

Reviewer Comments and checklist for Resnick manuscript:

Reviewer #1: The measure described in this paper addresses an important problem, supporting nursing assistants' in their provision of restorative care to nursing home residents. Being able to objectively describe nursing assistants' restorative care activities is foundational to understanding variations in care provision and to testing interventions to improve the lives of people who live in nursing homes.

The authors describe U.S. policies that impact nursing assistants' provision of restorative care. As an international reader, I believe that the measure is relevant beyond the U.S.A. The authors might consider clearly indicating that these are U.S. policies but that the philosophy of care is not limited to the U.S. Similarly, stating that the reimbursement system (p. 4, line 9) is a U.S. system, might clarify the context for international readers.

The "Factors that Influence Restorative Care Behaviors of NAs" section of the literature review describes the Restorative Care Behavior Checklist being based in self-efficacy theory. The literature review clearly links this theory to evidence about nursing assistants' work. It does not include information about the relationship between knowledge and performance of restorative care behaviors - but this is one of the factors in the theoretical model that is used for validity testing. The discussion section includes additional factors external to the nursing assistant that may influence restorative care behaviors. If, in the introduction to this section, the authors explicitly state that the checklist was developed based on self-efficacy theory and that this theory is a model of internal factors that influence an individual's behaviors, then the lack of discussion of the external factors in the literature review and the connection between this section and the discussion would make more sense. Otherwise, a more complete literature review would discuss the full range of factors that influence nursing assistants' work.

The theoretical model that is used for validity testing is summarized in a complicated diagram. I was not able to distinguish between the arrow-heads on the diagram, so the hypothesized relationships between constructs and variables was not clear to me. The hypothesized relationships are not clearly stated in the literature review or the validity testing subsection of the methods section. What does education mean in the figure?

The research methods were clearly described and justified. The detailed description of Rasch analysis and interpretation of results of Rasch analysis is necessary for readers who may be less familiar with this approach. I have a few suggestions:

1. Clarify whether measures other than the Restorative Care Behavior Checklist are published measures or whether they were developed by the authors for this study. There is no citation for the prior use of the Job Attitude Scale (p. 13, line 3-4).
2. Reformat the diagram of the theoretical model so that the reader can distinguish between arrow-heads and include some description of the hypothesized relationships in the text.
3. Include description of the residents in the nursing homes or the residents for whom care was observed as well as what time of the day care was observed. This would help with interpretation of the findings that some types of care were rarely observed. It would

also help with interpretation of discussion that the findings might be influenced by the kinds of patients who were observed.

4. Why is it not necessary to do any of the psychometric analysis on the summary score of the checklist?
5. Appendix A was not included in my version of the manuscript (p. 11, line 17), so I cannot comment on it.
6. Why did the authors decide against using kappa to test for item agreement? I would be interested in knowing about agreement on two levels: (1) whether or not the care activity was being provided, the denominator of the summary score; and (2) if the care activity was being provided, whether or not a restorative care approach was used, the numerator of the summary score. This would help with interpretation of the finding that several of the activities were rarely performed. Disagreement about whether an activity is being performed would seem to be less important than disagreement about whether a restorative care approach was being used. It is not clear from the manuscript whether the Rasch analysis could differentiate this.

The results section is comprehensive. I have a few suggestions to make it easier for the reader to understand the findings.

1. Data about education are missing from Table 1
2. There are duplicate lines in Table 2 (Number of restorative care activities performed and Number of restorative care activities not performed, with different rounding for SD of number of restorative care activities not performed). Please explain what these lines mean in the text. Why is it appropriate to report absolute number of restorative care activities? The scoring for the scale is the ratio of restorative care activities to observed activities (nursing assistants' use of restorative care when they had an opportunity to do so). I don't understand what is meant by "Overall they performed restorative care during 64% of the observed activities (evidence of)" (p. 16, line 1 to 2). I assume this is an average over the 386 nursing assistants and that there was variation among the nursing assistants. Please include information about variation between nursing assistants.
3. Explain what "missing" means in Table 3 and clarify the sample size. There are consistently 8 missing and the total observations (not applicable + performed + not performed + missing) is 395 (387 + 8 missing). The reported sample size is 386 (p. 10, line 14). This doesn't make sense.
4. The Validity Testing subsection concludes that the fit statistic for Communication OUTFIT is high but not important enough to threaten fit of the model. This is inconsistent with the description of the Rasch analysis in the methods section. A citation to support this conclusion as consistent in Rasch methods would assist the reader.

In the discussion section, I am interested in knowing whether the authors think that the low frequency of providing care related to eating, assistive devices and exercise might have impacted internal consistency. Is it possible that the hypotheses used for validity testing are wrong? It is not clear to me that the study, as reported, provides support for validity of the RCBC (p. 20, line 15). The convergent validity testing results do not support the conclusion of validity.

This paper is well organized and was a pleasure to read. I congratulate the authors on

their work and contribution to improving support for nursing assistant work in nursing homes.

Reviewer #2: Overall, a well written manuscript that was easy to read and follow. Other suggestions/comments/questions:

Page 4 lines 18-23. First time that an explanation of restorative care behaviors occurs. I would recommend defining restorative care early on in the manuscript for readers that may not be as well versed as the authors in the topic.

I would also recommend that the author(s) spell out nursing assistant and include the abbreviation early in the manuscript. I only saw one way or the other but never together to explain what the abbreviation stood for.

Page 8, line 16. I would also suggest spelling out the Restorative Care Behaviors Checklist and including the abbreviation the first time it appears.

Page 9, lines 4-6. Why the heterogeneity in the evaluators? I don't see how this potential impact was addressed in the results?

Page 10, lines 7-9. This seems a little vague. What do you mean by "known sources of information"? I would suggest being a little more specific here and providing citations.

Page 10 lines 22-23. How did you take into account the potential effects of 36% of the nursing assistants having prior training on restorative care?

Page 11 lines 13-17. How many total items are on the RCBC? I did not receive the appendix A you refer to.

Page 13 lines 15-20. The advantage to using the Rasch Analysis over KR-20 or Cronbach alpha is not clear.

Page 16 lines 21-22. Earlier you stated that exercise was one of the least likely activities to be observed?

Page 18 lines 14-15. Total number of items on the RCBC? Could there be subscales identified within the 10 categories of behaviors? Then, you would probably have an increase internal consistency at least for the subscales.

Figure 1. I would suggest putting only the 8 paths that were statistically significant rather than trying to depict visually all 58 hypothesized paths. It is very difficult to follow as it is currently designed.

Reviewer #3: Stat reviewer

This review is focused on measurement and statistical analysis.

In general, the analysis was described well and only some minor suggestions for clarity are listed below, with the exception of the AMOS model, which needs clarification that assumptions were checked, that the sample was large enough, and that the results make sense. Figure 1 cannot be read as there are too many paths. In addition, some suggestions for additional issues to include in discussion are presented below.

P 10 L 9: It would be useful to present a brief rationale for the predictors of age, gender, and ethnicity as concepts that would influence knowledge of restorative care, self-efficacy, and outcome expectations. In particular, it seems odd to consider gender as a potential predictor of these. Both age and ethnicity had little variation in the data and thus may not meet the assumptions for a predictor variable in the AMOS analysis. This is a minor weakness that could be addressed in a couple of sentences: one to clarify the rationale in this section, and another to state that assumptions for AMOS analysis were met in terms of variability, normality of the variables. See, for example, Kline p 84 who states that each group within a grouped exogenous variable must be screened separately to be sure it meets the assumptions. (Rex Kline, Principles and Practice of Structural Equation Modeling.)

P 11 L9: This description of the RCBC items is confusing. Suggest calling "interactions" in line 8 "activities" instead. This may clarify the meaning of "Activity not performed..." in line 9. As it reads now, it is not clear whether activity refers to a restorative care activity or a resident activity. In addition, the use of the term "restorative care activities", as in line 1 of page 12, adds to the unclear distinction between the activity done between NA and resident and the restorative care intervention (or behavior or interaction).

P 13 L10: From the wording in this line, the reader expects to see a Person Separation score and a Person Separation Reliability score. There are two scores in the results, one of which is called the Person Reliability score. Clarifying these terms and using them consistently would strengthen the manuscript and improve understanding of the statistical analysis. In this paragraph, provide the meaning or range of potential scores. Without this information, the result on p 18 that score was 6.63 cannot be interpreted. The score on p 18 for "person reliability" of .77 is presented in the next paragraph as the score for "person separation".

P 14: The INFIT and OUTFIT description is very clear

P 16 L19: Clarify the meaning of "mapping" which was not described in the analysis section.

P 17: Add a justification for testing such a complex model of 58 parameters with a sample of 386 subjects. Kline (see above) presents a rule of thumb of minimum of 10 subjects per parameter estimated with 5:1 making a model unstable. This would suggest that 38 parameters was the limit. Authors should assure the reader that assumptions for AMOS model were tested (see comment on variability above) and met. There is little discussion about the theoretical rationale for the large number of proposed paths in the model, and perhaps a more parsimonious model could be presented as the initial model, especially as the initial model did not produce good results. Figure 1 is impossible to read with so many intersecting paths. A simpler model, with rationale, would be easier to present. Alternatively, present assurance that the assumptions for the data were considered and that the sample size was adequate. Presenting the results from page 17 in either a figure or a table would facilitate understanding the significant pathways described in the text.

P 17 L19 State which model is referred to, the original or the trimmed model.

P 18 L 15, add a brief explanation of why a wide range of behaviors would influence the score of .77.

Discussion

Regarding the findings, a weakness of the study that should be noted was the lack of observations of eating, use of assistive devices, and exercise (87-93% not observed) of observations did not include these activities. One would expect that these would be prime examples of opportunities for restorative care interventions and the lack of data should be noted as unfortunate and perhaps a reason given for their absence in terms of methods of the study.

The discussion should deal with the issue of gender being the only predictor of engaging in restorative care. First, gender had almost no variability so it should be acknowledged that this is not a particularly useful finding. Second, the fact that none of the hypothesized predictors had an influence on engaging in restorative care should be acknowledged as unexpected and some reasons why posited. This could be tied in with the social ecological model paragraph. In addition, the finding that knowledge and self-efficacy led to outcome expectations, but not to actual restorative care behavior should be mentioned in the discussion along with possible reasons for this unexpected finding.

In regards to findings on validity testing on p 16, where exercise and eating are included with the most common activities in which NAs perform restorative care, these findings should be acknowledged as not useful when so few of these activities were observed, as shown in Table 3.

CHECKLIST FOR STYLE

Title Page

Supply affiliation and professional title for each author.

References

Other: The Chang & Lin, 2005 reference does not match the in-text citation of Chang, 2005. Please correct.