1. Introduction

Nurses routinely make judgments about the states of their patients and make decisions, choosing among interventions, in the care of their patients. Judgment and decision-making researchers, such as myself, study how these judgments and decisions are made, and when there is a need, we study how the accuracy and efficiency of such judgments and decisions can be improved. From my point of view, a taxonomy of nursing diagnoses is a tool to help nurses make better judgments and decisions. In this editorial I discuss nursing diagnoses as a taxonomic system intended to facilitate cognitive organization of information. I do not take a position, pro or con, on the issue of using nursing diagnoses in practice, but rather my focus is on the qualities that make a taxonomic system useful as aid in learning, memory, and decision-making. First, I will present some historical work from cognitive psychology related to clinical judgment and decision-making in order to provide the reader with a sense of my perspective. Second, I will review some research studies that I believe are representative of those conducted in developing the NANDA taxonomy over the years and I will critically evaluate them from the point of view of a cognitive scientist. The selection of studies for evaluation is, admittedly, biased. My intent is not to provide an exhaustive summary of the literature, but to provide examples that serve to illustrate the state of this research. Finally, I will outline an ambitious recommendation for research aimed at quantifying the quality of any nursing diagnoses taxonomy if it is to function as a tool for improving judgment and decision-making.

Nursing diagnoses categorize patients’ health problems for which nurses provide solutions through nursing interventions (Müller-Staub et al., 2007). According to the North American Nursing Diagnosis Association, nursing diagnoses are intended to provide nurses with the basis for selecting interventions to achieve outcomes for which the nurse is accountable (NANDA International, 2007). Accurate descriptions of patient problems or needs in the form of nursing diagnosis are seen by some as prerequisites in achieving high-quality patient care (Florin et al., 2005). The use of nursing diagnoses has the potential for facilitating comprehensive nursing documentation and helping to increase the efficiency of electronic health data management (Gordon, 2006). It is anticipated that aggregation of nursing data will enable the development of knowledge related to the quality and cost of nursing care and permit comparisons of quality and cost across localities and time periods (Lunney, 2006). Many nurse leaders believe that existing medical diagnostic classification systems, such as the International Classification of Disease (ICD) and Diagnostic and Statistical Manual for Mental Disorders (DSM) are insufficient for these purposes. These beliefs have produced international efforts since 1973 to develop and implement a standard taxonomy of nursing diagnoses.

Of great concern however, studies to date have shown that nurses are not very accurate at applying nursing diagnoses in practice (Hasegawa et al., 2007; see also Lunney, 2001 for review spanning over 30 years). The problem may be that nurses need to be better trained as diagnosticians (Carnevali and Thomas, 1993), however, it is also possible that the taxonomy itself is inadequate for its intended purpose. The organizing normative structure relied upon for labeling health problems of individuals or groups unquestionably affects the diagnosing behavior of the clinicians using it. Applying principles from cognitive continuum theory (discussed below) it is reasonable to expect that alignment of the normative structure of the taxonomy with subjective organizations of knowledge held by practicing nurses will improve diagnostic accuracy.

2. Clinical judgment, learning, memory, and taxonomies

In the seminal work by Elstein et al. (1978) four stages in the diagnostic reasoning process used by physicians were
identified: (1) cue acquisition, (2) hypothesis generation, (3) cue interpretation, and (4) hypothesis evaluation. Subsequent work has confirmed the generality of this process but revealed that accuracy is confined within domains of medical knowledge (see Dowie and Elstien, 1988 for review). Research by nurse scholars indicates that nurses who engage in diagnostic tasks use the same cognitive processes and arrive at conclusions in the same manner (Offredy, 2002, 2005; see also Carnevali and Thomas, 1993 for review). However, while physicians diagnose disease using the ICD or DSM taxonomies, nurses diagnose “human responses to actual and potential health problems/life processes” using the NANDA taxonomy. Therefore, the parameters of the diagnostic task faced by these two types of health care worker are qualitatively different because the taxonomies employed provide different sets of cues to acquire and interpret, as well as different sets of hypotheses to generate and evaluate.

Clinical judgment is often intuitive. To the extent that a taxonomy is intuitive and internally coherent, it is expected that people can learn it and use it to classify patients. Cognitive psychologists have long understood some fundamental aspects of how information is stored and retrieved from memory. Individuals organize information as they encode it into memory, and the greater this organization, the better their recall of the information (see Glass et al., 1979 for review of this early work). Much of this pioneering research in human learning and memory was done using lists of words. This work is relevant because NANDA nursing diagnoses (and classification systems, in general) are here viewed as large, normative organization schemes involving complex concepts expressed with words and intended for memorization and retrieval.

Two long-established principles from cognitive psychology are relevant to the current discussion. First, in a classic experiment by Bower et al. (1969) students were presented with a list of 112 simple words (e.g., foot, hand, piano, tuba, gold, iron) to be learned. Students saw four cards, each containing approximately 28 words. Students studied each word set for about a minute and then tried to recall as many words as possible. The key feature of the experiment was how the words were organized when presented to the students. In one condition, the words were organized into intuitive groupings, in the other condition the categorization of the words was non-intuitive. Students in the intuitive-organization condition recalled about 100 words whereas students in the non-intuitive-organization condition only recalled around 45. The results demonstrate the first principle; by merely manipulating the manner in which material is organized for presentation, learning and memory of the information can be altered substantially.

Second, Tulving (1962) demonstrated that when faced with the task of remembering a random set of words, people attempt to group the words into clusters based on alphabetizing, similarity in sounds, or idiosyncratic associations. What functions as an intuitive organizational scheme for one person may not prove useful for improving the learning and memory of another. “Subjective organization” was the term Tulving assigned to this second principle; when people do not perceive any obvious organization in a set of objects they will try to invent an organizational scheme that can be imposed on the material. This principle implies that the individual nurse, faced with assessing the health care problems and needs of patients, will organize signs and symptoms in order to decide on a best (or at least reasonable) course of action for providing care. Despite standardization in nursing education, nurses may show considerable variation in their knowledge structures of signs and symptoms. Through experience with many patients (i.e., feedback and reinforcement) the nurse refines these knowledge structures to correspond with the reality of his or her clinical practice.

One of the intended functions of the classification system is to assist the nurse in the diagnostic task by providing a rubric for organizing patient signs and symptoms into a diagnosis, and related diagnoses into classes and domains. The two principles imply that to the degree that the NANDA system (or any external classification system for that matter) is intuitive, it is expected to facilitate organization of information in memory. Conversely, when the organizational structure imposed upon the individual nurse is not aligned with his or her subjective organizations, recall of its content and relations will be diminished resulting in poor accuracy in its application.

More recent developments in cognitive theory suggest that alignment of subjective and normative organizations can improve performance. Cognitive continuum theory (CCT) is an adaptive theory of human judgment and posits a continuum of cognitive modes anchored by intuition and analysis (Hammond, 1986; see also Hammond, 1996 for review). CCT also posits a task continuum that is conceptualized as adjacent to the cognitive continuum. The task continuum is a range of different tasks that will benefit from different ratios of intuition and analysis. The theory specifies task features that are likely to induce cognitive modes at different points along the cognitive continuum. By matching the mode of cognition to the features of the task CCT provides a basis for improving the accuracy of clinical reasoning. As Hamm (1988) points out, because the correspondence between cognition and task features partially determines accuracy, CCT offers a helpful perspective for improving accuracy because it can guide the adjustments the clinician might make to his or her cognition or to the clinical task. The most intuitive and least analytical modes of cognition (uncriticized private judgments and group discussions) are the most commonly encountered in the medical clinic (Climo, 1984; Hamm, 1988). Offredy et al. (2008) report that these two highly intuitive modes of cognition dominate clinical reasoning by nurses as well. Based on CCT it seems reasonable to assume that adjusting the taxonomy of nursing diagnoses to be as intuitive as possible will maximize the likelihood that nurses can learn it and apply it accurately in practice.

Please cite this article in press as: Beckstead, J.W., Taxonomies of nursing diagnoses: A psychologist’s view. Int J Nurs Stud (2008), doi:10.1016/j.ijnurstu.2008.08.010
3. The NANDA diagnosis taxonomy: development and status

Müller-Staub et al. (2007) evaluated and compared four of the leading international nursing diagnosis classification systems under development (the International Classification of Nursing Practice, the International Classification of Functioning, the Nursing Diagnostic System of the Center for Nursing Development and Research, and the NANDA system) on three general but fundamental criteria: (1) a diagnosis classification should describe the knowledge base and subject matter for which the nursing profession is responsible, (2) each class fits within a central concept, and (3) each diagnosis possesses a description, diagnostic criteria, and related etiologies. These four systems were examined because of the widespread belief that these systems are comparable and usable for the same purposes. While all four systems had merit, the authors concluded that only the NANDA diagnosis system successfully met all three criteria, noting “NANDA is the only classification studied that defines nursing diagnoses conceptually, that is in terms of responses to health problems/life processes” (p. 711). The current discussion will therefore focus on the NANDA classification system of nursing diagnoses as it seems to have the greatest potential for being adopted worldwide.

The NANDA definition of a nursing diagnosis is: “A clinical judgement about individual, family or community responses to actual and potential health problems/life processes. A nursing diagnosis provides the basis for selection of nursing interventions to achieve outcomes for which the nurse is accountable” (NANDA International, 2007, p. 332). The NANDA taxonomy has been under development since the 1970s. In the 1990s the NANDA Taxonomy I was combined with the American Nursing Association to form NANDA Taxonomy II. Progress and revisions have been chronicled in national conference proceedings published annually. Many papers in these proceedings discussed important issues in taxonomic theory and construction; some of these ideas have been incorporated into the ongoing development of the system. Much of the research reported centers on validation of individual diagnoses using expert panels (an example is summarized below). Current efforts include linking diagnoses to classifications of nursing interventions and outcomes. What is lacking to date, however, is an empirical evaluation of the entire taxonomy as a classification system for facilitating information organization. In its current form, Taxonomy II, the system is described as being multiaxial, employing seven axes. The axes are defined as: Axis 1, the diagnostic concept; Axis 2, subject of diagnosis (individual, family, community); Axis 3, judgment (impaired, effective); Axis 4, location; Axis 5, age; Axis 6, time (chronic, acute, intermittent); Axis 7, status of the diagnosis (actual, risk, wellness, health promotion). Axis 1 is the principal element of the diagnostic statement and is purported to be hierarchical, aggregating some 3000 signs and symptoms into 188 diagnoses grouped thematically into 47 classes arranged within 13 domains.

The taxonomy has been constructed through attempts at consensus among NANDA members in attendance at annual meetings. Many diagnoses have been proposed and research to validate them undertaken, often in the form of dissertations. In general, the work to date has been mixed in terms of methodological rigor (examples are discussed below) with most being of limited if not detrimental value. To illustrate, one of the better researched is the diagnosis of activity intolerance (00092) “a state in which an individual has insufficient physiologic or psychological energy to endure or complete required or desired daily activity”, which was included in Taxonomy I in 1982. Kim et al. (1984) in the first study aimed at validating this diagnosis reported it to be one of the most frequently reported nursing diagnoses for cardiovascular patients. These authors found problems with the identification of defining characteristics including multiple terminologies for the same cue (sign or symptom), lengthy lists of cues with low frequencies of occurrence, and listing the same cue for multiple diagnoses.

Fitzmaurice (1987) attempted to determine nurses’ use of cues (signs and symptoms) in the clinical judgment of activity intolerance using a judgment analysis paradigm. The intent of the study was to address two questions: “How good is the prediction of activity intolerance using defining characteristics of NANDA?” and “How do nurses use and weigh defining characteristics to judge the likelihood of the nursing diagnosis?” A heterogeneous group of 25 expert nurses recruited from three distinct nursing organizations served as subjects. Six defining characteristics from Taxonomy I (fatigue, weakness, dyspnea, discomfort, heart rate, and blood pressure) served as cues and were used to construct 125 patient vignettes. Two defining characteristics (ECG changes for arrhythmias and ischemia) were not investigated “because EEG interpretation requires expertise that some nurses do not have”. This type of statement is troubling because it suggests that the nursing diagnosis taxonomy has been constructed without proper consideration of the knowledge structures and cognitive abilities of practicing nurses.

The paper is fraught with indications that the investigator lacked a fundamental understanding of the quantitative methods she used to form her conclusions; the most glaring of which is identifying cues as “dependent variables”. In a proper judgment analysis study, each subject would evaluate all vignettes and a regression analysis would be performed on each subject’s responses to obtain statistical estimates of the cue weights used by each person (Dowding and Thomson, 2003). In such analyses the multiple R is interpreted a measure of the individual’s cognitive control, or consistency in using information to make judgments. For example, Beckstead and Stamp (2007) reported values ranging from 0.687 to 0.928 in a sample of nurse practitioners judging the likelihood of patients developing heart disease based on eight cues. Individual differences in judgment policies may
also be examined by comparing standardized regression coefficients from different individuals. However, this traditional design and analytic procedure was not used by Fitzmaurice. Instead, each subject rated only 25 vignettes (each vignette was rated by six subjects) and judgments were pooled to conduct a single regression analysis on 750 judgments. The multiple \( R \) was only 0.343, but the author concluded that “the model provided a relatively good estimate of activity intolerance” and that “Approximately 12% of the nurses’ judgments were explained by the characteristics postulated by NANDA” (p. 319). This misapplication of judgment analysis led Fitzmaurice to conclude “Empirical findings do not support the theoretical model of the diagnosis of activity intolerance” (p. 322). Had she designed the study correctly to permit regression analysis at the individual level (and conducted these analyses), she most likely would have come to very different conclusions regarding cognitive control and the relationship of NANDA defining characteristics to diagnostic judgments.

MacLean (1989) attempted to determine the critical cues used for diagnosing activity intolerance by asking nurses to rate cues for their “importance for diagnosing moderate activity intolerance related to an imbalance between oxygen supply and demand” using a 1–7 rating scale. Low importance was defined as ratings of 1 or 2, moderate importance was indicated by ratings of 3 through 5, and high importance defined by values of 6 or 7. Rather than evaluate the utility of the eight characteristics provided by NANDA, the study began by compiling a list of 210 cues obtained from a review of research and practice literature which was “reduced to 127 by dropping redundant, vague, or imprecise cues, and lay terminology” presumably based on the subjective judgment of the author. A small three-round Delphi survey was conducted on a random sample \( (N = 79\) responding to all three rounds) of ANA membership possessing a master’s degree in nursing. The intent was to obtain consensus (defined by 75% agreeing on high, moderate, or low) on the importance of each cue and to select critical cues by retaining only cues with mean importance ratings >5 and inter-quartile ranges of <1. On the first round, nurses reached consensus on only three cues, while rating 86 of the 127 as moderate to highly important. There was no consensus on unimportant cues. In round two respondents were given feedback (mean ratings of the cues) and asked to revise their ratings; consensus was reached on the importance of 14 cues. Only one cue reached consensus as being unimportant. Citing failure to reduce the list to a manageable number of cues, a third round was conducted in which only the 23 cues with ratings >6 were presented. Consensus was reached on 19 of these but respondents indicated that “several were redundant” despite the fact that the author had removed redundant cues from the original list. Only three of the 23 most important cues were reported as being present more than 50% of the time when a patient experienced activity intolerance. All of the eight defining characteristics provided by NANDA were in the final 19; the additional terms being variations on these (e.g., “dyspnea” was replaced with “dyspnea”, “severe dyspnea”, “shortness of breath”, and “labored breathing”). So rather than clarifying which of the eight proposed defining characteristics were actually used by nurses, the study seems to have created more ambiguity by adding redundant terminology.

Unfortunately, these two studies are representative of much of the published work in developing the NANDA taxonomy. They are plagued with problems in conception and execution and there does not seem to be a cohesive progression of empirical validation and extension from one study to the next, but rather a sort of “restarting” of the process with each new undertaking. What it is needed is an empirical assessment of the entire classification system in its current form complete with comprehensive and constructive feedback to the developers and proponents.

Internal coherence. Internal coherence is a necessary property of a hierarchical taxonomic classification system. It is what makes such a system intuitive and helpful for cognitively organizing information. In a taxonomy of diagnoses, the basic unit or defining characteristic is the sign or symptom; a collection of these define a diagnosis. Two diagnoses that have some overlap in their defining characteristics are taxonomically related; two diagnoses that share no common characteristics are not. And for completion, two diagnoses that are comprised of exactly the same set of defining characteristics are taxonomically the same diagnoses.

If a taxonomy is to include two or more unrelated diagnoses, the classification system must have a means of identifying these diagnoses as being unrelated (in terms of overlapping characteristics) yet as being part of the larger system. The term “class” may be used in this regard. Diagnoses that share some of their defining characteristics form a class; diagnoses with no common characteristics are placed into different classes. Problem solved.

So long as diagnoses either do, or do not, share defining characteristics the system of putting them into different classes works. But what are we to do when two diagnoses share more of the same characteristics than do two other diagnoses? The solution is to add another level to our hierarchy to incorporate this finer distinction of relatedness. So we have unique groupings of signs and symptoms that form diagnoses, diagnoses that have “a lot” of overlap in their defining characteristics are placed in the same class, diagnoses that share “a little” of their defining characteristics are placed in different classes but in the same domain, and diagnoses that share no common characteristics are placed in different domains. This type of system is internally coherent and as such can function as an aid in the cognitive organization of diagnoses.

In the official description of the NANDA taxonomy, the developers note, “Some nursing diagnoses lend themselves to placement in more than one domain and class” (NANDA International, 2007, p. 253) implying that the classification system lacks internal coherence. This may...
explain, in part, why nurses have difficulty learning and applying the system.

There is evidence that the normative organization of signs and symptoms into the NANDA classification of nursing diagnoses is not as intuitive to nurses as it could be. Hasegawa et al. (2007) differentiated diagnostic competency into the nurse’s ability to use clinical judgment for determining nursing diagnostic labels and ability to determine the relevant defining characteristics of a patient’s actual state. They report that while nurses were highly competent at determining relevant symptoms, they were poor at determining the correct NANDA diagnostic labels. Myers and Spies (1987) analyzed the frequency with which nurses used NANDA diagnostic labels. Using a content analysis of 328 spontaneously generated labels produced by 54 staff nurses who viewed a video-tape of the same patient care situation, only 23% of the diagnostic labels were reported as being correct. Justice (1987) inspected the cues requested by 42 cardiac nurses to diagnose chest pain. Of note, 122 cues were requested, but only nine by at least 50% of the sample and only 2 of these were related to obtaining the correct nursing diagnosis. Only 21 nurses correctly diagnosed the pain as being of gastroesophageal origin. Castles (1982) found no evidence to support the claim that multiple nurses would arrive at the same nursing diagnosis when using the diagnostic system to assess the same patient at approximately the same time. Taken together these studies suggest that the diagnostic taxonomy provided by the NANDA system may not fully accommodate nurses’ subjective organization of patient signs and symptoms.

4. A constructive recommendation

First, the NANDA nursing diagnosis taxonomy should be analyzed using quantitative methods to determine its internal coherence and to localize sections of relative clarity/ambiguity within its (normative) structural configuration. Second, nurses’ cognitive (intuitive) organization of diagnoses ought to be measured and compared to NANDA’s normative structure. This could include perceived similarity of diagnoses based on their definitions as well as perceptions of the relevance and clinical utility of signs and symptoms used to define diagnoses. The correspondence between NANDA’s vast list of defining characteristics and those that nurses actually find relevant and useful for distinguishing the health care problems and needs of one patient from those of another could then be assessed. Cognitive continuum theory suggests that aligning these two configurations of knowledge will improve learning, utilization and utility of nursing diagnoses in nursing education, practice, and research.

A quantitative analysis of the normative structure of the existing taxonomy would provide valuable insight into its internal coherence. If this is found to be high then our attention should be directed at determining where nurses’ subjective organization of diagnoses deviates from this internally coherent structure and altering nursing education efforts. For example, regions within the taxonomy that show the greatest discrepancy between normative and subjective organization could be identified as priorities for curricular changes in the teaching of nursing diagnoses (both for nursing students and for practicing nurses in continuing education programs).

On the other hand, if the internal coherence of the taxonomy is found lacking, then our attention should be focused on making recommendations to NANDA that will improve utility and utilization by improving the taxonomy’s intuitive appeal to practicing nurses. These may include: deletion of diagnoses formed entirely from irrelevant defining characteristics with very low clinically utility, shortening lists of defining characteristics for some diagnoses, combining of diagnoses which are perceived as virtually identical, and reorganizing/reconceptualizing the hierarchical arrangements of diagnoses based on patterns of objective and subjective similarity.

It may be found that certain classes or an entire domain within the taxonomy are not at all coherent, however, when nurses’ subjective experiences with defining characteristics of the diagnoses in this region of the taxonomy are examined, we find that all (or most) are perceived as irrelevant and as having very low clinical utility. Elimination of these diagnoses would improve the overall coherence of the taxonomy without reducing its correspondence for intended users.

5. Conclusion

From my point of view, a taxonomy of nursing diagnoses is a tool that should help nurses make better judgments and decisions regarding the care of their patients. In the paragraphs above I have discussed some of the qualities that, in my opinion, are key for making such a tool a good one. These included internal coherence, intuitive appeal to its users, and a correspondence to the situations nurses face in practice. I chose to discuss the NANDA taxonomy as it is the one which seems to have the most promise as a useful tool, as concluded by Müller-Staub et al. (2007), and because there is an abundance of published material on its development.

Admittedly, this editorial is an evaluative statement on a state of affairs within nursing from the point of view of an outsider (i.e., a cognitive psychologist who studies judgment and decision-making). By offering this perspective, I hope to broaden constructive discussions among nurse scholars regarding the proper functions of a nursing diagnosis taxonomy. My personal motive for writing this essay is to stimulate interest in judgment and decision-making among nurse researchers. To those who wish to learn more about this fascinating field of study, I recommend the excellent introductory text edited by Thompson and Dowding (2002).

The field of psychology has much to offer nursing. Theories from clinical psychology have influenced many
nurse scholars (see Beckstead and Beckstead, 2006 for discussion). Carl Rogers and Abraham Maslow shared an optimistic view of people as being capable of self-care and self-determination given a secure, nurturing environment; themes that pervade many nursing theories. Cognitive psychologist Kenneth Hammond was one of the first to conduct scientific research aimed at understanding clinical judgments and decisions made by nurses (see Kelly and Hammond, 1964). Among the topics he examined were: types of cognitive tasks nurses faced in practice (Hammond et al., 1966a), information units nurses used in problem-solving (Hammond et al., 1966b), information-seeking strategies nurses used when assessing the state of their patients (Hammond et al., 1966c), and how nurses revised their judgments when presented with new information (Hammond et al., 1967). This line of research has been extended to advanced registered nurse practitioners by Beckstead and Stamp (2007). Psychometric theory and its measurement techniques are invaluable for researchers studying individual differences. Beckstead et al. (2008) discussed the importance of cross-cultural measurement invariance in international nursing research and illustrated a structural equation approach using the Women’s Role Strain Inventory.

Understanding how people use classification taxonomies, in general, is of interest to cognitive psychologists. Understanding the qualities that permit a nursing diagnosis taxonomy to function as an aid in clinical judgment and decision-making ought to be of interest to nurse researchers. Only to the extent that a taxonomy is internally coherent, corresponds well to its intended universe of cases, and is intuitive to its users, can we expect people to apply it successfully. Without these qualities any taxonomy is of limited value. If a taxonomy of nursing diagnoses is to produce the benefits anticipated by its advocates it is imperative that, which ever taxonomic system is eventually endorsed worldwide, it meets these criteria.

Conflict of interest

None declared.

References


Jason W. Beckstead*
University of South Florida, College of Nursing,
Tampa, FL 33612, United States

*Tel.: +1 813 974 7667; fax: +1 813 974 5418
E-mail address: jbeckste@hsc.usf.edu

14 August 2008

Please cite this article in press as: Beckstead, J.W., Taxonomies of nursing diagnoses: A psychologist’s view. Int J Nurs Stud (2008), doi:10.1016/j.ijnurstu.2008.08.010