

## NRES-D-10-00010R1: Response to Reviewers' Comments

<b>Editor Comments</b>	<b>Response</b>
<p>It seems especially important to keep the discussion section focused on the primary variables. The reviewers' view is that there are relatively few novel results. Do not try to conceal this by focusing on lesser topics.</p>	<p>In the original manuscript submission, in discussion section, there was one sentence at the beginning of paragraph four that addressed the findings related to the variables in the baseline model, or the variables other than physical work environment and job satisfaction. Therefore, we believe that our discussion was focused on the primary variables.</p> <p>In response to a reviewer's request during the first revision, the revised manuscript included an expanded discussion of two non-primary variables, including the relationship between 12-hour shift and race or ethnicity and job satisfaction.</p>
<p>Keep in mind that if additional space is required to fully describe your methods, you can prepare some materials for supplemental digital content, which will be posted on-line only. This feature allows you to use the words you need to build your case, but does not increase the length of the printed manuscript if it is accepted for publication.</p>	<p>We are submitting two supplemental digital content files. Additional validity information is included in Digital Supplemental Content 1. Digital Supplemental Content 2 includes bivariate correlations matrix requested by one of the reviewers (#3). All in text edits are marked in red font.</p>
<b>Reviewer 1 Comments</b>	<b>Response</b>
<p>First, I commend the authors for their careful consideration and response to the reviewers' comments. They have attempted to clarify the methods and analyses used in the study, and the writing is much clearer.</p> <p>That said, I am less enthusiastic about this paper than before. The major purpose of the study is to explore the relationship between the physical work environment and nurses' job satisfaction, controlling for all the other factors that have been found to influence nurses' job satisfaction. The analysis failed to show any association between the physical work environment and job satisfaction, as measured in this study.</p>	<p>Thank you.</p> <p>The authors believe that this study makes an important contribution to building of theory of RN job satisfaction. While in our study no significant association between physical work environment and job satisfaction was found when multiple job satisfaction confounders were controlled for, the extant studies that have found significant relationship between the two variables did not provide control of multiple confounders and thus may have found the</p>

<p>The lengthy discussion, however, is focused on the relationship of all the other independent variables and job satisfaction. Although a few of the relationships were important to point out (perceptions of members of racial minority groups and effect of 12 hour shifts), the rest is not pertinent to the focus of the study.</p>	<p>significant relationship between these two variables because they have tested an underspecified model.</p> <p>Apart from one sentence summary of the findings from the baseline model at the beginning of paragraph four in the two-page discussion section, there was no discussion of non-primary study variables in the original manuscript submission. The discussion of the relationship between 12 hours shift and race was included upon this reviewer's request. We therefore believe that our brief discussion section is focused on findings related to primary variables.</p>
<p>There was little attention paid to the various items on the physical environment scale, and on any measured differences across types of units.</p> <p>Bottom line, there was no association between physical work environment and job satisfaction, in spite of the bivariate results.</p>	<p>We have not performed unit comparisons in this study. We do report results of analysis of responses to the individual items on page 14 of the manuscript:  “...The majority of respondents did not agree that the interior design in their workplace created a cheerful or warm work climate (71%), had a contemporary appearance (77.8%), or had an optimal furniture arrangement (75.2%)...”</p> <p>We have tested a hypothesis based on previous research and theory from non-healthcare settings that suggest physical work environment and job satisfaction are related. The hypothesis was not supported in our study and we believe that it is important to report these findings, especially since no other study has tested this relationship controlling for multiple confounders of job satisfaction.</p>
<p>I remain concerned with the lack of generalizability of the study results, since the makeup of the nursing staff is not typical (high percentage of BSN).</p>	<p>Limited generalizability is stated as a limitation. We believe that other magnet hospitals employ primarily BSN graduates.</p>

<p>Some of the numbers do not add up. For example, if the 5 items on the job satisfaction scale have 7 levels each, there should have been 35 response categories, not 31 as described in the paper.</p>	<p>The numbers do add up in the following way: There are 31 categories contained in the full range of the scale scores from a minimum score of 5 to the maximum score of 35.</p>
<p>It is impossible to interpret the collapsed 5 level response scale. If it ranges from 0-4, would not a 2 equal neither satisfied nor dissatisfied? If not, what are the anchors at each end of the scale?</p>	<p>As indicated in the manuscript on page 8, the anchors for the general satisfaction item ranged from 1 (<i>very dissatisfied</i>) to 7 (<i>very satisfied</i>).</p> <p>On page 9 of the revised manuscript the following text was added:</p> <p><i>Following data collection, to aid interpretation of the collapsed response categories, anchors for the collapsed scale were assigned as follows: 0 (very dissatisfied), 1 (somewhat dissatisfied), 2 (neither satisfied nor dissatisfied), 3 (somewhat satisfied), and 4 (very satisfied).</i></p> <p>In this case, as this reviewer suggested, a score of 2 can in fact be interpreted as neither satisfied nor dissatisfied. This text was added on 13:</p> <p><i>“The job satisfaction scale score ranged from 0 to 4 and the respondents reported median level of job satisfaction at 2.0 (Interquartile Range = 2.0; 25th percentile, Q1 = 1.0; 75th percentile, Q3 = 3.0), indicating that, overall, the RNs in this sample were neither satisfied nor dissatisfied with their job.”</i></p>
<p>The same interpretation problem is presented for the physical environment scale. What does a 2.9 out of 9 mean? What are the anchors at each end of the scale?</p>	<p>The anchors were assigned to the Physical Comfort subscale of WES used to measure physical work environment in this study as suggested in Interpretive Report Form (Moos, 2008). On page 10 the following text was added:</p> <p><i>“Following data collection, to aid interpretation of the physical work environment scale scores, the following interpretive statements described by Moos (2008) were used to provide comparison of individual scores to the scores of work groups in general: 0,1,2 (considerably below average), 3 (well below average), 4 (below</i></p>

	<p><i>average), 5 (average), 6 (above average), and 7,8,9 (considerably above average)."</i></p> <p>On page 13 the following text was added:</p> <p><i>"The responding RNs reported a mean physical work environment score of 2.9 (SD = 2.2) on a scale ranging from 0 to 9 indicating that, overall, the RNs in this sample rated their physical work environment well below average compared to the score of work groups in general."</i></p>
I am also concerned that after decades of use, the internal consistency reliability of the tool remains below 0.70.	The suboptimal reliability of the physical work environment scale in this sample was stated as a limitation in the discussion section. At the time the study was conducted we were unable to locate other scales that provided a comprehensive measure of multiple physical work environment features. We agree that development of a new scale or refinement of the scale we used should be pursued in future research.
Additionally, I am not clear why both the independent and dependent variables would be identified as latent variables. However, this may be my lack of expertise with the methods used in the study.	Latent variable is any variable that can not be observed directly. Examples are preferences, attitudes, behavioral intentions, and personality traits. Such constructs can only be measured indirectly by means of observable indicators, such as questionnaire items designed to elicit responses related to an attitude or preference. Many of our independent variables (e.g., supervisor support, autonomy, variety) are measured with scales (i.e., observable indicators) and therefore we refer to them as latent variables as well our dependent variable (e.g., job satisfaction).
<b>Reviewer 2 Comments</b>	<b>Response</b>
The authors have responded thoroughly to most concerns expressed by reviewers. The study is well done, the article is well-written, and the knowledge produced useful to the field.	We agree, thank you.
The one remaining concern is the sensitivity of the instrument measuring the physical work	Although, comparing physical work environment scores across different RN units

<p>environment. It is good to learn that the measure captured the variance of RNs' subjective perceptions of the physical environment; however, the conceptual underpinning points more to the actual objective difference rather than individual's subjective responses to the environment.</p> <p>If this is measure is sensitive and specific the scores should vary systematically by unit (if the units actually were different in their physical environments).</p> <p>If the units were different and the tool did not capture the differences, perhaps the tool really doesn't capture the theoretical concept. If not - that could be the explanation for the lack of significant effect. Readers would appreciate knowing that about the tool.</p>	<p>was not within the scope of this study, evidence of sensitivity of the instrument to differentiate between units with different physical environment features is available from other studies. The following information was added as Supplemental Digital Content 1:</p> <p>“The Physical Comfort subscale of the WES (Moos, 2008) used to measure physical work environment in this study was not included in Kovner et al.’s study. However, evidence of validity from contrasting or known groups is available from two studies. Moos (2008) reported that, as expected, in a study of over 900 medical center employees, those who worked in a new building scored higher on the Physical Comfort subscale compared to employees who worked in an older, badly maintained building. Similarly, Dickens et al. (2005, p. 300) found that in comparing perceptions of work environment among four different nursing service units, the nurses who worked “in the least ‘fit for purpose’ building” had scored the lowest on the Physical Comfort subscale of the WES.”</p>
<p><b>Reviewer 3</b></p>	<p><b>Response</b></p>
<p>The author(s) have been responsive to many of the comments from the first review.</p>	<p>We agree, thank you.</p>
<p>However, my considerable concerns about the scoring of the dependent variable (job satisfaction) and the choice of data analysis strategies to test the model remain. I suggest they rethink their handling of this important variable, or perhaps see whether SEM using the original job satisfaction scoring would yield similar results. The decisions made about the dependent variable, type of analysis, the large number of control variables, and changes in the actual number of participants complicates the interpretation of lack of statistical significance and support for their hypothesis.</p>	<p>We believe that we have appropriately addressed our decisions about scoring of the dependent variable and choice of the analytic method in the previous response to the reviewer. We address the reviewer’s additional concerns in the responses below.</p>
<p>1. A cross sectional study is not a predictive study.</p>	<p>We agree that a cross sectional study does not test causation between independent and</p>

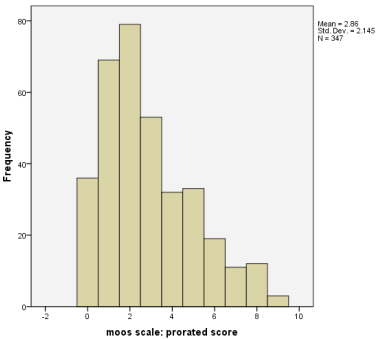
	<p>dependent variables as we have stated in the discussion section. However, we used the definition of predictive design from Burns and Grove (2005, p. 240) who explain that “Predictive designs are used to predict the value of one variable on the basis of values obtained from another variable or variables.” Further, they explain that “Predictive designs require the development of a theory-based mathematical hypothesis proposing the independent variables are expected to predict the dependent variable effectively. The hypothesis is then tested with regression analysis” (p.241). We believe that our study fits this description.</p>
<p>2. The N shifts from 305-308 for age, organizational and unit tenure to 329-346 for other demographics to 362 as the stated sample size in the abstract and methods sections.</p>	<p>The N shifts because not all subjects provided data for all variables.</p>
<p>3. Categories of education do not include diplomas in Nursing, and race/ethnicity does not include Hispanics. It is not clear why these are omitted or in which category they have been placed.</p>	<p>The authors knew before the study was conducted that this particular hospital did not employ diploma nurses and therefore the response options to the education question did not include diploma as an option. Hispanics are included in category <i>Other</i> as indicated in foot note of Table 2.</p>
<p>4. Discussion in the abstract misrepresents the findings of the study, with too much emphasis on the importance of perceived physical work environment. Although significant, the bivariate correlation is small (<math>r = .26</math>) and this significant relationship disappears when controlling for other things.</p>	<p>We believe that we have not misrepresented the findings of the study.</p>
<p>5. It is not clear why the job satisfaction score was categorized. The scale contains 5 items and the range in possible scores is 5 to 35. Using it in parametric statistical tests would be quite appropriate. Indeed, it would be consistent with how the author(s) have used other rating scale scores in this same manuscript.</p>	<p>We have provided an explanation in the previous response explaining why the job satisfaction score was categorized. Ordinal probit regression is a parametric analytic technique. In our view, the number of categories is not what determines a nature of a variable, it is in fact the inability to show that the distance between the different response categories is not exactly one which precludes one from defining the variable as truly continuous, interval level measure.</p>
<p>6. Categorizing job satisfaction complicates the analysis and makes it less digestible for many</p>	<p>We believe that ordered probit analysis is the appropriate analytic technique given the level</p>

<p>readers. If it was not categorized, multiple regression or SEM would be very appropriate and much more understandable for most readers.</p>	<p>of measurement for the dependent variable in this study.</p>
<p>7. The mean and standard deviation of the job satisfaction scale (before categorization) is not provided.</p> <p>The manuscript says that there was "scarcity of data at many of the original response values." Using quintiles to define the groups, while statistically appropriate, can result in putting subjects with very different original scores together in the same group. Providing means, standard deviations, and ranges on the original job satisfaction scale for each group would be helpful.</p>	<p>We include this additional information on page 13:  <b>“Prior to reducing the number of job satisfaction response categories, the scale score ranged from 5 to 35 with a mean of 23.4 (<i>SD</i> = 7.3).”</b></p> <p>Based on the distribution of original scores below, we believe that the distribution of the cut points throughout the range of the scale is reasonable and therefore the concern that subjects with very different original scores were put together in the same group is unwarranted in this case. The original scale range was 5-35 or 1-31 and the new categories are 0-4 with the following distribution:</p> <p>1-13 = 0 (n= 69)  14-17 = 1 (n= 78)  18-22 = 2 (n = 70)  23-26 = 3 (n = 61)  27-31 = 4 (n = 69)</p>
<p>8. If the job satisfaction variable is considered strictly ordinal and not appropriate for parametric statistics, it should not be used in the confirmatory factor analysis. In addition, including the dependent variable with the independent variables in a confirmatory factor analysis where the structure is specified as only a correlation between the IVs and the DV (considering both to be on the same causal level) results in a misspecified model.</p>	<p>The reviewer mentioned two issues. First, it was mentioned that ordinal variables cannot be used in parametric statistics. However, ordinal items or variables are appropriate for parametric analysis. Ordinal logistic and ordinal probit regressions are parametric procedures. Ordinal items are not strictly appropriate for parametric procedures based on the linear model (OLS regression) because the outcome variable is assumed to be continuous and unbounded (i.e., it can take on any real number from + to – infinity). Although we presented confirmatory factor analyses results treating the ordinal (Likert) items (for job satisfaction and the other measures) as continuous, there are state-of-the-art procedures to conduct a confirmatory factor analysis of ordinal items. These procedures are implemented in the Mplus software we used. Basically, rather than use a linear regression to</p>

link the ordinal items to their respective latent variable, probit regression is used. Also, the overall estimation procedure changes from full information maximum likelihood when items are continuous to a robust weighted least squares estimator when the items are treated as ordinal. It is useful to note that the main incremental fit indices (CFI=.97 & TLI=.98) are higher when appropriately treating the items as ordinal compared to the earlier analysis that treated them as continuous (CFI=.90 & TLI=.89). This is now mentioned in the validity and reliability paragraph on page 10.

The reviewer suggested the including the latent variable for the dependent variable (job satisfaction) in a confirmatory factor analysis with the latent variables for all of the predictors while allowing all of the latent variables to merely correlation would lead to model misspecification. This is not exactly true. In a CFA, one merely wants to assess the measurement properties of the measures representing the various latent variables. One wants to look at factor loadings, correlations among the factors, and overall model fit. To do this, all relevant latent variables need to be in the model and the model should be saturated (e.g., each latent variable is allowed to correlate with each of the others). In this way, model fit only represents the “measurement model.” Once constraints are placed on the relations between the latent variables (such as directional relations and fixing relations to zero), we introduce the structural model. Here model fit represents blending of fit for both the measurement and structural model. It is not a clean assessment of just the measurement model. Although few individuals do so, model fit can be assessed for just the measurement model, just the structural model, and for the joint combination of measurement and structural model (which is what is normally reported when testing a structural equation model).

<p>9. Validity of the scales is not fully addressed with factor analysis (confirmatory or exploratory). It indicates how well each item/indicator is tied to its latent concept, but it does not indicate that the items/indicators measure the desired latent concept.</p> <p>Only hypothesis testing or known groups comparisons can provide strong evidence that the latent concept is what the analyst thinks it is. Evidence of validity in previous studies should be presented for each scale, especially for the WES by Moos.</p> <p>The yes/no format of the WES often results in a scale that is not reliable or valid, especially in samples of more educated participants who find it difficult to choose "yes" or "no" when their perception really is "sometimes."</p>	<p>We agree.</p> <p>Additional evidence of validity for all scales from other studies is presented in Supplemental Digital Content 1.</p> <p>We agree that development of a new measure or refinement of this measure to address its limitations should be pursued in future research. Validity of the scale from contrasting (or known) groups is submitted as part of Supplemental Digital Content 1. We acknowledge the suboptimal reliability of the scale as a limitation in the discussion section. Despite its limitations, we could not locate another instrument that measured physical work environment features as comprehensively as the instrument we used.</p>
<p>10. While the author(s) have adequately addressed the shared variance within units, it is likely that the items in the perceived physical environment indicator also have some shared variance due to using the same hospital for recruitment of all participants. The amount of shared hospital variance will vary with the extent to which the participants' perceptions reflect real architectural features, hospital-wide policies, and operating decisions by hospital administration. More discussion of this limitation would better highlight this big limitation. The solution would be to sample RNs from more hospitals - considering the N for power analysis to be the number of hospitals.</p>	<p>We agree that due to sampling of the subjects from a single hospital, we can not detect the impact of shared hospital variance in our model. We have added the following on 18:</p> <p>“Sampling from a single hospital also precludes the researchers’ ability to detect the impact of shared hospital variance in the tested model which could have impacted the outcome of the analysis.”</p>
<p>11. Table 1 can be shortened considerably by deleting demographic variables and their definitions (from age through unit tenure, and</p>	<p>We appreciate this reviewer’s suggestion for modifying Table 1. However, another reviewer requested that response options for some of the</p>

<p>unit type through patient load). Maybe combine others with Table 3?</p>	<p>demographic variables be included. Therefore, we decided to leave the demographic variables in Table 1.</p>
<p>12. Many would put ER, OR, PAR (recovery room) into the ICU category as critical care units, although the patient load is very different.</p>	<p>We have used the categorization that the studied hospital uses.</p>
<p>13. P 11, lines 244-245: were the series of MR analyses among the independent variables only (with each serving as the "y") or were they using job satisfaction as the "y"?</p>	<p>The series of MR analysis were among independent variables only where each independent variable in turn serves as dependent variable and the others are independent variables. This procedure was performed as suggested by Tabachnick &amp; Fidell (2007). The following was added on page 12:</p> <p>‘Based on multicollinearity analysis, in which tolerance levels, expressed as 1– squared multiple correlations (SMC) were examined in a series of multiple linear regressions among the independent variables in which each independent variable in turn served as dependent variable...’</p>
<p>14. p 12: the mean physical work environment score was quite low for a scale that ranges from 0 to 9. Do higher scores represent greater or lower physical comfort? In addition, the scale's standard deviation is quite large (compared to the mean), suggesting the scores were skewed or perhaps bimodal.</p>	<p>Higher scores represent greater comfort and this explanation was added on page 10:</p> <p>“Higher scores corresponded with more favorable perceptions of physical work environment.”</p> <p>The examination of the bar graph does not indicate bimodal distribution:</p>  <p>Although, this independent measure is positively skewed, ordered probit does not have distributional assumptions of normality and linearity and given the size of the studied</p>

	sample, the physical work environment variable distribution should not affect the outcome of the analysis.
15. Describing or displaying the correlations among the independent variables would be helpful in interpreting the results.	Correlation matrix displaying the correlations among the independent variables is submitted as Supplemental Digital Content 2.