Improving Survey Development and Objective Measurement in Nursing Research: Construct Modeling and Rasch Diagnostics Analysis

MARY P. BOURKE, RN, MSN, PhD.
ASSISTANT DEAN OF GRADUATE PROGRAMS
INDIANA UNIVERSITY KOKOMO
Sigma Theta Tau International’s 29th Nursing Research Congress
Objectives

Objective 1: Understand how Construct Modeling is used to identify the characteristics of a construct under study, and therefore a guide for instrument development.

Objective 2: Understand how Rash Model diagnostics are used to evaluate the effectiveness of instruments developed for objective measurement.
Problem: Nurses and physicians at this Midwest regional medical center had on-going conversations about the difficulty with communication.

Due to the quest for Magnet Status, administration was very interested in identifying problems from the nurses perspective. There is a lack of research in the area of nurse physician communication.

It was determined, that through construct modeling, characteristics of communication could be identified and items written to measure these characteristics.

Construct Modeling allows a researcher to focus on the institutional assessment needs (context) versus creating an instrument to generalize to a population.
“Communication between nurses and physicians does not flow as it should.” (Greenfield, 1999)

“Perspectives are often so vastly different between the two professions that finding common ground on which to base communication may not be possible.” (Stein-Parbury & Liaschenko, 2007)

“Nurses may be more dissatisfied in their attempts to communicate with physicians than vice versa because of the subordinate role that nurses hold relative to medicine and administration authorities.” (Aiken, Smith, & Lake, 1994)

“Although touted as a patient safety tool, the timeliness of communication was not associated with communication satisfaction. Nurses are more satisfied with understanding, open, and accurate communication, especially with attending-level physicians.” (Manojlovich & Antonakos, 2008)

“The hierarchical boundary that frequently separates nurses from physicians can inhibit communication.” (Edmondson, 2003)
Purpose: A partnership was formed between a large regional medical center and nursing faculty from a large university. The partnership-facilitated research was created for the purpose of identifying nurses perceptions of communication problems between physicians and nurses.

A team was formed that included one nursing faculty and five staff & administrative nurses.

After several meetings, characteristics of communication between physicians and nurses were identified.
Method: We began the process of constructing an instrument based on the Construct Modeling approach. “Construct Modeling is a framework for developing an instrument.” (Wilson, 2005) The Building Blocks of Construct Modeling are as follows:

The construct map: The Construct Map: The construct map process is a way to conceptualize a construct. A construct is the object or concept we want to measure such as communication.

The design plan: “A procedure, or design, that allows the observations to be made under a set of standard conditions that span the intended range of the item contexts”. (Wilson, 2005) Items were designed based on the construct map. The items had a lengthy review process using content experts and professional peers. Several iterations of the instrument were developed based on feedback.

The outcome space: Likert type response set. (Strongly Disagree, Disagree, Agree, Strongly Agree)

Measurement model: Rasch measurement.
Dimensions of Communication
Identified-communication is the activity of conveying meaningful information.

Abuse
Competent
Self Esteem
Anxiety
Fear
<table>
<thead>
<tr>
<th>Respondents</th>
<th>Responses to Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registered Nurses with strong agreement to perceptions of communication between physicians and nurses.</td>
<td>Ordered series of levels: Characteristics inherent in the construct as defined within this context:</td>
</tr>
<tr>
<td>Registered Nurses with moderate agreement to perceptions of communication between physicians and nurses.</td>
<td>Items identify characteristics of Abuse</td>
</tr>
<tr>
<td>Registered Nurses with strong disagreement to perceptions of communication between physicians and nurses.</td>
<td>Items identify characteristics of Fear</td>
</tr>
<tr>
<td>Registered Nurses with moderate disagreement to perceptions of communication between physicians and nurses.</td>
<td>Items identify characteristics of Anxiety</td>
</tr>
<tr>
<td></td>
<td>Items identify characteristics of Self Esteem</td>
</tr>
<tr>
<td></td>
<td>Items identify characteristics of Questioning Competent Practice</td>
</tr>
</tbody>
</table>
Population: RN’s employed at a regional medical center (89 RN’s responded). A research day was organized and most RN’s attended giving us an adequate sample size.

Example Items-Design Plan:

- I feel uneasy when I call physicians at home.
- I have experienced a physician questioning my competency.
- A physician has personally belittled me in the presence of my co-workers.
- I feel some physicians exhibit intimidating behaviors.
- Co-workers discourage me from contacting certain physicians due to their past behaviors.

The outcome space-Likert type response set. (Strongly Disagree, Disagree, Agree, Strongly Agree)
Measurement Model

Analysis: Rasch analysis allowed a more appropriate analysis of the data and provides detailed information about the psychometric properties of an instrument.

Normally, traditional test theory is used for survey data ignoring rules for ordinal data. Gravetter & Wallnau (2010) describe the limitations of ordinal data as the inability to measure the distance between each category consequently, it is inappropriate to use traditional statistics. Ordinal values only indicate the direction from one score to another.

Bertoli-Barsotti (2005) describes Rasch Analysis as an acceptable alternative paradigm to traditional tests for ordinal data such as Chi Square, Mann-Whitney, Wilcoxon and Kruskal-Wallis.

When ordinal data fit the Rasch model requirements, Rasch “transforms ordinal-level data into estimates of item/person parameters on interval-level scaling.” (p.72.)

Rasch provides a mathematical framework to determine unidimensionality of a construct being measured on a less than/more than scale.
Item Polarity is investigated using the Point-Measure Correlation. This identifies if the responses to a particular item align with the construct we are measuring.

Tool diagnostics were performed as follows: category frequencies and average measures; as well as, threshold estimates, probability curves, and category fit statistics.

Tool diagnostics make explicit the developer’s intentions and the respondents’ record of their attitudes, behaviors, or achievement on the construct of interest. (Bond & Fox, 2007) The series of diagnostics reveal the same story of instrument functioning.
Item Polarity is a dimensionality diagnostic specific to the rating scale model. It indicates if the responses to a particular item align with the overall construct under measurement.

See table 1. in handout (PR-MEASURE correlations that are negative or close to zero identifies problematic items not consistent with the construct. We want to see moderate to high (.37 to .65) correlations.
Assess category functioning by looking at category frequencies. Frequencies indicate how the respondents used each item’s response categories. (See Table 2. in handout.)

Average measures – In the rating scale model, the average measures is the average of the ability estimates for all persons in the sample who chose that particular response category, with the average calculated across all observations in that category. Average measures should increase monotonically across response categories. (Bond & Fox, 2015) (See Table 3. in handout.)
Thresholds or (Step calibrations) are the “difficulties estimated for choosing one response category over another”. For example, how difficult was it for nurses to endorse one item over another. ("Calibration is the procedure of estimating person ability or item difficulty by converting raw score odds to logits on a Rasch measurement scale.” Bond & Fox, 2015)

Step calibrations must increase monotonically. If not the diagnosis is disordered. (See table 4. in handout.)
A Probability Curve is a way to visually inspect the distinction between thresholds. Each category should have a distinct peak. Problematic thresholds that are disordered will be obvious. Distances of thresholds should increase monotonically by 1.4 logits and no more than 5 logits. (See Table 5 in handout.)
Fit statistics provide a criteria to determine the quality of rating scales. Infit and Outfit mean squares < than 0.7 indicate overfit to the Rasch Model (to good to be true-Guttman) over 1.3 indicate more misinformation than information. (Linacre, 2009) (See Table 6 in the handout)

Validity and reliability are articulated within the context of the Rasch model. (See handout)
Guidelines for Collapsing Categories

Collapsing of Categories—Collapse a category into an adjacent category. For example collapse Strongly Agree data into Agree.

1) Does it make sense?

2) Attempt to create a more uniform frequency distribution.

See handout.
Reliability in the Rasch Model is estimated for both persons and items.
The person separation index is used for estimating the spread of persons on the measured variable.
The item separation index is used to measure reliability.

Validity—Construct and Internal Validity. (See handout.)
Differential Item Function—Comparing the estimates across two or more distinct groups of interest such as male for female to examine if items have significantly different meanings for the different groups.

The results of DIF contribute to the investigation of instrument functioning.
Thank you for attending

References-See handout.