Measuring Knowledge Utilization in Health Care

Carole A. Estabrooks
Lars Wallin
Margaret Milner

Abstract

In this paper we address the need for methodological advances in the research utilization field focusing on the area of measurement. Unresolved measurement challenges present an important and practical problem. An inability to adequately measure research utilization calls into question studies that claim to demonstrate either its causes or its effects. In this paper we: (1) briefly review the concept of research utilization and its meanings, (2) review the requirements of good measurement instruments, (3) review existing research utilization instruments in nursing (the field with which we are most familiar), and (4) discuss implications and future requirements for scholarship in this field.

Key words: research utilization, measurement, good measurement instruments, nursing

The authors wish to thank Connie Winther for her input on methods and assistance and for reviewing an earlier version of the manuscript. Dr. Estabrooks (post-doc fellow) and Wallin (post doctoral fellow) receive career support from the Canadian Institutes of Health Research (CIHR).

Associate Professor, Faculty of Nursing, University of Alberta (Canada) & Academic Co-Director, Canadian Center for Knowledge Transfer. carole.estabrooks@ualberta.ca

Postdoctoral fellow, Faculty of Nursing, University of Alberta (Canada)

Graduate student, Faculty of Nursing, University of Alberta (Canada)
1. Introduction

A significant attribute of the post-modern world is the reliance on science and research-based knowledge. In an economy increasingly viewed as knowledge based, knowledge is seen as an important social and organizational asset (Blackler, 1995; Brown and Duguid, 1991; David and Foray, 2002; Nahapiet and Ghoshal, 1998). Explaining knowledge use becomes a central question of contemporary scholars from a variety of disciplines.

The tradition of knowledge utilization and related research in the health sciences is among recent efforts to root policy and practice decisions in science. The evidence-based medicine movement (Evidence-Based Medicine Working Group, 1992) emerged in its current form in the early 1990s, although its roots were put down much earlier and quickly evolved into more general calls for the adoption of an evidence-based decision-making culture at all levels of the health care system (Gray, 1995; Evidence-Based Decision Making Working Group, 1997; National Forum on Health, 1997). In the Canadian context, these efforts eventually contributed to pressures to downsizet the welfare state and increase the efficiency and effectiveness of those components that were not eliminated or privatized. Since the creation of the Canadian Health Services Research Foundation in 1997 (http://www.chsrf.ca) and the Canadian Institutes for Health Research in 2000 (http://www.cihr-irsc.gc.ca) the knowledge utilization agenda has accelerated at an ever increasing rate in Canada.

Under the broad umbrella of knowledge utilization, research utilization is related to evidence-based decision-making in the health professions, as well as, to similar concepts in sociology, agriculture and business, such as innovation diffusion (Rogers, 1995). It is also closely connected to the field of evaluation utilization that deals with the performance and outcomes of program evaluation (Johnson, 1998). Knowledge, in our case science, is however not translated to usefulness by itself - complex processes and deliberate efforts are required. We define research utilization as that process by which specific research-based knowledge (science) is implemented in practice.

Research utilization is commonly assumed to have a positive impact on patient outcomes by assisting with eliminating ineffective practices and implementing more effective procedures (Sackett et al., 1998; Grob, 2000). However, this process is not straightforward. Many studies point to our uncertainty about the effect of medical and nursing research on clinical practice (Bero et al., 1998; Oxman et al., 1995; Estabrooks, 1998; Rycroft-Malone et al., 2002). Over the past three decades investigators in nursing have studied research utilization and proposed several factors thought to influence the uptake of research. However, similar to the literature on evaluation utilization (Johnson, 1998), we find a number of weaknesses in this literature. For example, the identification of factors that influence the research utilization process has not been followed by the development of theoretical models that specify relationships among factors or offer clear guidance for the development of implementation strategies. Of concern to us in this paper is the lack of development in the area of measuring research utilization, both generally and in nursing in particular. Whether research utilization is conceived as product (resulting from particular sets of factors) or process (resulting in particular health outcomes) our inability to measure it satisfactorily raises questions about claims of its importance.

Our purpose in this paper is to address the need for methodological advances in the research utilization field focusing on the area of measurement. Unresolved measurement challenges present an important and practical problem. An inability to adequately measure the construct of interest (research utilization) calls into question studies that claim to demonstrate either its causes or its effects. In this paper we: (1) briefly review the concept of research utilization and its meanings, (2) review the requirements of good measurement instruments, (3) review existing research utilization instruments in nursing (the field with which we are most familiar), and (4) discuss implications and future requirements for scholarship in this field.

2. Research Utilization in Health Care

The objective of increasing the use of research in health care is to improve patient outcomes. A society’s values and its financial resources constitute obvious and basic elements of a health care system’s capacity to meet patient needs. There is evidence to suggest that several additional
Factors within the health care system effect outcomes. These include health professionals' educational level, organizational climate, leadership, nurse-patient relationships, staff well-being and quality improvement strategies (Shortell, Bennett, and Byck, 1998; Shortell et al., 2000; Aiken et al., 2002; Aiken et al., 2001; Aiken et al., 2003; Arnetz, 1999). We hypothesize that research utilization is one indicator of an optimum practice environment, an environment that leads to improved patient outcomes. The relative importance of that indicator remains to be evaluated, but we believe it is a factor whose importance will be more apparent as the body of research examining it grows.

At the same time that it is relevant to consider research utilization as one of many factors influencing health outcomes it is also important to identify the constituent elements of the research utilization process itself. Evidence-based medicine has been described as having five steps: (1) converting information needs into answerable questions, (2) tracking down the best evidence for answering the questions, (3) critically appraising that evidence, (4) implementing the results of this appraisal in clinical practice, and (5) evaluating care performance (Sackett et al. 1998). This general model reflects the commonly accepted ingredients of knowledge utilization as it is generally understood among health care providers. Although it can be challenged on, for example, its assumption of linearity in the change process, the model has evolved to where it is the cornerstone of the evidence-based movement in modern health care. The literature on evidence based care, as well as, research utilization focuses heavily on the critical appraisal phase, with the implementation phase being less well understood. Despite the need for expertise and resources in the appraisal process, we and others argue that the major stumbling block to achieving more research based clinical practice lies in the research implementation phase (Bero et al., 1998; Rycroft-Malone et al., 2002; Grimshaw and Russell, 1994; Feder et al., 1999; Thomas et al., 1998). Our knowledge of what constitutes successful research implementation strategies in practice is sparse. In effect, we treat this phase like a black box phenomenon where research implementation is viewed as more art than science – when the opposite is most urgently needed.

3. Types of Research Use

Several conceptualizations of research utilization have been proposed. Larsen (Larsen, 1980, p. 424) proposed that knowledge utilization be classified as instrumental and conceptual. Rich (Rich 1975, Rich 1977) and Weiss (Weiss 1979) also discussed these two kinds of knowledge utilization. Symbolic utilization was added by Beyer and Trice (Beyer and Trice 1982). Steier (1985) introduced instrumental and conceptual use into the nursing literature. Estabrooks (1999b) empirically verified instrumental, conceptual and symbolic research use by nurses.

Instrumental research utilization is that concrete application of research either in making a specific decision or specific interventions such that the research is translated into a material and usable form (e.g., a protocol or guideline). Instrumental use is a direct type of research use. In conceptual research utilization, the research may change one's opinion or mind set about a specific practice area but not necessarily one's particular action. It is a more cognitive application of research and is believed to occur more often in practice than instrumental applications of research. Symbolic research utilization is the use of research knowledge as a political tool in order to influence policies and decisions.

Currently, there is no unified theory of research utilization in nursing or elsewhere. Nor are there established competing research utilization theories suited to local contexts. The relative lack of research utilization theory and the failure of investigators to incorporate theory when it is available into instrument construction has hampered development of good research utilization measures. For example, in nursing while there are many models of research utilization, among them, CURN (Horsley, Crane, and Bingle, 1978), Goode (Goode et al., 1987) Ottawa (Logan and Graham, 1998), Steier (Steier, 1994), NCAST (Nursing Child Assessment Satellite Training) (Barnard and Hoech, 1978), Horn (Goode and Bulechek, 1992), Iowa (Titter et al., 1994), PARHIS (Kitson, Harvey, and McCormack, 1998; Kitson et al. 1996), when we examined research utilization instruments, we found that they draw either on Roger's (Rogers 1995) work or do not identify a theoretical underpinning.

4. Rogers' Innovation Diffusion Theory

Rogers' theory of innovation diffusion is widely used in the nursing
literature as a basis for measuring research utilization. Brett (1987), Winter (1990), Kirchoff (1982) and Lia-Hogberg et al. (1999) developed instruments based on this work. Rogers' model describes a sequence of five stages an individual undergoes in adopting an innovation. He stipulated that an individual would pass through the stages of awareness, persuasion, decision, implementation, and finally confirmation in the adoption process. According to Rogers (1995), in the awareness stage the individual comes to know about the innovation. In the persuasion stage the individual is either favorably or unfavorably disposed to the innovation. Based on this disposition, she engages in activities leading to a choice at the decision stage. The implementation stage follows a decision in favor of the innovation. Implementation involves a change in behavior as the innovation is put into practice. At the confirmation stage the individual reinforces beliefs about the innovation and the decision. Though Rogers' model is widely used as a basis for measuring research utilization in nursing, Estabrooks (1998) has previously argued that a drawback is using Rogers' model is the untested assumption that the innovation in classical diffusion theory is equivalent to research findings in the context of clinical nursing practice. That is, is innovation as product equivalent to knowledge as product?

5. Measurement Theory

A defensible theoretical basis is foundational to the development of measurement tools and to their validity assessment, providing clarity for the construct being measured. There are many theories available to those undertaking instrument development for psychological and social constructs. Among them, the three most commonly implemented are classical test score theory, item response theory, and generalizability theory. Each theory carries its own set of assumptions upon which the measure is premised. Classical test theory (Hambleton and Jones, 1993; Ellis and Mean, 2002) maintains that every measurement is an additive composite of two components — true ability of the respondent on that construct and random error. Classical test theory forms the foundation of reliability theory. Item response theory (Hambleton, 1991; van der Linden and Hambleton, 1997) is a model-based theory that assumes that a response to an item is functionally related to a latent (unobservable) trait, presumably measured by the item. The theory postulates that as an individual's standing on the latent trait increases, the probability of a correct response increases. This is often used when there is a correct, or more desirable answer. Generalizability theory (Brennan, 1983, 2000; Shavelson and Webb, 1991) is an extension of classical test theory as well as an application of some analysis of variance procedures to measurement models involving multiple error sources. It was designed with the intent to remedy what is often considered a major drawback of classical test theory, the assumption of a single source of random error. Additionally, there are a number of other criteria for good measures with which researchers must come to terms, among them:

Minimal Self-Report and Social Desirability Effects. A reliable and valid measure should be able to capture the intended "truth" about the measurement object. To achieve minimal effects of self-report and social desirability it is necessary that test questions are constructed in such a way that they are minimally "value laden." The instrument should have an embedded mechanism to determine levels of social desirability, as well as handling other effects such as acquiescence and a tendency to choose the center of the scale.

Unambiguous Scale. The scale of measurement should not be ambiguous to either respondents and users. The scale should be easily understandable to stakeholders, and the types of inferences made about the results of measurement should be logical and understandable to all.

Acceptability within the research field. The ideal instrument would be reusable and generalizable across settings. The theory of the construct and the way in which the construct is operationalized by the instrument should be widely accepted by the researchers in the field. Ideally, other researchers would adopt the tool so that generalization and cross-study comparison of results across projects would be more compelling.

6. Existing Research Utilization Measures

The need to better understand the problems of implementing research triggers requirements for measuring this implementation or research use. Most investigations in the area have described characteristics that facilitate research use and modeled these as predictors of research use. In the nursing field, several instruments attempting to measure research use have resulted from these efforts.

We created a set of rules for inclusion of studies that narrowed our focus to the actual use of research in practice. Articles that reported the use or development of an instrument in which words or phrases such as
<table>
<thead>
<tr>
<th>Citation</th>
<th>Measuring Knowledge Utilization in Health Care</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table 1</strong> Measures of Research Utilization</td>
<td><strong>Table 1</strong> Measures of Research Utilization</td>
</tr>
<tr>
<td>Citation</td>
<td>Topic</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>1. Hur, H. I. 1997</td>
<td>Use of nursing practice research findings</td>
</tr>
<tr>
<td>2. Hur, H. I. 1998</td>
<td>Organizational antecedent mechanisms and adoption of innovation by nurses</td>
</tr>
<tr>
<td>3. Corin, L. A., &amp; Seoep, A. G. 1999</td>
<td>Innovation adoption behavior among nurses' Nursing</td>
</tr>
<tr>
<td>4. Hur, H. I. 1999</td>
<td>Information seeking, research utilization, and heterogeneous research utilization of pediatric nurses' education</td>
</tr>
<tr>
<td>5. Michel, V. &amp; Smith, W. Y. 1995</td>
<td>Communication and use of research findings among practice nurses</td>
</tr>
<tr>
<td>6. Bruggen, A. C. 1999</td>
<td>Swedish nurses' awareness of, attitudes and use of selected research findings</td>
</tr>
<tr>
<td>7. Katalin, et al. 1996</td>
<td>Overview of Research-Based Practice by Durham Hall Nurses</td>
</tr>
</tbody>
</table>

*The measurement concepts we have found in the research utilizations are shown with questions relating to research utilization in practice. Articles included in the table are those that applied the concept of a specific nursing practice described as an instrument for measuring research utilization. The following concepts are used in conjunction with research utilization: evidence-based practice, evidence-based practice, and evidence-based practice. The table shows a summary of the existing statements in the literature on the use of systematic reviews (Cochrane, et al, 1997; Dobbins, Cocklin, & Barnsley, 2001; Leufkens, Schraff, & Weimann, 2003; Shenton, 1999; Wallin, et al, 2000).
<table>
<thead>
<tr>
<th>Citation</th>
<th>Sample N</th>
<th>Subjects</th>
<th>Framework</th>
<th>Specific AE Measurement</th>
<th>Scoring</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other multi-item instruments [3]</td>
<td>20</td>
<td>Registered nurses</td>
<td>Nursing</td>
<td>6 items</td>
<td>6-point Likert scale: strongly disagree, disagree, agree</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

**Research Utilization Questionnaire (RUQ)**

<table>
<thead>
<tr>
<th>Citation</th>
<th>Sample N</th>
<th>Subjects</th>
<th>Framework</th>
<th>Specific AE Measurement</th>
<th>Scoring</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citation</td>
<td>Sample Size</td>
<td>Subjects</td>
<td>Framework</td>
<td>Specific R1 Measurement</td>
<td>Scoring</td>
<td>Reliability</td>
<td>Validity</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>----------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>21</td>
<td>312</td>
<td>Registered nurses</td>
<td>Not specified</td>
<td>Research Education Questionnaire</td>
<td>Not defined</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>22</td>
<td>400</td>
<td>Registered nurses</td>
<td>Unpublished</td>
<td>The Research Use in Nursing Questionnaire, 1993</td>
<td>In practice and research practice (n = 47)</td>
<td>Unreliability and unreported</td>
<td>Not reported</td>
</tr>
</tbody>
</table>

Single item(s) in questionnaire:
- How often do you use research in your practice (on a scale of 1 to 5, where 1 = never and 5 = always)?

<table>
<thead>
<tr>
<th>Citation</th>
<th>Sample Size</th>
<th>Subjects</th>
<th>Framework</th>
<th>Specific R1 Measurement</th>
<th>Scoring</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>500</td>
<td>Registered nurses</td>
<td>Not specified</td>
<td>Nurse's Research Involvement Survey</td>
<td>Not defined</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>24</td>
<td>600</td>
<td>Registered nurses</td>
<td>American Nurses Association, 1999</td>
<td>Research Involvement Survey</td>
<td>Number of times participated in activity</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>25</td>
<td>200</td>
<td>Occupational nurses</td>
<td>Not specified</td>
<td>Not defined in research survey</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>26</td>
<td>150</td>
<td>All grades of hospital nurses (UK)</td>
<td>Not specified</td>
<td>Two items in questionnaire: one asking nurses to rate the extent of research findings in nursing practice in the last two years, the other asking nurses to rate the frequency with which they use research findings in nursing practice</td>
<td>Not reported</td>
<td>Content validity by expert panel and pilot study</td>
<td></td>
</tr>
</tbody>
</table>

**Measuring Knowledge Utilization in Health Care**
Although we risk missing interesting work on the direct use of research by excluding such papers, we rationalized this decision based on the complexity and scope of nursing guidelines, the significant variance in the rigor with which they are developed, and the evidence base that supports them. We also excluded a number of papers that one might anticipate being included where multi-item instruments were used to measure several factors related to research use, but not research use in itself. These included such factors as individual and organizational determinants, and barriers to research use (Camilleri and Huffman, 1998; Ehrenfeld and Eckertling, 1991; Funk et al., 1991; Adamsen et al., 2003; Carroll et al., 1997; Griffiths et al., 2001; Kamwendo, 2002). Other papers were excluded due to lack of clarity in the method or result sections specifically regarding if and how research use was measured (Curtin and Jaramazovic, 2001; Davies, 1999; Ketefian, 1975; Kirchoff, 1982).

We located three commonly used multi-item instruments published in 20 papers and three multi-item instruments cited in one paper each. Another 20 published papers used single-item questions to measure actual research use. A more detailed description is presented below.

Nurses Practice Questionnaire (NPQ)

Brett (1987) developed the Nurses Practice Questionnaire (NPQ) based on Roger's (1983) innovation diffusion theory. The NPQ consists of brief description of 14 specific nursing practice innovations and it is the most widely used research utilization instrument in the nursing literature. Based on Roger's (1983) stages of innovation adoption, seven questions measuring the nurse's stage of innovation adoption are posed for each of the nursing practice innovations. The questions attempt to examine Roger's five stages of the adoption of a specific practice innovation - level of awareness, persuasion, decision, implementation, and confirmation.

The NPQ and its modifications have been used in ten studies (see Table 1). The reliability assessments reported in all these studies are limited to Cronbach's alpha test of internal consistency and little is reported that suggests validity has been established. A number of assumptions underpin the NPQ. The NPQ assumes research use has a
linear progression through the identified stages and that awareness of the innovation (the research finding) and a positive attitude (persuasion) towards innovation reflect use. The term ‘use’ is not clearly operationalized in the NPQ. In fact, with few exceptions in this body of literature (e.g., Estabrooks, 1999b; Rodgers, 2000), an operational definition of the term research use was not given in the studies.

Finally, the NPQ (as do all instruments we assessed) assumes that implementing research is good despite the fact that many research findings are decontextualized and in some cases impractical in the face of existing resources and particular patient populations.

The Research Utilization Questionnaire (RUQ)

Champion and Leach (1989) developed the Research Utilization Questionnaire (RUQ). This instrument is made up of 42 self-descriptive statements comprising four subscales of which research utilization is one. Research utilization was measured using a summated 5-point Likert scale. The scale comprises 10 items measuring the degree to which a nurse felt she incorporated research findings into practice. Sample items in the instrument include: I base my practice on research, I apply research results to my own practice, I use research to manage my practice.

This instrument and its modifications have been used in seven research utilization studies (see Table 1). Similar to the NPQ, psychometric assessment has been limited to Cronbach’s alpha test of internal consistency and some limited assessments of content validity. “Use” is not well defined and the instrument does not discriminate different forms of research use. We found no evidence that the development of this instrument was based on an established theoretical framework making it difficult to understand how the research utilization process was conceptualized. The items included in this instrument contain value-laden phrasing making it prone to social desirability biases.

Edmonton Research Orientation Survey (EROS)

The Edmonton Research Orientation Survey (EROS) developed by Pain, Hagler, and Warren (1996) is a self-report questionnaire measuring practitioners’ participation in and orientation toward research. It was developed in the context of rehabilitation medicine specialties (e.g., physiotherapy, occupational therapy). We included it because researchers have included nurses in their samples. This instrument has four subscales of which the ‘Using Research/Evidence Based Practice’ is one and is composed of 10 items using a 5-point Likert scale. Examples of items that measure research use from the EROS include: reading the research literature has changed the way I practice, and hearing research presentations has changed the way I practice.

The instrument has been used in three studies. As with the other instruments a drawback of the EROS is the vagueness of the term use. No theoretical framework or operational definitions are reported. While the instrument was tested for construct validity and reports acceptable reliability ratings for the research use subscale, it assumes that involvement in and understanding of research implies a higher level of research use. The influence of social desirability through self-report is again potentially present. Additionally, the questions ask respondents to reflect on their use of research over the past five years, raising questions of recall accuracy. As with the NPQ and RUQ, the EROS does not explicitly integrate measurement theory in its development.

Other Multi-Item Instruments

We located three other multi-item instruments developed to measure research use. We found one report describing each instrument. Beoesser et al. (1994) developed an instrument and called it the Research Utilization Questionnaire (RUQ) based on a critical review of the literature. The instrument was composed of statements and open-ended questions with regard to use of products and methods in relation to research use. Products included searching the literature, engaging others to use research findings, critique of research studies, and the creation of clinical action plans from findings. The underlying assumption is that these products infer research use, which the authors did not define or conceptualize in the study. The instrument also examined the frequency of use of knowledge sources as an indicator of research use, assuming that this variable predicts use without empirical evidence to support such claims. The authors provided no report of validity or reliability testing of this instrument.

Pelz and Horsley (1981) reported results from an intervention study measuring direct and indirect research utilization as part of the Conduct and Utilization of Research in Nursing (CURN) research development project. The instrument was specifically developed for the project where an intervention team was placed in experimental hospitals to determine their effect on the research utilization process. Sample questions (5 items in total) related to frequency of instrumental research use. Examples include: you evaluated a research study to determine its value for practice, and you discontinued or rejected a practice activity because of knowledge included in the results of research studies. Questions that measured research use less directly included the extent to which the respondent did...
back as three years prior to the study, raising questions of recall accuracy.

Common attributes of measures

Across all reviewed studies there was a focus on the measurement of the individual nurses’ use of research. Organizational dimensions of research utilization were not commonly studied. If organizational variables were examined the unit of analysis was the individual suggesting serious unit of analysis errors exist. Also common were retrospective survey designs, heavily dependent on self-reporting and recall. We located few experimental designs and no longitudinal designs. The conceptualization of research use was almost always unclear. Respondents were not given descriptions of the various kinds of research use (i.e., instrumental, conceptual, symbolic), leading to uncertainty in interpreting study outcomes.

A number of assumptions were apparent in these instruments – that research utilization is always good (all research is applicable to practice), that research utilization is linear (a sequential process of dissemination – critical appraisal – implementation), that evidence-based practice is an individual responsibility, and that research utilization is a valuable outcome in and of itself. Perhaps the most striking assumption in these studies and embedded in the instruments was that of rationality, i.e., the assumption that decisions to use research are the decisions of rational actors in rational environments.

6. Discussion

Our review of research utilization measurement in nursing has revealed that both single item measures and more widely used instruments reflect a common set of problems. First, there are general problems that include: lack of research utilization or appropriate other theory, lack of construct clarity; lack of explicit measurement theory, lack of advanced psychometric assessment, a presumption of linearity, an absence of longitudinal work (neither repeated measurement in studies or frequent use of any instrument) and potentially influential but unacknowledged assumptions. Second, these instruments were also commonly susceptible to a number of other problems, including self-report and recall biases, social desirability biases, under-developed scaling approaches, and lack of clarity regarding the unit of analysis (and resultant ecology fallacies).
Lack of research utilization or appropriate other theory

We have established that little in the way of research utilization or other related theory has guided the development of research utilization measures to date in nursing. A more interesting avenue may be to pursue what sorts of ‘models’ should guide us. In 1986 Van de Ven (van de Ven, 1986) argued against the emergence of a unified model, suggesting that competing theories would emerge that would serve us better. Potentially useful theory and theoretical models include those in the field of clinical decision making (Tavakoli, Davies, and Thomson, 2000; Dowding and Thompson, 2003), decision making (Groen and Patel, 1985; Kahneman, Slovic and Tversky, 1982; Patel, Arora, and Kaufman, 1999), critical thinking (Facione, 1992), information processing theory (Newell and Simon, 1972), self-efficacy (Bandura, 1977), behavioral change (Prochaska and DiClemente, 1986; Prochaska and DiClemente, 1992; Oldenburg, Glanz, and French, 1999), organizational theory (Brown and Duguid, 2001; Czarniawska and Sevén, 1996; Denis et al., 2002; Orlikowski, 2002) and evaluation utilization models (Greene, 1988a; Greene, 1988b; Cousins and Leithwood, 1993; Patton, 1997; Johnson, 1998). Finally a group of models less commonly associated with research or knowledge utilization, but which we believe are of significant importance to this area are policy analysis frameworks (e.g., Stone, 1999; Hogwood and Gunn, 1984; Lomas, 2000, Jenkins-Smith and Sabatier, 1993).

Lack of construct clarity

Construct validity is premised on a well-grounded theoretical conceptualization of research utilization. In addition to theorizing, there are additional methodological approaches to achieving it. Cronbach and Mehl’s nomological network (1953) is one well accepted approach. However it carries strenuous demands for strong theory. Kane’s (1992) constructed arguments (sets of logical, testable postulates that are used to fill in gaps in theory) approach is less demanding of strong theory and consequently may be more suitable in the research utilization field.

Lack of explicit measurement theory

We have identified a consistent absence of measurement theory in the development of reviewed instruments. We argue that using an appropriate theory (or theories) would benefit the development of any research utilization instrument and is strongly needed for guidance in future development. Choosing which theory to use depends on many factors, all stemming from the inferences that one wishes to make about research utilization once it has been measured. The hypothesized inferences will drive how a study is designed and the way in which data are collected and analyzed. The measurement model must fit with the construct (in line with the desired inferences), and will inform the way in which the instrument development data are collected.

Lack of psychometric assessment.

All of the current studies lack significant psychometric assessment of used instruments. The use of the multi-item instruments has commonly been accompanied by testing homogeneity using Cronbach alpha, while most of the single-item instruments have no measure of reliability. Validity in some cases was assumed because investigators drew on original instrument development and in fewer cases by actually assessing it. In Lacey’s pilot study (1994) validity was strengthened by semi structured interviews and in Pain’s EROS study (Pain, Hagler, and Warren, 1996) evidence of construct validity was reported. However, over half of the studies did not mention validity. As there is frequent use of various types of scales it is surprising that psychometric testing is rarely reported. This is, however, in line with an ill-defined construct of interest and emphasizes Dunn’s (1983) claim of two decades ago that construct validity is a serious and unresolved problem in the field.

Presumption of linearity

The presumption that research utilization is a linear process is most obvious in the NPQ instrument, but is present in the other instruments and single-item questions. This presumption is closely linked to the idea of rationality and presumes that the use of research follows a planned course — identification, critical appraisal, and implementation. This understanding may be analytically useful but in practice, does not occur. Further the least attention has been devoted to the most complex part of the process, implementation. Recognized guidelines frequently need adaptation to local conditions. Further, the complexity of most health settings and the intensity of care needs contribute to change being a circuitous not a linear route. Such conditions need to be factored into the measurement of research utilization.

Unacknowledged assumptions

A number of unacknowledged assumptions adversely influence the measurement of research utilization. They do so by creating conditions where respondents may understand questionnaires as idealistic and artificial, rather than as an organic part of a practice requiring many forms
of knowledge in addition to science. In particular, assumptions of linearity and rationality run counter to the complex and often messy conditions replete with competitive interests and influences prevailing in most health care facilities today (Kitson, Harvey, and McCormack, 1998; Rycroft-Malone et al., 2002).

Estebovsk (2001), Norman (1999), and Rich & Oh (2000) have argued persuasively that a rationality bias is one of the most serious impediments to successful advances in the field. We have observed that this particular bias persists through much of the historical and contemporary literature in the broad field of knowledge utilization, although alternative approaches are emerging in fields such as organizational studies.

We located three published papers expressly addressing the measurement issue in this field (Dunn, 1983; Rich, 1997; Weiss, 1981). These authors have each offered guidance on approaches to move this field forward. Their arguments illustrate the complexity and wide range of possibilities required to conceptualize and measure research utilization. Each calls unmistakably for conceptual clarity. Commonly they also address the necessity of pluralism in procedures and methods of measurement. Separately, Dunn (1983) emphasizes the need for strengthening reliability and validity given the subjective properties inherent in knowledge utilization. Weiss (1981) argues for the need for specific foci, that is, on studies, people, issues or organizations and that these foci determine method. Rich (1997) provides a comprehensive overview of issues influencing research utilization.

7. Future Directions

Rich (1997) examining the knowledge utilization field broadly across many disciplines, claimed that little new empirical work had been reported in the knowledge utilization field in the last decade and that a bias toward measuring what was easy to measure persisted. We found little to refute Rich’s claims. Although there has been a measurable increase in empirical work on research utilization in nursing in the past fifteen years, little new information has been reported and few methodological advances are apparent – the picture regarding the measurement of research utilization is perhaps bleakest of all. Further, in our review of knowledge utilization literature across many fields and disciplines we found little to distinguish those fields from nursing when

searching for methodological advances.

In aiming for the development of a research utilization measurement instrument in nursing we suggest the following be re-examined. First, Dunn’s (1983) overview of instruments for measuring knowledge (or information) use, in which he presents instrument used in other fields, e.g., the information utilization scale (Larsen, 1982); the stages of concern scale (Hall et al., 1979); the evaluation utilization scale (Johnson, 1980); and the overall policy impact scale (van de Vall and Bolas, 1982). These scales and indices, as well as, newer instruments may offer productive ideas on problems in the measurement of research utilization in our field of interest -- nursing. Another valuable contribution that could inform development of an instrument for measuring research in nursing may be Knott & Wildavsky’s (1980) “seven standards of utilization” – perception, cognition, reference, effort, adoption, implementation and impact. The thorough work done by Johnson (1998) in examining implicit and explicit theoretical models for evaluation utilization may also have bearing for such efforts in our field.

We advocate the study of research utilization as both process and product because both are necessary to assess its impact on health outcomes. However, the most fruitful courses of study probably lies in the examination of research utilization as process. Whether one is examining knowledge (or research) utilization as product or process, however, the most pressing issue ahead of us lies in establishing construct clarity in order that defensible measurement can follow. Partitioning the research utilization process into manageable components and clearly delineating the focus of attention will ensure the appropriate measurement approach and selection.

Meaningful advancement in this field requires the intervention studies aimed at the implementation phase of research be mounted to test strategies believed to enhance research utilization. These studies need to incorporate longitudinal designs in order to evaluate sustainability and they need to be embedded in sound theorizing. The clarion call for good theory in this field remains unanswered. If we are to measure research utilization – we are going to have to answer it.
of knowledge in addition to science. In particular, assumptions of linearity and rationality run counter to the complex and often messy conditions replete with competitive interests and influences prevailing in most health care facilities today (Kitson, Harvey, and McCormack, 1998; Rycroft-Malone et al., 2002).

Essabrooks (2001), Norman (1999), and Rich & Oh (2000) have argued persuasively that a rationality bias is one of the most serious impediments to successful advances in the field. We have observed that this particular bias persists through much of the historical and contemporary literature in the broad field of knowledge utilization, although alternative approaches are emerging in fields such as organizational studies.

We located three published papers expressly addressing the measurement issue in this field (Dunn, 1983; Rich, 1997; Weiss, 1981). These authors have each offered guidance on approaches to move this field forward. Their arguments illustrate the complexity and wide range of possibilities required to conceptualize and measure research utilization. Each calls unmistakably for conceptual clarity. Commonly they also address the necessity of pluralism in procedures and methods of measurement. Separately, Dunn (1983) emphasizes the need for strengthening reliability and validity given the subjective properties inherent in knowledge utilization. Weiss (1981) argues for the need for specific fact, that is, on studies, people, issues or organizations and that these fact determine method. Rich (1997) provides a comprehensive overview of issues influencing research utilization.

7. Future Directions

Rich (1997) examining the knowledge utilization field broadly across many disciplines, claimed that little new empirical work had been reported in the knowledge utilization field in the last decade and that a bias toward measuring what was easy to measure persisted. We found little to refute Rich’s claims. Although there has been a measurable increase in empirical work on research utilization in nursing in the past fifteen years, little new information has been reported and few methodological advances are apparent — the picture regarding the measurement of research utilization is perhaps bleakest of all. Further, in our review of knowledge utilization literature across many fields and disciplines we found little to distinguish those fields from nursing when searching for methodological advances.

In aiming for the development of a research utilization measurement instrument in nursing we suggest the following be re-examined. First, Dunn’s (1983) overview of instruments for measuring knowledge (or information) use, in which he presents instrument used in other fields, e.g., the information utilization scale (Larsen, 1982); the stages of concern scale (Hall et al, 1979); the evaluation utilization scale (Johnson, 1980) and the overall policy impact scale (van de Val and Bolas, 1982). These scales and indices, as well as, newer instruments may offer productive ideas on problems in the measurement of research utilization in our field of interest — nursing. Another valuable contribution that could inform development of an instrument for measuring research in nursing may be Knott & Wildavsky’s (1980) “seven standards of utilization” — reception, cognition, reference, effort, adoption, implementation and impact. The thorough work done by Johnson (1998) in examining implicit and explicit theoretical models for evaluation utilization may also have bearing for such efforts in our field.

We advocate the study of research utilization as both process and product because both are necessary to assess its impact on health outcomes. However, the most fruitful courses of study probably lies in the examination of research utilization as process. Whether one is examining knowledge (or research) utilization as product or process, however, the most pressing issue ahead of us lies in establishing construct clarity in order that defensible measurement can follow. Partitioning the research utilization process into manageable components and clearly delineating the focus of attention will ensure the appropriate measurement approach and selection.

Meaningful advancement in this field requires that intervention studies aimed at the implementation phase of research be mounted to test strategies believed to enhance research utilization. These studies need to incorporate longitudinal designs in order to evaluate sustainability and they need to be embedded in sound theorizing. The clarion call for good theory in this field remains unanswered. If we are to measure research utilization — we are going to have to answer it.
References


Greene, J. C. (1988a). Communication of results and utilization in


Rich, R. F. (1975). Selective utilization of social science related...


