (June 13, 139-140)

Nurses did not tend to think of an instrument as having an inherent measurement error. For instance, when explaining the fluctuations in oxygen saturation measurements produced by a figure probe instrument, Joan talked about a patient's condition changing, and about the need for validity triangulation with more accurate data ("validity triangulation" is a concept of evidence taken up below in the subsection "Accuracy"):

Joan: It [the sats] can change often, all the time, a very little bit. But if all of a sudden the person were to get extremely short of breath, it can drop to a significant number, very quickly. And that will be alarming.

Glen: That's good information, because now I can ask, "When you take the reading, how do you know it is the right reading rather than one of the fluctuations?"

Joan: This oxygen finger probe can be backed up by a blood test of the oxygen. And that one will be more accurate. (May 29, 17-23)

Gia (June 13, 16-20) mentioned the temperature of a patient's hands as a factor that might cause fluctuations in a finger probe reading. Similarly according to Jamie, the margin of error (the \pm value) in a haemoglobin measurement was not caused by the instrument itself but was caused by other factors that could affect the measurement using a sensitive instrument:

Jamie: Depending on, again, what's happening, what kind of surgery they've had. In some kind of surgeries we expect them to bleed a moderate amount. Other surgeries, you don't. (June 18, 94-95)

The same issue (measurement fluctuations not being caused by the instrument's inherent measurement error) emerged during a discussion about fluctuations in Gia's measurement of blood pressure:

Glen: ... Does it fluctuate an awful lot? So, when you told me 180, is it sort of like a weigh scale and the digitals are changing from 75-85 and you just sort of have to look at it for a few seconds?

Gia: Yes. Often it will change frequently depending on what they [patients] do. I mean, your heart rate often goes up with movement and all of that. So, it does fluctuate a lot, especially when people are in arrhythmias (a rhythm other than what is normal) and it often is irregular. So you expect that if someone is in atrial fib, or something like that, the pulse will change a lot because it is irregular. And whatever it is at that second, that's what the machine will pick up. So it might say 78, 76, 50, or 100. It may go like that every second. (June 4, 87-95)

When she spoke about rounding off fluctuating heart rate data, she said:

Gia: It usually depends on how erratic it is. If it is very erratic, then I probably would, but you don't often see the huge fluctuations as I had given you as an example. You don't often see that. You *can* see that but often it will be 177, 179, you know, a 182. So I would tend to say 180, I print a strip and often it says on the strip what it is, at that moment. (June 4, 108-111)

...

Gia: ... But in a hospital after surgery, we have so many playing factors in blood pressure. For instance, pain control (morphine), will drop your blood pressure. Epidurals (for example, ...) affect blood pressure. So to us, acceptable blood pressure in the hospital post-op is different than what acceptable blood pressure would be in a doctor's office. So even if someone is a little high or a little low, especially on the low side, we take into consideration what we've given them [i.e. pain control medication]. So, even if we see a blood pressure of 100 on 60, often we are not concerned because we have caused that blood pressure. And it also depends on what they are normally. Some people run with a low blood pressure that is normal on your everyday day-to-day. (June 4, 124-132)

When discussing the possible fluctuations in a blood sugar count produced by a Glucometer, Joan believed the measurement did not fluctuate. She explained this by the fact the Glucometer was a very accurate instrument (accuracy is a different but closely related major concept of evidence, taken up below).

Glen: I'm just wondering. Let's say my job was operating this machine, and I did a test and it read 2.3. In your experience, how much would that 2.3 fluctuate if I did it a half minute later?

Joan: On the same person you mean?

Glen: Yes.

Joan: Oh, very little. It would be very accurate. It's a very accurate reading.

Glen: So it doesn't fluctuate much?

Joan: From minute to minute? No. From hour to hour? Of course. Because depending on what they [patients] have eaten. But it's a very accurate test. (June 11, 64-72)

Here again we see an instrument's measurement error being ignored while a nurse focuses on factors related to a patient's unique individuality and well being. On the other hand, Joan remembered once taking a second measurement in a case of a manual blood pressure reading:

Glen: Is there a situation that you recall that you would have taken it and said, "Well, I'm not sure," and actually took it again. Or is this so accurate you just need to take one –

Joan: The manual, you mean?

Glen: Yes.

Joan: Oh no, I've rechecked myself in some instances. Also, depending on the patient, you might check both arms and compare. (June 18, 55-62)

...

Glen: You mentioned sometimes you would take a reading on both arms, sometimes you'd just repeat it –

Joan: If it were abnormally low, I would check both arms, definitely. So say, 85 on 40 would be low. You know, it gets lower than that.

Glen: We're just talking about where you'd double check.

Joan: A low number. If I got an 85 over 45 blood pressure, I would re-check it.

Glen: Repeat that reading, either on the same arm or different arm?

Joan: The other arm. You know, you could do it a couple of times, and even come back and do it again in a few minutes. (76-84)

When the Unit Manager distinguished between the Dynamap (the Critikon portable blood pressure machine) and the Welch-Allyn portable blood pressure machine, she described occasions when nurses actually measured one instrument against another (measurement triangulation).

UM: I've seen people take a blood pressure with a Critikon, check it with a Welch-Allyn, and then if they are still in doubt they'll do a manual. But there is enough doubt about the Critikon's accuracy that the nurses are not really that confident in its measurement. (October 14, 60-71)

This type of event did not arise in the 24 nurse interviews conducted during this research project (a circumstance mentioned in the subsection "Limitations"). Except for a few isolated instances, however, the concept of evidence "repeated measures" did not seem to guide a nurse's actions.

In science, as mentioned above, the degree of reliability (or an instrument's sensitivity) is conventionally expressed in terms of an instruments' error of measurement, the \pm value associated with a measurement. This simple value often masks statistical assumptions and reasoning that underlie the concepts "measurement error" and "confidence limits." As already indicated, the nurses did not seem to have had a need to consider \pm values during any of the events they discussed. However, when I specifically asked them to consider the \pm value for a measurement they had taken, it turned out (as the excerpts below indicate) the \pm value that scientists associate with measurement error was marginalized or ignored due to two other issues more important to a nurse's clinical reasoning: (1) the variation in a reading is accounted for by a patient's *unique differences*, differences that supersede any \pm value inherent in an instrument reading; and (2) the variation in a reading is accounted for by *changes* in a patient's environment or body system, changes that supersede any \pm value in an instrument reading. In Gia's discussion (above), she accounted for fluctuations in heart rate measurements by mentioning a patient's arrhythmias and the context of what is normal for the unique individuality of a patient. In Jamie's discussion (above), he rationalized a haemoglobin measurement fluctuation by describing contextual factors (e.g. type of operation). Later in Jamie's interview, I altered my approach from talking about measurement fluctuations to talking about changes between consecutive readings and how big a change would cause concern on his part. We discussed a patient whose haemoglobin count had dropped 5 points from 90 to 85.

Glen: ... to you that's within a range where you wouldn't be concerned enough to take action, but you would point it out to people.

Jamie: Yes, I'd point it out and I'd just be having a look at the patient to see if they didn't look anaemic and pale and wiped out. (June 18, 121-124)

...

Glen: ... If it was 90 to 88 [a drop of 2 points], or something, is that worth bringing to someone's attention?

Jamie: No.

Glen: Okay, so –

Jamie: I don't think I would. The doctors would probably check it every day and if they didn't, well, I don't think I'd be too worried about it, because it fluctuates a bit, day to day.

Glen: Okay, that is what I was wondering, too. Normally, to what degree does it fluctuate, without surgery and things like that? So plus or minus two is sort of a very safe range –

Jamie: Yes –

Glen: that you wouldn't even bring it to someone's attention.

Jamie: No. Probably even three or four.

Glen: Okay, so my example with five just moved beyond the “normal” range for you.

Jamie: Yes. A range of five is what I’m kind of thinking. (133-145)

In other words, a change of about ± 4 was a minor change in a haemoglobin count, and would not cause Jamie to go to the next level of appropriate action; but this change is thought to be due to a patient’s condition changing, not due to an uncertainty in the measurement. A change of 5 units, however, caused Jamie to go to the next level of appropriate action because it indicated a concern for the well being of the patient.

Similarly, Sarah considered a drop of 6 in a haemoglobin count to be “drastic” and she took action (her patient had gone from 83 to 77). But a change within ± 1 would have been considered insignificant by Sarah.

Glen: And if ten minutes later another sample (let’s say from the other arm) was sent down, would the reading of 77 be in both of them? Or is there kind of a fluctuation –

Sarah: There may be a fluctuation of one or two, not a lot. I would say around the same for sure. It’s hard to say. ... But it would be around the same. Because it’s pretty much –

Glen: Plus or minus one?

Sarah: Yes, but not a lot.

Glen: So, if it was 83 and it came back as 82, you’d say it hasn’t changed for me.

Sarah: Yes.

Glen: 81, 80, the fact that it was –

Sarah: It’s just that this was so drastic. If it was, say 83, and it came back, and let’s say it was 81, well, that’s not much of a change. (June 16, 59-70)

I interpreted Sarah to mean that the instrument that produced the haemoglobin count did not have a error of measurement, and that fluctuations reflected something happening with the patient. A change of ± 1 was considered by Sarah to be an insignificant change in *the patient’s condition*, rather than being within the error of measurement of the instrument.

Another complicating factor related to fluctuating measurements on the surgical ward was whether the measurement fell within the normal range of measurements for the particular circumstances of the patient’s unique individuality, or whether the measurement fell outside the normal range. For instance, when talking to a patient about their blood pressure, Jamie rounded off the reading to the nearest 5 when the reading was within

normalcy for a patient, but Jamie did not round off the reading when it was outside the normal range (June 18, 68-78). Although Joan never rounded off measurements, she too considered a variation in the systolic pressure of ± 5 to be insignificant when it lay within the patient's normal range, but was sensitive to an even smaller change (e.g. ± 2) when it lay outside the normal range (June 18, 88-99). Terry considered ± 10 to be an insignificant change in the systolic pressure within a patient's normal range (June 8, 149; June 20, 110), and ± 5 for the diastolic pressure (June 20, 116). In the following exchange, he underscores the practical reasons for nurses to ignore the concepts "measurement error" and "repeated measures."

Glen: So if this person who in this circumstance had a blood pressure of 160 over something and you came back half hour later and it was a 170 over whatever, you would think, "Well, maybe that's in the reading," rather than –

Terry: Well maybe it's in the reading but, once again, you are going to ask, "And what else?"

Glen: "And what else." Okay.

Terry: What else is causing that? How is he lying? Was he sitting up the last time the blood pressure was taken? Because if your body is *sick* and weak, it doesn't compensate for lying down and standing up. (June 20, 117-124)

Of prime importance to a nurse is "what else?" because the central purpose of a nursing unit is "to improve the condition and comfort of the patient" (Chloe, May 26, 31), not to justify a measurement on the basis of the instrument's reliability. In other words, the surgical ward was patient-oriented more than it was measurement-oriented. From the surgical Unit Manager's way of thinking, "this is what encompasses the *art* of nursing" (UM, October 14, 47, emphasis in the original). In contrast to a surgical ward, science-rich workplaces that are product-oriented must rely heavily on an instrument's reliability to claim a certain product quality, thus making these workplaces more measurement-oriented than is the case for a surgical ward.

Jamie expressed this point of view when he steered our conversation away from error of measurement in a blood pressure reading, to the topic of a patient's uniqueness as a cause for fluctuating readings from an instrument. He seemed to use "accuracy" in this context to mean reliability, as defined by Gott et al. (2003).

Glen: Some instruments, they just never give you quite the same reading, so you actually have to read it a second or third time and take the average, you know, that kind of thing?

Jamie: I think we probably do that a lot in nursing. I don't think anything we do here is 100% accurate, because it varies so much from person to person. The accuracy is not really all that important in some of the things we do (in some of the readings), because people want to be in homeostasis so everything just balances. I think as long as everything is balancing for the person, accuracy is not really that big a deal. If I was talking to someone whose blood pressure was 119 over 79, I'll just say their blood pressure is 120 over 80. (June 18, 68-76)

...

Jamie: As long as they are within a certain range, I am not really too concerned about the exact accuracy. Once it gets out of a specific range, like you say, I'm more concerned with how accurate it is. (81-83)

Gott et al.'s (2003) third approach to establishing reliability of a measurement was multiple observers taking identical readings with the same instrument. On the surgical ward, this procedure did not appear to occur, likely because there was neither the time nor the need. Pervasively, nurses did seem to repeat measurements (e.g. a patient's symptoms or vitals) when they first took on responsibility for a patient (e.g. after a shift change). Nurses did so not to double check the previous nurse's measurement, but to continue a collection of data on a patient to look for a pattern or trend that might be important if a nurse had to decide whether or not to go to the next level.

In summary, the six surgical ward nurses who participated in this study were guided by two key concepts of evidence associated with reliability: measurement triangulation, and how a measurement was taken (i.e. instrument use). Key concepts of evidence that were seldom relevant to the nurses were: measurement error, repeated readings with the same instrument, and repeated readings by different observers. These results must be qualified by other concepts of evidence that nurses used but were not listed in Gott et al.'s (2003) compilation of concepts of evidence: (1) normalcy range, that is, a reading either lies within or outside what is normal for a person; though for some measures such as a potassium level, several categories outside of normalcy were considered (e.g. low and very low, or high and very high), or heart rate (e.g. low and dangerously low, or high and dangerously high); and (2) a patient's unique individuality, a concept that helped define normalcy, and that overshadowed the concept of non-repeatability. As a consequence, when the nurses worked with evidence on a surgical ward there was little generalizability, but instead, there was always transferability to the specific context at hand.

In the context of nursing, measuring is more than simply representing a physical condition of a unique individual patient (e.g. blood pressure). A measurement often represents a *changeable* (i.e. dynamic) condition of that patient. This unpredictable variability within a patient's complex body can affect an instrument reading, which may or may not have critical implications for the patient. This was Terry's concern when he asked, "And what else?" (June 20, 120). By contrast, in many industries what is being measured (e.g. gas pressure) is generally assumed to remain static during the measurement process, and consequently, the fluctuations in measurements are attributable to the measurement process itself, that is, the error of measurement. But for the nurses in this study, fluctuations in measurements were either attributed to the changing condition of the patient being measured, or to the inaccuracy of an instrument, in which case a nurse engaged in validity triangulation

with a more accurate instrument that used a different process to arrive at a measurement (a topic to which we now turn).

Accuracy

The quality of a scientific measurement is determined by its reliability and by its validity. Validity is concerned with: “Does the reading actually measure what is claimed to be measured?” (Gott et al., 2003, 9.2). For instance on a surgical ward, the sat monitor’s finger probe produces a reading that is claimed to measure a patient’s blood oxygen saturation. But it may also inadvertently measure a patient’s smoking behaviour (yellow fingernails), a patient’s hand temperature, or a patient’s haemoglobin count (Gia, June 13, 16-20). Depending on the patient, the finger probe may not yield a valid measure of blood oxygen saturation.

Police measure the blood alcohol level of a person by using a breathalyser and then they can crosscheck that measurement with a blood test. Crosschecking with a different process to measure the same variable is “validity triangulation,” a concept of evidence that was illustrated in some of the events already described in this Report.

Validity is a broad concept of evidence often discussed in science-based industries in terms of how close a piece of evidence comes to the “true” value; in other words, a measurement’s *accuracy* (Gott et al., 2003, 6.1). The two concepts, validity and accuracy, are very closely related. The word “accuracy” refers to a less abstract concept and it appeared in the nurses’ transcripts, while the more abstract “validity” never did.

Although the two concepts of reliability and accuracy differ, they are related when one judges whether or not some data should be considered as evidence. One’s confidence in a measure’s accuracy will be influenced by the measure’s reliability, for instance, unreliable readings do not engender the belief that an averaged datum is particularly accurate.

During my early interviews with the nurses, they talked about a machine in terms of its *accuracy*, which then became a topic of discussion in a later interview with several nurses. I asked, “Of all the machines that are used on this ward, which one do you think is the *most* accurate? And which one do you think is the *least* accurate? As one should anticipate, the word “accuracy” had nuances of meanings depending on the context.

The nurses’ interviews strongly reflected the belief that nurses should not trust a machine, especially when a nurse can take a manual reading (e.g. for heart rate, blood pressure, etc.). When talking about a heart rate reading of 140, displayed on a Dynamap digital screen, the following exchange occurred.

Chloe: At that point, it was a case of implementing some other means of gathering evidence. I've always been taught never assume your machine is right. Get your hands on, take a radial pulse, get your stethoscope out and listen yourself directly to the heart.

Glen: That's why you used the stethoscope.

Chloe: Yes. By then, the rate I heard myself was 170. So then when we attached the cardiac monitor, we saw it [the patient's heart rate] was all over the map and still going up. (June 7, 115-120)

A similar view was expressed by Joan, Jamie, and Gia (already quoted above).

Joan: ... Anything measured manually, to me, is going to be the most accurate. So, if I were to do a blood pressure with a machine and do it manually, the manual blood pressure, the one that you hear with your own ears, to me, is always more accurate. (June 18, 18-20)

Jamie: ... When you use the Dynamap, the actual electronic one that just reads it, you totally trust the machines. If it malfunctions at all, you won't really know just by looking at the reading. It could be right or wrong. But if you are actually listening to the blood pressure yourself, with a stethoscope manually, then it's a whole lot more accurate than relying on a machine. You are sort of doing it a manual way. You are just using a device that assists you, you are actually using your own ears and assessing it yourself. The electronic Dynamap just does it totally on its own. So, you're right or wrong depending on how accurate the machine is. (June 18, 14-21)

Gia: I think that around here our gut judgement is everything. And just because it's a machine, doesn't mean that it's always right. (June 13, 136-137)

However, when Gia talked about a patient with a problem heart, hooked up to a 3-lead cardiac monitor (not a Dynamap), and talked about a medical paper trail, she seemed to put more trust in this particular machine:

Gia: ... But in the heat of the moment, when everything's happening, I tend to trust the monitor unless something would spark me otherwise. If I took a radial pulse and it didn't match the monitor, I would tend to think that the radial pulse I was getting is wrong because when it's that quick it's difficult to count every beat; whereas, the monitor would actually pick up every beat.

Glen: But, when you have a moment, then you would go ahead and print it out just to have a –

Gia: A copy of it to put in the chart for proof, or evidence, ... [we'd have proof] that that was really on the monitor.

Glen: In business, they call it a paper trail.

Gia: Right. It's our medical paper trail. (June 4, 71-79)

Here Gia has provided an additional social context for dealing with data: permanent records create evidence to be used in the distant future (e.g. perhaps if a review took place), in addition to be used immediately to decide on an intervention. In the specific context described by Gia, a computer generated printout had greater value

than a nurse's manual reading. Except for this one occasion in which Gia questioned the accuracy of a manual heart rate measurement taken "in the heat of the moment," the nurses never once mentioned human error inherent in a manual reading when they discussed accuracy. The Unit Manager, on the other hand, spoke specifically about human error in a manual reading:

UM: ... There is so much subjectivity in a manual measurement of blood pressure. Yes, you can say, "I'm confident that his blood pressure was 100 over 70, because that's what I heard." But maybe now that I'm older my hearing isn't as good as it used to be, and some young 22 year-old might take it and all of a sudden it's 120. Well that's a significant difference that a machine would probably have picked up. I think that, although we've been traditionally taught that manual and tactile measurements are the most accurate, in some ways we don't realize how subjective some of them are. (October 14, 55-61)

Once again, context is everything in nursing. The crucial role of context in the evaluation of evidence is represented by the outer circle of Gott et al.'s (1999) model for measurement, data, and evidence (Figure 2).

Just above, Gia stated that a 3-lead cardiac monitor was more accurate than a manual reading. Her concept of accuracy seemed to be related to the *detail* provided by the machine.

Gia: It [the 3-lead cardiac monitor] takes a reading of your heart and translates it to the monitor. And it makes waveforms on the screen. To us, every wave means something different that the heart is doing. And depending on how many of those waveforms you get in a certain amount of squares (which is time) that tells us what the heart rate is. (June 4, 52-55)

Joan and Jamie agreed.

Joan: Like I said before, those cardiac monitors, they pick up a heart rate quite well. If a heart rate becomes irregular and quite irregular, it will bounce all over the place. But the thing is we'll see the rate because we'll see a rhythm on the monitor. The number, it can bounce all over the place, but we will actually be able to see a complex pattern: the heart rhythm on the screen. And that's what we will look at more or less, and we'll take a manual to get an actual number, an actual reading. (June 18, 109-114)

Jamie: You can visually see what the heart rate is, and if they've even *got* a heart rate. (June 24, 74)

However, Jamie's view was tempered by his realization of limitations in the detail available for clinical reasoning.

Jamie: Cardiac monitors here aren't that accurate, because we're not really a CCU or ICU, so we're not fully trained at assessing these monitors correctly. The monitors here are not really that accurate because they're old and they don't give us a very good picture.

Glen: They would be able to show you what the rate is, so they're accurate enough to show you a dramatic change.

Jamie: Yes, but they're not accurate enough to show us some of the things that are going on within the heart, and some of the functions of the heart. (June 24, 87-93)

...

Jamie: The monitors are usually accurate to distinguish what rate the heart is going at. They're usually pretty accurate at that.

Glen: That's accurate, and you wouldn't have to do a manual under those circumstances?

Jamie: I would just check, just quickly. Because it only takes a second to put a finger on the pulse and make sure it is going at that rate. But I wouldn't do it every 15 minutes. I probably would check just once or twice to be sure, and then I wouldn't check after that. (96-101)

In this context, Jamie, too, appeared to conceive of accuracy in terms of the specific detail provided by an instrument.

Limitations to the accuracy of detail from heart monitor readings came from the Unit Manager's caution:

UM: ... the monitor could be displaying a rhythm and the patient could be dead. That is a possibility. Mind you, it's not probable. If the nurse doesn't realize that all that monitors do is show electrical activity, then we have a basic problem of not really knowing scientifically how that machine works. (October 14, 94-97)

Terry, along with the other nurses, believed a manual reading of blood pressure was more accurate than a Dynamap reading.

Terry: ... because it [Dynamap measurement of blood pressure] was an ambiguous reading or a reading that was high, we wanted to get a very accurate reading. So then we go to a manual blood pressure reading with a BP cuff and stethoscope. (June 8, 58-60)

His belief was predicated on the unreliability of Dynamap blood pressure readings, discussed above in the subsection "Reliability."

As previously mentioned, nurses tended to believe that instruments had no inherent measurement error, and hence when a patient's condition remained static, consecutive readings would be the same because the instrument was so accurate; for example, Joan's discussion about a Glucometer (quoted above) in which she concluded, "But it's a very accurate test" (June 11, 72).

A topic directly related to accuracy is how well an instrument is functioning. Normally this quality is assured by a routine calibration of an instrument. This process entails using concepts of evidence such as (Gott

et al., 2003): end points, intervening points, zero point, and scales. However, instrument calibration is not within the jurisdiction of nursing.

Glen: Again, I want to learn a little bit more about the monitor as an evidence-gathering device. Whose responsibility is it to make sure that the monitor is calibrated so when it says 180, it's really 180. Because, that's not your job, is it?

Gia: No. And you know, I don't know what the routine is; if they have to be calibrated every year or every two years. But we have a clinical engineering faculty on staff in every hospital and they take care of repairs and all that kind of maintenance. (June, 4, 80-85)

As a consequence, the concept of evidence called "validity" is primarily an engineering responsibility in a hospital, except in the cases where nurses are cognizant of variables (i.e. specific contexts) that would jeopardize an instrument's accuracy (e.g. yellow fingers caused by smoking interfere with a patient's sats reading).

Other Concepts of Evidence

The compendium of concepts of evidence proposed by Gott et al. (2003) includes a substantial number of entries that clearly lie outside the purview of nursing. For instance, nurses do not concern themselves with: instrument calibration, the scales that underlie instruments, sampling, statistical treatment of data, and many topics related to the design of experimental investigations (for ethical reasons). On the other hand, concepts of evidence applicable to nursing included reliability, validity, data presentation, and relevant societal aspects (e.g. credibility of evidence, practicality of consequences, power structures, and acceptability of consequences), as illustrated in earlier sections of this report.

One question concerning Gott et al.'s compendium remains: Do nurses use concepts of evidence not found in the compendium of concepts of evidence? Three were noted earlier: a normalcy range for a patient, the unique individuality of the patient (the object of measurement), and the variability within a patient's complex body (Terry's "And what else?"). However, another very different type of concept of evidence emerged from the nurses' transcripts.

In the science-rich workplaces studied by Duggan and Gott (2002), people measured and assessed physical attributes of various entities. In the present study, however, people measured and assessed both physical and *emotional* attributes of patients. The nurses' transcripts clearly indicated that human emotions defined an important subset to concepts of evidence not found in Gott et al.'s compendium, a subset related to such fields as psychology, sociology, and anthropology. This new subset is acknowledged here because of its

role in the science-rich workplace of the surgical ward, but a detailed explication of specific, emotion-related, concepts of evidence requires further investigation beyond the parameters of the present study.

In several events recounted by nurses, emotion-related observations were assessed as evidence from which to make clinical decisions on how to improve the condition and comfort of a patient.

Chloe: Evidence [concerning the emotional state of patients] is often based on your own observations of people and what you're told in the reports. (June 1, 7-8)

A patient's improvement may be influenced by the interaction between the patient and their visiting relatives or friends. A nurse must therefore attend to these visitors to benefit the recovery of a patient. In one encounter, Chloe found herself dealing with extremely stressed relatives. The patient had undergone an amputation the night before.

Chloe: Over the course of the day the family became increasingly anxious to the point of being abusive and obstreperous, according to the night nurse's report. I guess they had been up for three nights in a row by then (an elderly wife and three daughters from ...). So there were some notes (made and recorded) of some conversations that had taken place between the emergency room staff and the family, and then between the RNs up here and the family. (June 1, 12-17)

...

The night before they had been very aggressive and abusive to the point that the surgeon had actually said to us that if these things happened again, call security and have them removed from the hospital. (50-52)

She described her first encounter as follows:

Chloe: When I entered his room in the morning, the patient was very comfortable and there was his elderly wife who just looked absolutely shattered. So I started to do my assessments and introduce myself to her. It was a situation in which I tried to make assessments of the patient and ask him about his level of pain or level of well being, but she would start to talk and answer for him. He was a perfectly lucid man. (June 1, 22-26)

The capability to notice symptoms for someone being emotionally shattered, as the wife was, required "watching for body language a lot" (Chloe, June 1, 81).

In order to collect more emotion-related data, Chloe needed to interact with the wife in a way that would help Chloe's patient heal. Some of Chloe's procedural and declarative understanding involved in this encounter seems best described as "intuition" or "intuitive knowledge," a component of clinical reasoning (Higgs & Jones, 2002, p. 7). The context for her interaction with the wife was the clinical decision to "establish a relationship with his wife by asking her some questions" (June 1, 28-29). Chloe's procedural capabilities were guided by a

subset of concepts of evidence related to the domains of psychology, sociology, and/or anthropology. As a result, the patient's wife was successfully cared for by Chloe over the next several hours and the wife did not impede the patient's recovery.

Chloe: They [family members] haven't been admitted and they don't have their name above a bed. But, they're just as important for the sake of the well being and recovery of the one in the bed. (June 1, 91-93)

A similar scenario unfolded when the daughters turned up later that day and became verbally upset over their father's Foley catheter (June 1, 50ff). The scenario eventually ended by all the relatives expressing their heart-felt appreciation and confidence in the hospital (73-78). In summary, Chloe collected data (body language mostly), she processed data, she collected new data by taking the nurse-visitor interaction to the next level, she enacted several interventions, and she monitored the results by collecting on-going data related to the emotional well being of the patient's relatives.

Although a patient's emotional well being is constantly on the mind of a nurse who focuses on the patient's physical attributes, sometimes the focus does shift to the patient's emotional attributes, as it did for Sarah (June 23). Her patient was found wandering around the ward unsafely at 7:30 a.m., a time when most patients continue to rest. Sarah apprised herself of his physical attributes (e.g. his old age, his low potassium level of 3.1 measured the day before for which he was being bolused, and his meds). However, Sarah attended to his psychological attributes, for instance, his tone of voice, his body language (e.g. no eye contact, pacing around, and bags under his eyes), and to his social behaviour (e.g. irrational conversations), all of which served as evidence for the conclusion that he was confused. The intervention Sarah initiated was not a physical one so much as a purely emotional one. She engaged him in an authentic human-to-human conversation, rather than in a professional nurse-to-patient conversation.

Sarah: I think that maybe he needed someone to talk to because his family hadn't been in for a couple of days. We're often so busy, we just run in and out [of a patient's room], not having time to just talk. (June 23, 96-98)

When Sarah focused on the patient's emotional attributes, she collected data (feedback) and assessed those data with concepts of evidence strictly related to the patient's emotional well being. Within a short time, the patient calmed down and began to rest comfortably. It was later during the same shift that his confusion was attributed to a physical set of circumstances unknown to the nurses that morning.

A different type of emotion-related event was called a "PR situation" (public relations) by Chloe (June 4) when she spoke about two similar incidents that happened simultaneously on her night shift after visiting

hours. In each case, a life partner (spouse) wanted to stay the night and comfort the patient by holding them in their arms, which can only be done when both are in the same bed together. The benefit to the patient had to be weighed against the possible disruption to the other patients sharing the 4-bed wardroom. The appropriate initial intervention (i.e. to request the visitor partner to leave the ward) needed to be achieved in a way that made the visitor feel supported so there would be no detrimental effect on the emotional well being of the patient. This night shift event became much more socially charged due to the fact that one of the visitor-patient couples was a same-sex pair who initially challenged Chloe's request by citing discrimination, being unaware of the identical intervention with an opposite-sex pair on the same ward. Chloe eventually solved both problems with each of the visitors by recognizing their basic concern for the patient and by proving a credible alternative to their staying the night. Success can be credited to her procedural understanding, her emotion-related concepts of evidence, and her intuition.

In this situation, *sensitivity* is a necessary quality in a nurse, but the word "sensitivity" has a much different meaning in this context than it has in Gott and colleague's (2003) catalogue of concepts of evidence: "The sensitivity of an instrument is a measure of the amount of error inherent in the instrument itself" (section 4.5). Emotional sensitivity and instrument sensitivity are two very different concepts of evidence; reflecting the difference between emotional attributes and physical attributes considered by nurses in their evidence-based clinical reasoning.

In addition to exploring emotional sensitivity, future research into emotion-related concepts of evidence in nursing may want to investigate the roles played by such concepts as empathy, equity, and respect, and to explore nurses' "aesthetic perception of significant human experiences" (Higgs & Jones, 2002, p. 27).

Scientific Knowledge-in-Use

Earlier in this report the following points were made: (1) research strongly suggests that most scientific understanding required in a science-rich workplace is learned on the job (Chin et al., in press; Coles, 1997; Lottero-Perdue & Brickhouse, 2002); (2) a pragmatic distinction can be made between scientific ideas and professional knowledge of nursing, on the basis of generalizable decontextualized knowledge versus transferable contextualized knowledge, respectively, although the distinction may be vague in some specific instances; and (3) the context of nursing predictably predisposes a nurse to drawing upon professional knowledge rather than scientific knowledge.

This prediction for nurses is supported by extensive research into the use of scientific knowledge in everyday science-related problem solving and decision making (Davidson & Schibeci, 2000; Dori & Tal, 2000;

Goshorn, 1996; Lambert & Rose, 1990; Macgill, 1987; Michael, 1992; Tytler, Duggan & Gott, 2001; Wynne, 1991). Thirty-one different case studies of this type of research were reviewed by Ryder (2001) who firmly concluded: When people need to communicate with experts and/or take action, they usually learn the scientific knowledge *as required*. The qualification “as required” needs clarification.

Even though people seem to learn science in their everyday world as required, this learning is not often the “pure science” (canonical content) transmitted by university science courses. Research into the application of scientific knowledge to everyday events has produced one clear and consistent finding: *most often, canonical scientific knowledge is not directly useable in science-related everyday situations*, for various reasons (Cajas, 1998; Furnham, 1992; Jenkins, 1992; Layton, 1991; Layton, Jenkins, Macgill & Davey, 1993; Ryder, 2001; Solomon, 1984; Wynne, 1991). For instance, when investigating an everyday event for which canonical science content was directly relevant, Lawrenz and Gray (1995) found that science teachers with science degrees did not use scientific knowledge to make meaning out of the event, but instead used other content knowledge such as values.

This research result, along with the 31 cases reviewed by Ryder (2001), can be explained by the discovery that canonical science knowledge must be *transformed* (i.e. deconstructed and then reconstructed according to the idiosyncratic demands of the context) into knowledge very different in character from the “pure science” knowledge of university science courses (Jenkins, 1992, 2002; Layton, 1991), as one moves from “pure science” for explaining or describing, to “practical science” for action (e.g. professional knowledge of nursing). Most nurses would face a formidable task if they were required, in addition to all their other demands, to deconstruct abstract scientific concepts and reconstruct them to fit the demands of an idiosyncratic event on a surgical ward.

Thus, empirical evidence contradicts scientists’ and science teachers’ hypothetical claims that science is directly applicable to one’s everyday life. What scientists and science teachers probably mean is that scientific concepts can be used to abstract meaning from an everyday event. The fact that this type of intellectual abstraction is only relevant to those who enjoy explaining everyday experiences this way (i.e. those who have a worldview that harmonizes with a worldview endemic to science; Cobern & Aikenhead, 1998) suggests that scientific explanations may very well be seen as irrelevant by those who do not naturally explain their everyday world in scientific ways.

How do nurses tend to think? Cobern’s (1991) in-depth research with 20 nursing students taking a university science course showed that most of the students did not share the materialistic and reductionistic worldview towards nature as their science instructor did, but instead held an aesthetic (beauty and design) or

experiential worldview towards nature. Several students did not even connect science with knowledge of the natural world. Only a small minority of students conceived of nature in a scientific way (i.e. those who had a worldview that harmonized with a worldview endemic to science).

In the present study, the transcripts of five nurses were almost devoid of references to scientific knowledge (except for the use of anatomical terms, an issue discussed below). As noted in the “Limitations” subsection in this report, one cannot be certain whether the nurses spoke in a lay genre to me as an outsider, or spoke in a professional genre to me as a science person. My interpretation of the interviews favoured the latter state of affairs.

The transcripts of one nurse, Terry, were replete with descriptions and explanations from a scientific worldview perspective. In the following exchange, Terry made his viewpoint very clear when he stated, “You really have to understand the physics of what’s going on with those chest tubes.”

Terry: ... And that’s also monitoring what’s happening with those chest tubes.

Glen: That’s when the evidence comes in.

Terry: Oh, absolutely. And for that *you really have to understand the physics of what’s going on with those chest tubes*. You have to understand why those chest tubes are there in the first place. Chest tubes are put in for two major reasons: either a haemothorax (“haemo” meaning blood, “thorax” meaning thoracic cavity) or pneumothorax (air in thoracic cavity). Then you have an open or closed haemothorax or pneumothorax. And “open” means it is open to the extreme environment through a hole through the rib cage through the intercostal spaces between the ribs. ... (June 15, 37-44, emphasis added)

Gia (June 20), on the other hand, described how she successfully solved a chest tube problem but her account was formulated on commonsense professional knowledge in nursing (i.e. what patients do when they pull their chest tube equipment along to the bathroom) rather than a scientific explanation of differential gas pressures in closed or open systems. This is not to say that Gia could not describe how a scientist would explain her patient’s situation (she was not asked for that information), but rather, a scientific explanation for her was not relevant to the problem-solving task at hand. Gia represents the large majority of nurses in Cobern’s (1991) study.

Terry’s scientific worldview descriptions and explanations included: (1) conceiving blood pressure in terms of a hydraulic closed-system in which the heart was the pump, leading deductively to systolic and diastolic blood pressures (June 8, 16-26); (2) conceiving BP cuff size in terms of surface area (June 8, 174); (3) conceiving the act of breathing, in part, as differential air pressure in an open system (June 15, 39-58); (4) conceiving of pain in terms of mechanistic features of the sympathetic and parasympathetic nervous systems (June 20, 47ff), (5) conceiving of an edematous patient in terms of a series of closed systems within the body

(June 25, 89-105); and (6) conceived the alveoli as a place for “the oxygen and carbon dioxide to exchange through osmosis” (June 25, 60). Although almost every event discussed by Terry was communicated within a scientific frame of reference (genre), he also drew upon professional knowledge of nursing as did his peers; for example, citing the empirical relationships between pain and blood pressure (June 20, 46), and between breathing/coughing and bringing a patient’s temperature down (June 20, 169).

My interpretation of Terry’s claim that a nurse needs to understand science so “you know if something is going wrong” (June 15, 70) is that *Terry* needs to understand science because he likely explains nature from a scientific worldview perspective. Because I can share his perspective with him, communication between us was effective. The use of scientific knowledge in nursing may very well be to facilitate communication among professionals who happen to share a scientific worldview. This use represents a very limited view of the application of scientific knowledge to nursing, restricting it to a small minority of nurses.

The other five nurses appeared to engage in clinical reasoning without expressing a need to draw upon scientific descriptions and explanations. Only three short exceptions occurred during their 20 interviews: Sarah’s (June 16) mechanistic description of the role of haemoglobin in the body, Gia’s (June 4) explanation for the beta-blocking effect of Metoprolol, and Joan’s explanation of how she solved a discrepancy (a technical problem associated with a patient’s medication):

Joan: It [the patient’s reaction] was not working as well as I thought it should, whereas it normally had worked in other situations. So when I looked at other things [information about the patient], I noticed that in his other medications he was taking a beta-blocker. Combivent works on a beta-receptor in the system. He was on a medication for his heart as a beta-blocker. So you’re blocking the beta-receptor when this medication works on the beta-receptor and so it was possibly not working for that reason. This is what I figured out. (May 15, 24-29)

On the other hand, when Gia described the event in which a patient’s nervous system reacted negatively to the medication Indocid, Gia mentioned that she would store her newly discovered empirical relationship “in the vault for future reference” (May 22, 45), which I interpreted as a reference to her professional knowledge of nursing. At that moment in the interview, I steered the conversation towards the topic of scientific knowledge.

Glen: Were you at all curious about the actual mechanism that explains how that medication works in the body? Or why the central nervous system seems to close down to some extent?

Gia: I never had time to look it up, but that would be interesting.

Glen: Does that affect how you would observe things?

Gia: Probably, I think you will have a more in-depth *understanding* of it, so you could probably recognize other signs and symptoms of an Indocid reaction. So, probably that will be helpful if I got into it deeper and actually knew the mechanism. But I don't at this time.

It seemed as if a scientific explanation would hypothetically have had potential value, but at that moment, it was not salient to the clinical reasoning in which Gia was engaged. She was perfectly capable of learning the scientific mechanistic explanation but it did not seem particularly relevant in this context. One can only speculate that her worldview perspective was not a scientific one as Terry's seemed to be. However, it was beyond the scope of this research study to inquire into the worldviews or self-identities of the nurses.

Although all six nurses made use of anatomical terms (appropriate noun labels) as they talked about events, this ability to apply scientific/nursing vocabulary is not considered in this study to be a demonstration of understanding scientific knowledge. Instead it is taken as evidence for the *procedural* capability to communicate unambiguously with other health professionals (i.e. to participate in the culture of nursing). However, perhaps the use of anatomical terms by nurses is one of those vague areas between two categories in Figure 1: "scientific knowledge" and "procedural capability."

Another pattern emerged from the interview data. Successful clinical reasoning did not necessarily draw directly upon scientific knowledge. Whenever a nurse mentioned a numerical observation (e.g. a blood pressure of 140 over 80), the numerals never had units associated with the measurement. Units of measurement were not apparently relevant to the nurses' clinical reasoning. (It is important to note in passing that in the 24 interviews conducted, no nurse happened to describe an event that focused on a nurse measuring out a medication for a patient, an event that would certainly have involved the proper units along with the quantitative amount of medication. Measuring medications was not a topic that arose.) All nurses could identify some of the measurement units when I specifically asked what they were. Only Terry remembered all but one of the units of measurement when asked what they were. It happened that no nurse could identify the units for a haemoglobin count. On a surgical ward, measurement units likely get in the way of efficient data management and communication. Perhaps protocol does not include measurement units in this context because the units do not change. As discussed earlier in this report, more important than the units themselves are: a measurement's relation to what is normal or what is critically abnormal, its relation to the context, and its relation to a pattern of data. In contrast to the "unitless" measurements used in clinical reasoning by nurses, units of measurement are central to scientific thinking. Thus, clinical reasoning in nursing and scientific reasoning differ in this regard.

The results from this study support earlier research that questioned the applicability of scientific knowledge to a nurse's knowledge-in-use, except for the rare nurse who happens to have a worldview in

harmony with the worldview underlying scientific knowledge. However, the preliminary nature of this small study points to the need to investigate this issue further with a greater number of nurses in more diverse roles (e.g., in other hospital units, in community clinics, and in homecare units). The evidence to date certainly supports the claim that the context of nursing predisposes most nurses to draw upon the professional knowledge of nursing rather than upon scientific knowledge, when engaged in clinical reasoning.

An alternative interpretation arises from a perspective on professional knowledge of nursing that does not partition it from scientific knowledge because both are evidence-based practices.

UM: I look at evidence-based practice as something that becomes part of you. So maybe I can't remember all the scientific ideas (e.g. the loop of Henley), but I still know that a water pill does its job. After thinking about it from *that* point of view, I think the nurses have assimilated the scientific principles that they learned, and they've taken the common denominator of: "This is how I understand that this is your water pill." (October 14, 8-12)

...

If I'm talking to a patient I might not say, "This is your Furosenide pill." I'm more apt to say, "This is your water pill." I'm not going to go into all the intricacies of how that works (the sodium/potassium pump, etc.), I'm probably just going to say, "It takes the water off, so it alleviates fluid on your heart and helps take fluid off your feet." I think sometimes when we're responding to the public, we don't come across as being scientific experts. But I think to really fully understand what we do, there has to be some grounding there somewhere. (14-20)

If one clearly conceives professional knowledge of nursing and scientific knowledge as two different systems of thought, a problem arises:

UM: But it's like it almost changes the discourse. It no longer becomes a discussion about scientific principles as much as it actually becomes a system all by itself; nursing, if you will; whereby it uses all the principles from other disciplines but has developed most [principles] around science. (October 14, 32-35)

The issue raised here (i.e. whether or not to partition professional knowledge of nursing from scientific knowledge) is reminiscent of the "science versus technology (engineering)" issue debated in science education during the 1970s and 1980s (Gardner, 1994; Layton, 1991). Today in academia, science and technology are generally conceptualized as two distinct ways of knowing, even though they interact and borrow from each other extensively (Collingridge, 1989) and can be indistinguishable in certain R&D projects (Jenkins, 2002).

The Canadian public, hospital patients included, do not generally distinguish between science and technology (Ryan & Aikenhead, 1992), and the public tend to confer prestige and expertise on scientific discourse and methods. Therefore, the *science* of nursing has a crucial role to play along side the *art* of nursing,

in the public forum. In the science education research community, however, professional knowledge of nursing is distinguished from scientific knowledge just as engineering is distinguished from science. This distinction does not in the least denigrate the intellectual expertise required of nurses; the distinction only acknowledges key differences, a perspective that has implications for curriculum development, not for public confidence in nursing.

Summary

The research question that guided this investigation asked, “While taking note of the specific declarative knowledge used by acute-care nurses in a hospital (knowledge-in-use associated with the technical field of nursing and the abstract field of science), is there a core set of concepts of evidence that can be identified?” There was a core set of concepts of evidence that appeared to be shared by all six nurses on the surgical ward.

Concepts of evidence related to reliability were: measurement triangulation, normalcy range, uniqueness of the patient measured, and variability within the patient’s array of physical attributes. On the other hand, the nurses seldom had reason to draw upon the following key concepts of evidence: repeated readings with the same instrument, measurement error, and multiple observers.

The nurses’ concepts of evidence related to validity centred on: accuracy; validation triangulation; and a general predilection for direct, sensual, personal access to a phenomenon over indirect, machine-managed access. The concept of evidence called “data presentation” (Gott et al., 2003, 16.0) surfaced during the interviews as well, when nurses spoke about graphing data to detect trends.

The above concepts of evidence have a common characteristic: they all deal with physical attributes of patients. Missing from Gott et al.’s (2003) compendium of concepts of evidence, but apparent in the surgical unit, is a set of emotion-related concepts of evidence related to psychology, sociology, and anthropology (e.g. cultural sensitivity). This is an area for future research perhaps.

The nurses’ concepts of evidence functioned within two interrelated contexts: (1) taking it to the next level, and (2) initiating a procedure or intervention. Both contexts exist for the prime purpose of healing patients (bounded by the realities of available time, resources, and interactions with other professionals in a hospital). Before engaging in either of these two types of processes (i.e. taking it to the next level and initiating a procedure or intervention), nurses considered the credibility of their observations, which they tended to evaluate as being either sufficient or insufficient. The parallel distinction is made in Gott et al.’s (1999) model (Figure 2) between data and evidence. The surgical nurses in this study appeared to evaluate their data in three different

ways. Data (readings, symptoms, measurements, or observations) became evidence when: (1) a datum was corroborated by other data, (2) trends in the data were perceived, and (3) there was a consistency or inconsistency between a datum and its context. In some instances, these three ways worked in various combinations to confer the status of evidence on data. Figure 2 captures the dynamic nature of nurses' measurements, data, and evidence, contextualized by the social functions and moral consequences that subsequent clinical action might bring (the outer ring of the model).

What conceptual content in physics has a direct role in nursing, given the abundance of instruments and physical procedures utilized by nurses? The answers "Some" and "None" are both correct but depend on the worldview of an individual nurse. The perspectives embraced by Terry and the Unit Manager, for instance, indicate *some* role for physics content, even if that role serves as communication invoking universal abstractions from physics, or as a source for the assimilation of physics principles. On the other hand, for the large majority of the nurses in this study (a proportion consistent the research literature reviewed above), "None" seems to be the evidence-based answer. A knowledge of physics may enhance communication among some nurses, but clinical reasoning appears to draw heavily or exclusively on the professional knowledge of nursing. Some current nursing content may very well be earlier deconstructions and reconstructions of physics content in a context of specific interest to most nurses, but unrecognizable in its present form as physics content to a purist in physics. The *technical professional* deconstruction/reconstruction of that physics knowledge may be relevant to nurses' clinical reasoning, but the original physics knowledge itself is not.

Acknowledgement

I appreciably acknowledge the cooperation and support of the Research Services Unit of the Saskatoon Health Region, and especially the surgical ward Unit Manager and the six nurses who generously gave their time and expertise to make the project a success.

References

- Cajas, F. (1998). Using out-of-school experience in science lessons: An impossible task? *International Journal of Science Education*, 20, 623-625.
- Chin, P., Munby, H., Hutchinson, N.L., Taylor, J., & Clark, F. (in press). Where's the science?: Understanding the form and function of workplace science. In E. Scanlon, P. Murphy, J. Thomas, & E. Whitelegg (Eds.), *Reconsidering science learning*. London: Routledge.
- Cobern, W.W. (1991, April). *The natural world as understood by selected college students: A world view methodological exploration*. A paper presented at the 64th annual meeting of the National Association for Research in Science Teaching, The Abbey at lake Geneva, Wisconsin.
- Cobern, W.W., & Aikenhead, G.S. (1998). Cultural aspects of learning science. In B.J. Fraser & K.G. Tobin (Eds.), *International handbook of science education*. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 39-52.
- Cole, S. (1992). *Making science: Between nature and society*. Cambridge, MA: Harvard University Press.
- Coles, M. (1997). What does industry want from science education? In K. Calhoun, R. Panwar & S. Shrum (Eds.), *Proceedings of the 8th symposium of IOSTE*. Vol. 1. Edmonton, Canada: Faculty of Education, University of Alberta, pp. 292-300.
- Collingridge, D. (1989). Incremental decision making in technological innovations: What role for science: *Science, Technology, & Human Values*, 14, 141-162.
- Davidson, A., & Schibeci, R. (2000). The consensus conference as a mechanism for community responsive technology policy. In R.T. Cross & P.J. Fensham (Eds.), *Science and the citizen for educators and the public*. Melbourne: Arena Publications, pp. 47-59.
- Dori, Y.J., & Tal, R.T. (2000). Formal and informal collaborative projects: Engaging in industry with environmental awareness. *Science Education*, 84, 95-113.
- Duggan, S., & Gott, R. (2002). What sort of science education do we really need? *International Journal of Science Education*, 24, 661-679.
- Eijkelhof, H.M.C. (1990). *Radiation and risk in physics education*. Utrecht, The Netherlands: University of Utrecht CDβ Press.
- Eijkelhof, H.M.C. (1994). Toward a research base for teaching ionizing radiation in a risk perspective. In J. Solomon & G. Aikenhead (Eds.), *STS education: International perspectives on reform*. New York: Teachers College Press, pp. 205-215.
- Furnham, A. (1992). Lay understanding of science: Young people and adults' ideas of scientific concepts. *Studies in Science Education*, 20, 29-64.
- Gardner, P. (1994). Representations of the relationship between science and technology in the curriculum. *Studies in Science Education*, 24, 1-28.
- Goshorn, K. (1996). Social rationality, risk, and the right to know: Information leveraging with the toxic release inventory. *Public Understanding of Science*, 5, 297-320.

- Gott, R., Duggan, S., & Roberts, R. (1999). *Understanding scientific evidence*. http://www.dur.ac.uk/~ded0www/evidence_main1.htm.
- Gott, R., Duggan, S., & Roberts, R. (2003). *Understanding scientific evidence*. <http://www.dur.ac.uk/~ded0org/Evidence/cofev.htm>.
- Higgs, J., & Jones, M. (2002) *Clinical reasoning in the health professions* (2nd ed.). Boston: Butterworth Heinemann.
- Jenkins, E. (1992). School science education: Towards a reconstruction. *Journal of Curriculum Studies*, 24, 229-246.
- Jenkins, E. (2002). Linking school science education with action. In W-M. Roth & J. Désautels (Eds.), *Science education as/for sociopolitical action*. New York: Peter Lang, pp. 17-34.
- Kuhn, T. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Lambert, H., & Rose, H. (1990, April). *Disembodied knowledge? Making sense of medical knowledge*. A paper presented at the Public Understanding of Science conference, London Science Museum.
- Lawrenz, F., & Gray, B. Investigation of worldview theory in a South African context. *Journal of Research in Science Teaching*, 32, 555-568.
- Layton, D. (1991). Science education and praxis: The relationship of school science to practical action. *Studies in Science Education*, 19, 43-79.
- Layton, D., Jenkins, E., Macgill, S., & Davey, A. (1993). *Inarticulate science? Perspectives on the public understanding of science and some implications for science education*. Driffield, East Yorkshire, UK: Studies in Education.
- Lottero-Perdue, P.S., & Brickhouse, N.W. (2002). Learning on the job: The acquisition of scientific competence. *Science Education*, 86, 756-782.
- Macgill, S. (1987). *The politics of anxiety*. London: Pion.
- Michael, M. (1992). Lay discourses of science, science-in-general, science-in-particular and self. *Science Technology & Human Values*, 17, 313-333.
- Ryan, A.G., & Aikenhead, G.S. (1992). Students' preconceptions about the epistemology of science. *Science Education*, 76, 559-580.
- Ryder, J. (2001). Identifying science understanding for functional scientific literacy. *Studies in Science Education*, 36, 1-42.
- Solomon, J. (1984). Prompts, cues and discrimination: The utilization of two separate knowledge systems. *European Journal of Science Education*, 6, 277-284.
- Tytler, R., Duggan, S., & Gott, R. (2001b). Public participation in an environmental dispute: Implications for science education. *Public Understanding of Science*, 10, 343-364.
- Wynne, B. (1991). Knowledge in context. *Science, Technology & Human Values*, 16, 111-121.

Figure 1. Knowledge-in-Use Held by Acute-Care Nurses for Use in Clinical Reasoning

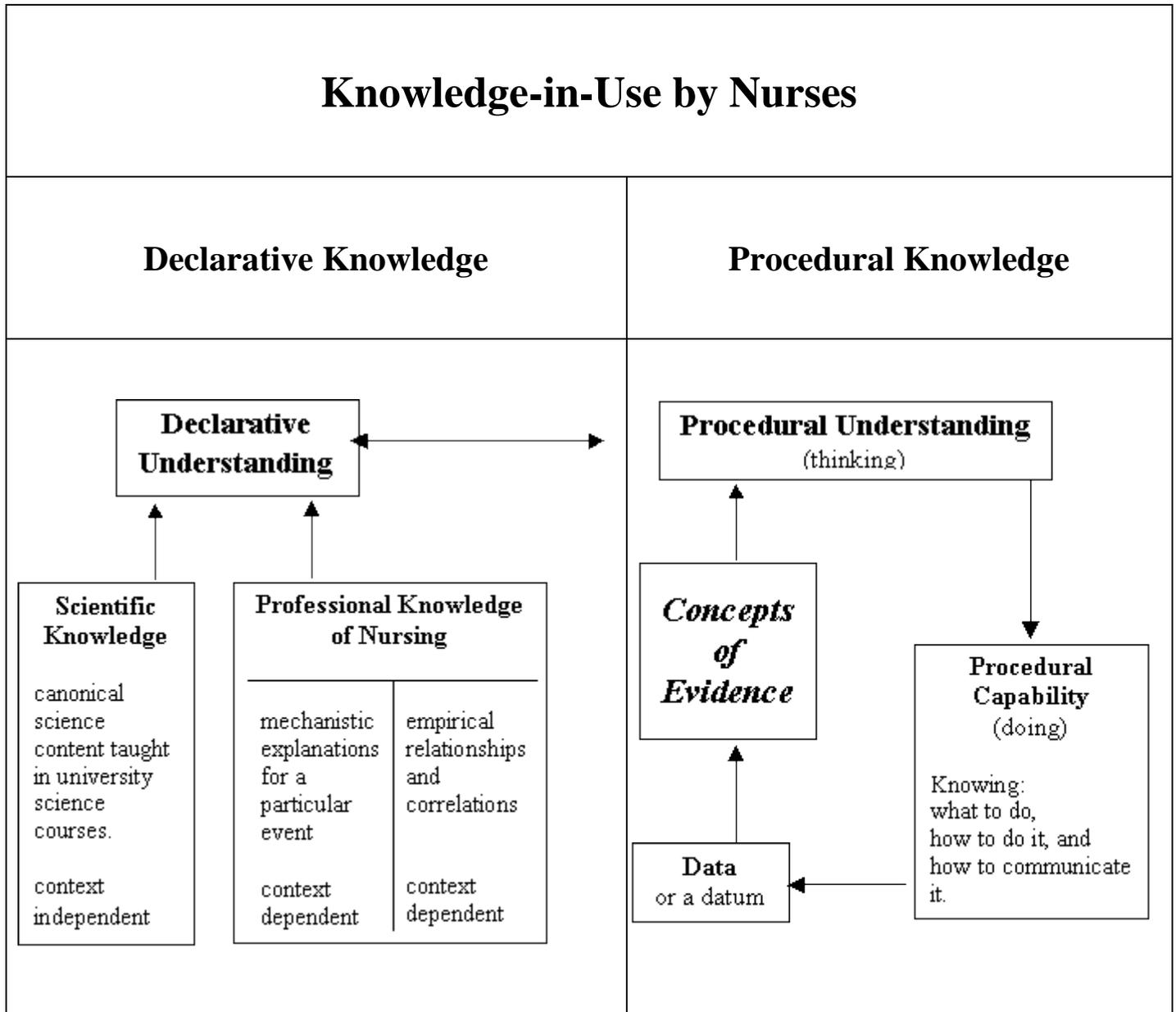
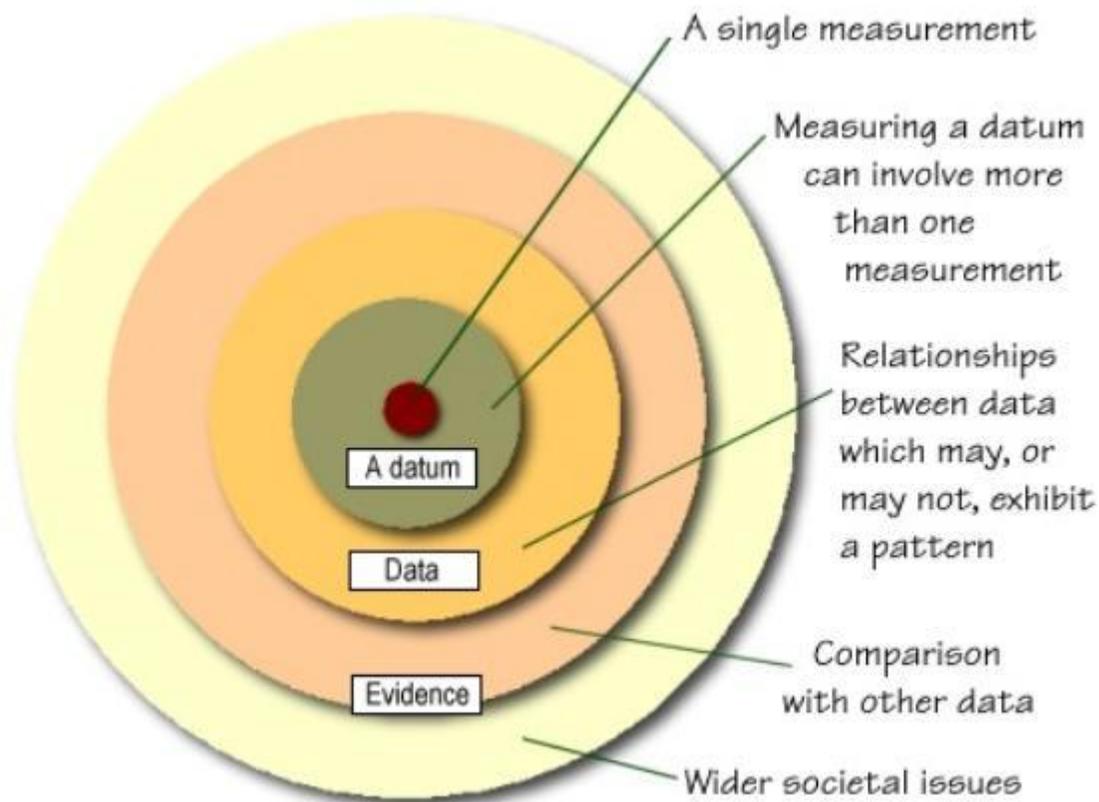


Figure 2. A Model for Measurement, Data, and Evidence



From Gott, Duggan & Roberts (1999).

Figure 3. A Scheme Depicting Different Types of Levels in “Taking It to the Next Level”

<p>1.</p> <p>Symptoms</p>	<p>2.</p> <p>Vitals</p>	<p>3.</p> <p>Targeted Tests</p> <hr/> <p>3.c. residents or doctors</p> <hr/> <p>3.b hospital specialists</p> <hr/> <p>3.a. nurses</p>
---------------------------	-------------------------	---