*Title Page & Acknowledgements

A Multilevel Confirmatory Factor Analysis of the Practice Environment Scale (PES): A Case Study

Byron J. Gajewski, PhD
Associate Professor
Department of Biostatistics
University of Kansas Schools of Medicine & Nursing
3901 Rainbow Blvd.
Kansas City, Ks 66160

Diane K. Boyle, PhD, RN
Associate Professor
University of Kansas School of Nursing

Peggy Miller, PhD, RN
NDNQI RN Survey Coordinator
University of Kansas School of Nursing

Frances Oberhelman, MSN, RN
Graduate Research Assistant
University of Kansas School of Nursing

Nancy Dunton, PhD
Research Associate Professor
University of Kansas School of Nursing
3901 Rainbow Blvd.
Kansas City, Ks 66160

Corresponding author:
Byron J. Gajewski, PhD
Address for reprint requests:
3901 Rainbow Blvd.
MS 1026
Kansas City, Ks 66160
Phone: 913-588-1603
E-mail: bgajewski@kumc.edu

Acknowledgements: Partial funding for all authors comes from a contract with the American Nurses Association.

Keywords: NDNQI®; ANOVA; Cluster; RN job satisfaction; multilevel analysis

ABSTRACT

Background: The aim of this paper is to describe multilevel confirmatory factor analysis via a case study for investigating the validity of the Practice Environment Scale (PES), a measure of the nursing practice environment. The PES data are collected from registered nurses (RNs) in nursing units in hospitals that are members of the National Database of Nursing Quality Indicators® (NDNQI®). NDNQI collects RN and patient information to aid in quality improvement, and research at the nursing unit level. The PES is collected from the individual RN, but items are worded so that analyses can be conducted at the individual, unit, or hospital level. Therefore, we desire to examine the validity of the PES at both the individual and unit level.

Approach: The PES was administered to 72,889 RNs from 4,783 nursing units (16 unit types, e.g., critical care and obstetric) in 2007. The PES has 31 items that represent five different domains. We fit a multilevel confirmatory factor analytic model with a structure based on the five domains. From this model, we estimate the between unit loadings and within unit loadings to investigate factorial, convergent, and discriminant validity at both the unit and RN levels. To investigate criterion-related validity, we correlated the five PES domains with the seven NDNQI-Adapted Job Enjoyment items at the unit and RN levels (also using a multilevel model).

Results: The multilevel factor analysis provides evidence of factorial, convergent, discriminant, and criterion-related validity at both the unit and RN levels.

Discussion: The PES is a valid instrument for use in quality improvement and research both at the unit and individual RN levels.

A Multilevel Confirmatory Factor Analysis of Practice Environment Scale (PES): A Case Study

Recognizing concerns about registered nurse (RN) job satisfaction, RN staffing, and quality of patient care within acute care hospitals, the American Nurses Association (ANA) established the National Database of Nursing Quality Indicators® (NDNQI®) (ANA, 1995). The NDNQI serves as both a resource for quarterly national benchmarking to member hospitals on unit level nurse-sensitive indicators such as RN staffing and patient outcomes, and a repository for data linking RN workforce characteristics and patient outcomes (Montalvo, 2007).

The NDNQI RN Survey on job attitudes and the work context is offered to member hospitals annually. The NDNQI RN Survey options include the Practice Environment Scale (PES) of the Nursing Work Index (NWI) (Lake, 2002) and the NDNQI-Adapted Job Enjoyment (JE) scale (Boyle, Miller, Gajewski, Hart, & Dunton, 2006; Taunton et al., 2004). Both the PES and JE scales measure RN workgroup, or unit level constructs, consistent with all indicators within the NDNQI (e.g., staffing, fall rates). RN workgroup perceptions of the nursing practice environment and job satisfaction are complex, multilevel constructs. The multilevel perspective of organizational research was recognized in the development of the PES (Lake) and JE (Boyle et al.; Taunton et al.) scales. As individual RNs in workgroups and organizations are exposed to common features, events, and processes, over time they develop consensual views of the workgroup and organization through interacting and sharing (Kozlowski & Klein, 2000). Although both the PES and JE scales were developed using best practice psychometric methods (Forbes & Taunton, 1994; Nunnally & Bernstein, 1994; Verran, Gerber, & Milton, 1995; Verran, Mark, & Lamb, 1992), periodic re-evaluation is needed to ensure continued support for validity.

There are several types of validity evidence (Goodwin, 2002) which include evidence based on instrument content (experts evaluations); evidence based on internal structure (e.g., item analysis, factor analysis); and evidence based on relations to other variables (e.g., convergent, discriminant, criterion). A confirmatory factor analyses (CFA) addresses many of these validations simultaneously. Focusing on the PES, a five factor CFA analytic model with adequate fit indices provides evidence of *factorial validity* (internal structure). Statistically significant relationships between the manifest variables (items) and the latent variable (subscale) provide support for *convergent validity*, whereas low to moderate correlations among the subscales provides evidence of *discriminate validity*. JE is an external variable to the PES, so it can be correlated with PES to provide evidence of *criterion validity*.

The PES was developed for RNs to be informants about the organization, which is conceptualized as the nursing unit in the NDNQI RN Survey. While the correlation structure between items from the same RN is modeled with a single level CFA (usual practice), the correlation of RNs within the same unit is ignored. A multilevel CFA will address the problem of RNs within units.

Although researchers in nursing are leaders in developing instruments that have good psychometric properties, the literature lacks *multilevel* CFAs. We searched PubMed, CINAHL, Proquest, and Nursing Reference Center (October, 2008) using the words “multilevel analysis”, “random coefficient analysis” or “model,” “mixed effect models,” “covariance component model,” or “Hierarchical Linear Model”. While there are many *multilevel* studies (e.g., Adewale et al., 2007; Cho, 2003; Estabrooks, Midodzi, Cummings, & Wallin, 2007; Lake, 2006a; Lake, 2006b; Wu, 1995, 1997; Wu & Wooldridge, 2005) in the nursing literature, we did not locate a *multilevel CFA. In this paper we aim to demonstrate a multilevel CFA using NDNQI data as a case study.* We have rich data on the PES from 72,889 RNs within 4,783 nursing units in 2007. A by-product of this aim is that we are able to provide further validity evidence for the PES.

A multilevel CFA provides factor loadings explaining relationships *between* units (at the RN workgroup level) and loadings explaining relationships *within* units (at the individual RN level). This model can be thought of as a decomposition of the total covariance matrix into a between unit covariance matrix and a within unit covariance matrix. The one-way analysis of variance (ANOVA) can be used as a foundation for explaining multilevel CFA principles.

Multilevel Confirmatory Factor Analysis

For pedagogical purposes assume that the instrument of interest has nine items that make a two factor structure (Figure 1): $Y_1, \ldots, Y_4$ are four responses from latent factor η_1 (e.g., staffing and resource adequacy, SRA) and $Y_5, \ldots, Y_9$ are from latent factor η_2 (e.g. quality of care, QC). A popular assumption is that latent factors are standardized (mean=0, variance=1). According to this single-level two-factor analytic model, the factors are correlated as represented by a double arrow. The single-level factor analytic model has a covariance matrix that is constrained by the two factors. This single-level model ignores the variation between units and assumes that RNs within the same unit are independent. From a univariate modeling perspective, it is similar to defining an analysis of variance model (ANOVA) that only expresses variation from RN to RN as shown in Figure 1 – assuming Y is an arbitrary univariate item.

The single level is not an adequate representation of the variation from unit to unit. In particular, a more realistic ANOVA models the variation between units (Figure 2). In this ANOVA model, the total variation in Y is expressed as variation *between* units in addition to variation *within* units. This relative variation is measured by the ratio of the between variation and the total variation (intraclass correlation coefficient, ICC). The higher the ICC, the more similar are the responses among RNs within a particular unit – an assessment of the independence assumption. In cases where the items have an ICC>0.01, a multilevel (two-level)

CFA should be considered in lieu of a single-level (one-level) CFA. Ignoring the two level structure can lead to over estimates of the standard errors for CFA parameters (i.e., factor loadings) and inflation of Type I errors - which can lead to spurious results. Fitting a multilevel CFA leads to an analysis that involves both *within* latent factors and *between* latent factors. *Within* and *between* loadings assess validity for the RN as well as the unit.

For the ANOVA model, method of moments (MOM, e.g., Forbes & Taunton, 1994) is a popular approach for estimating the between and within variances (Figure 3.1). Succinctly, the ANOVA table elements (sum of squares within (SSW), sum of squares between (SSB), degrees of freedom (df), and mean squares (MSW or MSB)) are equated to their respective variance expected mean squares and solved to estimate within and between variances. The ICC is estimated directly from these variance estimates. Muthen (1989) proposed calculation for the two-level CFA that has analogs to the previous ANOVA calculations (Figure 3.2). First, a multivariate within covariance matrix is calculated by summing all of the covariance matrices within each of the units. Second, a covariance matrix is calculated using the data aggregated to the unit level. This produces a multivariate version of the ANOVA table, and the same algebra results in estimated between and within covariance matrices that can be used for estimating a two-level CFA. Muthen's approach, known as multi-level Muthen (MUML), enjoys good statistical properties (Yuan & Hayashi, 2005).

NDNQI RN Survey Data: Practice Environment Scale and Job Enjoyment

In this section we apply MLMU CFA to the 2007 NDNQI RN Survey data as a case study to demonstrate the approach, as well as establish validity of PES at both the RN and unit levels. First we discuss in more detail the PES and JE.

The PES-NWI (or just PES) was derived from the Nursing Work Index, an instrument originally designed to measure nurses' job satisfaction and perceptions of quality of care (Aiken & Patrician, 2000). The PES includes 31 items in five subscales that characterize the nature of professional nursing practice (Lake, 2002). Participants are instructed to “indicate the extent to which you agree that the item is PRESENT IN YOUR CURRENT JOB,” and provided a 4-option, Likert-type response scale, ranging from strongly disagree (1) to strongly agree (4). The subscales and example items follow.

- Nurse Participation in Hospital Affairs, 9 items: Opportunity for staff nurses to participate in policy decisions.
- Nursing Foundations for Quality of Care, 10 items: Working with nurses who are clinically competent.
- Nurse Manager Ability, Leadership, and Support of Nurses, 5 items: A nurse manager who is a good manager and leader.
- Staffing and Resource Adequacy, 4 items: Enough staff to get the work done.
- Collegial Nurse-Physician Relations, 3 items: Physicians and nurses have good working relationships.

Evidence supports the PES as a valid and reliable instrument for the measurement of the nursing practice environment (Lake, 2002).

The NDNQI-Adapted Job Enjoyment scale is 7 items that measure the extent to which nurses on the unit like their jobs in general (Taunton et al., 2004). The scale includes the stem “nurses with whom I work would say that,” and a 6-option, Likert-type scale, ranging from strongly disagree to strongly agree. An example item is: they find real enjoyment in their jobs.

Psychometric analysis indicates high reliability and robust validity at the individual and workgroup level (Boyle et al., 2006; Elliott, 2006; Taunton et al.).

The PES and JE were administered to 72,889 RNs from 4,783 RN workgroups (16 unit types, for example: critical care and obstetric) in 2007. Using Mplus software, we fit a multilevel confirmatory factor analytic model with a structure based on the five domains. From this model, we estimate the between unit loadings and within unit loadings to investigate construct and convergent validity at both the unit and RN levels. To investigate criterion-related validity, we correlated the five PES domains with the seven NDNQI Job Enjoyment items at the unit and RN levels (also using a multilevel model).

Results

Mplus summary statistics for PES items are presented in Table 1. Means range from 2.50 to 3.31, near the midpoint of the item's range (1-4). Standard deviations (STD) range from 0.58 to 0.67, slightly under what would be expected from items that vary across the entire response options (range/4=3/4=0.75). ICCs range from 0.08 to 0.28, with 90% of the items having ICCs larger than 0.10 and thus high enough to require a multilevel model (ignoring the ICC will result in 50% inflation of perceived information,see for example Campbell, Donner, & Klar, 2007). The ICCs presented by Mplus compare closely with those calculated using the formulas presented in Figure 3, as illustrated with *PesSR01*(see ANOVA, Table 2). The variation across units is significant ($F_{4747, 68141}=5.05$, $p<.001$) . The approximate average number of RNs per unit is $n \approx 72{,}889/4{,}783=15.24$. This leads to $\text{var}(W)=MSW=0.45$ and $\text{var}(B)=(MSB-MSW)/n=(2.27-0.45)/15.24=0.1194$ and finally $ICC=0.1194/(0.1194+0.45)=0.21$.

Mplus software generated both the single-level (traditional) and the two-level (multilevel) CFAs with output summarized in Table 3. The loadings are represented by B because they are essentially regression coefficients where an item is regressed on a domain. Because the domain is standardized (mean=0, SD=1), B is interpreted as the increase (or decrease) in the score of an item for every standard deviation increase in the domain. For example, in the single-level model B=0.45, saying for every standard deviation increase in Hospital Affairs, the item PESHA01 increases by 0.45. Therefore, among the HA items, PESHA06 has the highest association with B=0.63. Further, output from Mplus guides in testing statistical significance of B because the standard error (SE) can be used to calculate Z=B/SE. If Z>1.96 or Z< -1.96, then we would say that the item is statistically significantly associated with the domain (p<.05, *convergent validity).* From the single-level output in Table 3, it appears that all items are significantly associated with their respective domain. The results also suggest factorial validity (CFI=0.90 & RMSEA= 0.066).

However, as demonstrated by the ICCs (Table 1), responses of RNs in the same unit are correlated. A better CFA would be a two-level CFA (output in Table 3). Rather than just *within* B, the model also gives a *between* B. To understand the differences in inference for the B's of the two-level model relative to the single-level model, we turn our attention to Figure 4 where comparisons are made of within Bs and Zs (dark squares) and between Bs and Zs (white squares) from the two-level model versus within Bs and Zs from the one-level model. The trend is that the two-level model reduces the within Bs and Zs because it properly accounts for the correlation of RNs in the same unit. The within versus the between Bs and Zs drop substantially because the between Bs have fewer units than RNs. In all instances the within *convergent validity* is preserved, and further, the between *convergent validity* is established. A factorial validity is demonstrated in the two-level model (CFI=0.90 & RMSEA= 0.042).

The two-level CFA also provides correlations among the subscales both within and between (Table 4), which demonstrate adequate discrimination among subscales, with the exception of the between correlation of Hospital Affairs and Quality of Care (0.92). Others have documented problems with Quality of Care (Li et al., 2007).

The final step involved fitting a second-order CFA providing a super-domain called PES that is realized through the other domains while simultaneously fitting the JE model. This resulted in a between correlation of JE and PES of 0.93 and a within correlation of 0.77.

Discussion

For the CFAs fit in this case study we assume that the items are multivariate normally distributed. Given the ordinal nature of the data, ordinal models may be fit in Mplus (e.g., Grilli & Rampichini, 2007). We did not consider ordinal models in this paper because large sample sizes make these models cumbersome to run. Furthermore, the sample sizes are large enough, at both between and within levels, for the central limit theorem to justify inference from the normal methods as the parameter estimates are normally distributed.

One question a reader may have is: given the high between correlation of JE and PES, why not just use the simpler JE scale? One important reason is that the PES scales provide more actionable information for nurse administrators than the conceptually singular JE scale.

Conclusion

This paper demonstrated the utility of multilevel CFAs in the assessment of validity of the PES, both by accounting for the correlation of RNs within the same unit, and by supplying evidence of validity at both the RN and the unit levels of analysis.

References

Adewale, A.J., Hayduk, L., Estabrooks, C.A., Cummings, G.G., Midodzi, W.K., & Derksen, L. (2007). Understanding hierarchical linear models: Application in nursing research. *Nursing Research, 56*, S40-S46.

Aiken, L.H., & Patrician, P.A. (2000). Measuring organizational traits of hospitals: The Revised Nursing Work Index. *Nursing Research, 49,* 146-153.

American Nurses Association. (1995). *Nursing care report card for acute care.* Washington, DC: Author.

Boyle, D.K., Miller, P.A., Gajewski, B.J., Hart, S.E., & Dunton, N. (2006). Unit type differences in RN workgroup job satisfaction. *Western Journal of Nursing Research, 28*, 622-640.

Campbell, M.J., Donner, A., & Klar, N. (2007). Developments in cluster randomized trials and Statistics in Medicine. *Statistics in Medicine, 26*(1), 2-19.

Cho, S. (2003). Using multilevel analysis in patient and organizational outcomes research. *Nursing Research, 52*, 61-65.

Elliott, C.G.S. (2006). Using aggregated micro-level data as measures of macro-level phenomena: The case of the NDNQI-RN Satisfaction Survey. Unpublished Doctoral Dissertation, University of Kansas, Lawrence.

Estabrooks, C.A., Midodzi, W.K., Cummings, G.G., & Wallin, L. (2007). Predicting research use in nursing organizations: A multilevel analysis. *Nursing Research, 56*(4), S7-S23.

Forbes, S. & Taunton, R.L. (1994). Reliability of aggregated organizational data: An evaluation of five empirical indices, *Journal of Nursing Measurement, 2*(1), 37–48.

Goodwin, L.D. (2002). Changing conceptions of measurement validity: An update on the new standards. *Journal of Nursing Education, 41*(3), 100-106.

Grilli, L., & Rampichini, C. (2007). Multilevel factor models for ordinal variables. *Structural Equation Modeling, 14,* 1-25.

Kozlowski, S.W.J., & Klein, K.J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K.J. Klein & S.W.J. Kozlowski (Eds.), *Multilevel Theory, Research, and Methods in Organizations* (pp. 3-90). San Francisco: Jossey-Bass.

Lake, E.T. (2002). Development of the Practice Environment Scale of the Nursing Work Index. *Research in Nursing & Health, 25,* 17-188.

Lake, E. T. (2006a). Multilevel models in health outcomes research Part I: Theory, design, and measurement. *Applied Nursing Research, 19*(1), 51-53.

Lake, E. T. (2006b). Multilevel models in health outcomes research Part II: Statistical and analytic issues. *Applied Nursing Research, 19*(2), 113-115.

Li,Y., Lake, E.T., Sales, A.E., Sharp, N.D., Greiner, G.T., Lowy, E., Liu, C., Mitchell, P.H., & Sochalski, JA (2007). Measuring nurses' practice environments with the Revised Nursing Work Index: Evidence from registered nurses in the Veterans Health Administration. *Research in Nursing & Health, 30*, 31–44.

Montalvo, I. (2007). The National Database of Nursing Quality Indicators® (NDNQI®). *The Online Journal of Issues in Nursing, 12*(3).

Muthen, B. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54*, 557-585.

Nunnally, J.C., & Bernstein, I.H. (1994). *Pyschometric theory* (3rd ed.). New York: McGraw-Hill.

Taunton, R.L., Bott, M.J., Koehn, M.L., Miller, P.A., Rindner, E., Pace, K., Elliott, C., Bradley, K.J., Boyle, D., & Dunton, N. (2004). The NDNQI®-Adapted Index of Work Satisfaction. *Journal of Nursing Measurement, 12*, 101-122.

Verran, J.A., Gerber, R.M., & Milton, D.A. (1995). Data aggregation: Criteria for psychometric evaluation. *Research in Nursing & Health, 18,* 77-80.

Verran, J.A., Mark, B.A., & Lamb, G. (1992). Psychometric examination of instruments using aggregated data. *Research in Nursing & Health, 15,* 237-240.

Wu, Y.B. (1995). Hierarchical linear models: A multilevel data analysis technique. *Nursing Research, 44*, 123126.

Wu, Y.B. (1997). An application of hierarchical linear models to meta-analysis in nursing research. *Nursing Research, 46*, 295-298.

Wu, Y.B., & Wooldridge, P. (2005). The impact of centering first-level predictors on individual and contextual effects in multilevel data analysis. *Nursing Research, 54*, 212-216.

Yuan, K.H. & Hayashi, K. (2005). On Muthen's maximum likelihood for two-level covariance structure models. *Psychometrika, 70,* 147-167.

Figure Legends

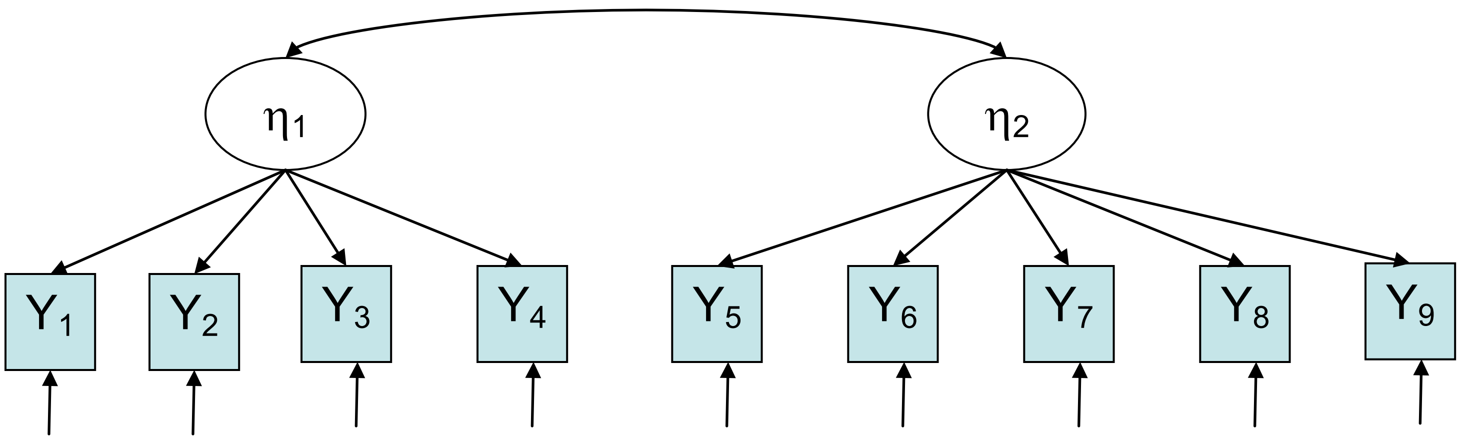
Figure 1. Classical factor analytic model with two factors.

Figure 2. Multilevel factor analytic model with two factors.

Figure 3. Estimating between and within variances in single and multilevel factor analysis

Figure 4. Comparing single to multilevel CFAs.

Figure 1.



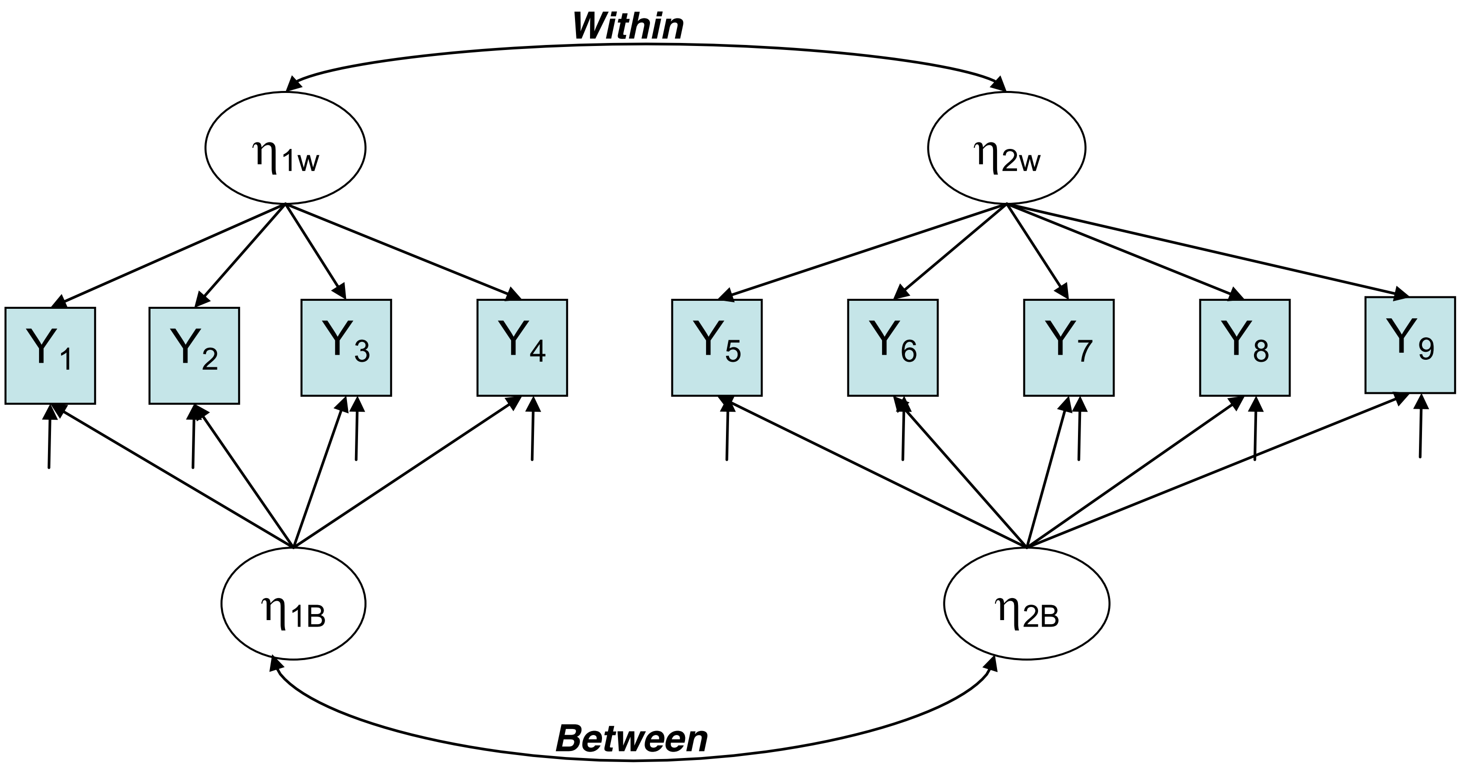
ANOVA analog

$Y = \mu + W$

μ is the overall mean of Y

var(W) is the variation from RN to RN

Figure 2.



ANOVA analog

$Y = \mu + B + W$

μ is the overall mean of Y

var(B) is the variation *between* units

var(W) is the variation from RN to RN *within* units

$ICC = \text{var}(B)/\{\text{var}(B)+\text{var}(W)\}$

Figure 3.

(1) ANOVA analog

SV	df	SS	MS	EMS	F
Between	J-1	SSB	MSB	var(W)+$n$var(B)	MSB/MSW
Within	N-J	SSW	MSW	var(W)	

Using method of moments (MOM)

var(W) ≈ MSW

var(B) ≈(MSB-MSW)/n

ICC = var(B)/(var(B)+var(W))

(2) Two-level confirmatory factor analysis (CFA) – Muthen Maximum Likelihood (MUML)

SV	df	SS	MS	EMS
Between	J-1	SS_B	S_B	$\Sigma_W + n\Sigma_B$
Within	N-J	SS_W	S_W	Σ_W

Using method of moments (MOM)

$\Sigma_W \approx S_W$

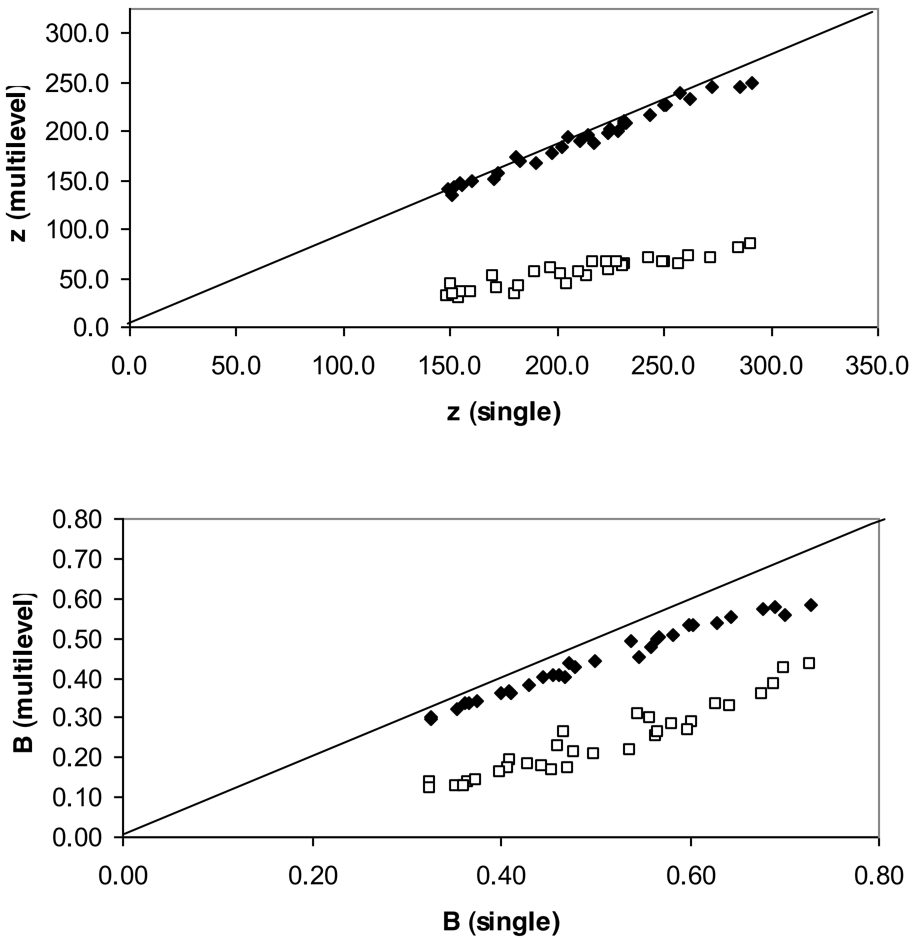
$\Sigma_B \approx (S_B - S_W)/n$

One can use MOM to approximate the *Within* and *Between* Covariance Matrices

- *Within* unit Confirmatory Factor Analysis
- *Between* unit Confirmatory Factor Analysis

Note: "J" is the number of units. "n" is the approximate number of RNs per unit. "N" is the total number of RNs.

Figure 4.



Dark squares are Within Bs or Zs, white squares are Between Bs or Zs.

Table 1. PES item summary statistics, n=15.24 and J=4,783.

Domain	Item	Mean	STD	ICC
	Nurse Participation in Hospital Affairs (HA)			
HA	PesHA01 career development-clinical ladder	2.89	0.73	0.20
	PesHA02 opportunity to participate in policy	2.68	0.77	0.15
	PesHA03 CNO visible and accessible	2.52	0.86	0.17
	PesHA04 CNO equal in power and authority	2.89	0.69	0.12
	PesHA05 opportunities for advancement	2.76	0.69	0.13
	PesHA06 administration listens and responds	2.58	0.81	0.18
	PesHA07 staff nurses involved in governance of hospital	2.87	0.70	0.15
	PesHA08 staff nurses serve on committees	3.14	0.58	0.12
	PesHA09 administrators consult with staff	2.50	0.80	0.14
	Nursing Foundations for Quality of Care (QC)			
QC	PesQC01 staff development for nurses	3.00	0.71	0.16
	PesQC02 high standards are expected	3.31	0.60	0.08
	PesQC03 clear philosophy of nursing	2.95	0.63	0.13
	PesQC04 nurses who are clinically competent	3.19	0.61	0.12
	PesQC05 active quality assurance program	2.97	0.58	0.12
	PesQC06 preceptor program	3.20	0.64	0.16
	PesQC07 nursing, not medical model	2.93	0.65	0.10
	PesQC08 up-to-date care plans	2.87	0.66	0.12
	PesQC09 patient assignments foster continuity	2.86	0.72	0.14
	PesQC10 use of nursing diagnoses	2.81	0.66	0.10
	Nurse Manager (NM) Ability, Leadership, and Support of Nurses			
NM	PesNM01 supervisors supportive	2.94	0.79	0.19
	PesNM02 mistakes as learning opportunities	2.83	0.75	0.15
	PesNM03 nurse manager good manager	3.00	0.87	0.25
	PesNM04 praise and recognition	2.69	0.80	0.17
	PesNM05 manager backs up staff	2.94	0.85	0.22
	Staffing and Resource (SR) Adequacy			
SR	PesSR01 adequate support services	2.70	0.75	0.21
	PesSR02 time to discuss patients with nurses	2.81	0.68	0.15
	PesSR03 enough RNs to provide quality care	2.56	0.84	0.27
	PesSR04 enough staff to get job done	2.51	0.80	0.28
	Collegial Nurse-Physician (NP) Relations			
NP	PesNP01 MD RN good relationships	3.07	0.62	0.16
	PesNP02 lot of team work	2.91	0.69	0.16
	PesNP03 RN MD collaboration	2.88	0.65	0.13

Table 2. The ANOVA for *PesSR01*.

Source of Variation	df	SS	MS	Expected Mean Square	F
Between	4747	10798.74	2.27	var(W)+$n$var(B)	5.05
Within	68141	30699.48	0.45	var(W)	

Table 3. Single- (traditional) & Two-level (multilevel) CFAs for manifest-latent variables.

		Single-level CFA			*Two-level CFA*					
					Within			Between		
		B	SE	Z	B	SE	Z	B	SE	Z
HA	BY									
	PESHA01	0.45	0.003	180.3	0.41	0.002	173.8	0.17	0.005	32.8
	PESHA02	0.57	0.003	224.7	0.50	0.002	203.2	0.25	0.004	57.5
	PESHA03	0.56	0.003	189.6	0.48	0.003	168.6	0.30	0.005	55.2
	PESHA04	0.41	0.002	170.3	0.36	0.002	150.5	0.19	0.004	50.2
	PESHA05	0.47	0.002	204.7	0.44	0.002	193.2	0.17	0.004	43.5
	PESHA06	0.63	0.003	243.1	0.54	0.002	216.8	0.33	0.005	69.3
	PESHA07	0.50	0.002	214.2	0.44	0.002	195.6	0.21	0.004	51.2
	PESHA08	0.37	0.002	182.8	0.34	0.002	169.6	0.13	0.003	40.5
	PESHA09	0.58	0.003	223.7	0.51	0.003	198.5	0.28	0.004	64.8
QC	BY									
	PESQC01	0.43	0.003	171.8	0.38	0.002	158.0	0.18	0.005	39.1
	PESQC02	0.33	0.002	150.8	0.30	0.002	135.2	0.14	0.003	43.1
	PESQC03	0.48	0.002	232.2	0.43	0.002	208.1	0.21	0.003	63.5
	PESQC04	0.33	0.002	149.2	0.30	0.002	141.8	0.12	0.004	31.3
	PESQC05	0.41	0.002	210.8	0.37	0.002	190.2	0.17	0.003	55.5
	PESQC06	0.35	0.002	154.8	0.32	0.002	147.7	0.13	0.004	29.1
	PESQC07	0.45	0.002	202.2	0.40	0.002	183.1	0.18	0.003	53.7
	PESQC08	0.37	0.002	159.8	0.34	0.002	149.1	0.14	0.004	35.1
	PESQC09	0.40	0.003	155.1	0.36	0.002	144.4	0.16	0.005	34.9
	PESQC10	0.36	0.002	152.0	0.34	0.002	143.1	0.13	0.004	32.9
NM	BY									
	PESNM01	0.64	0.002	261.8	0.55	0.002	232.9	0.33	0.005	72.0
	PESNM02	0.57	0.002	231.4	0.50	0.002	209.1	0.26	0.004	62.3
	PESNM03	0.69	0.003	250.8	0.58	0.003	227.8	0.38	0.006	65.9
	PESNM04	0.60	0.003	230.9	0.53	0.003	209.7	0.29	0.005	61.0
	PESNM05	0.68	0.003	249.7	0.58	0.003	226.0	0.36	0.005	65.6
SR	BY									
	PESSR01	0.55	0.003	216.9	0.45	0.002	188.0	0.31	0.005	66.0
	PESSR02	0.46	0.002	197.7	0.41	0.002	178.3	0.23	0.004	58.5
	PESSR03	0.73	0.003	285.3	0.58	0.002	244.5	0.43	0.005	80.4
	PESSR04	0.70	0.002	291.3	0.56	0.002	249.2	0.42	0.005	83.0
NP	BY									
	PESNP01	0.47	0.002	228.0	0.41	0.002	201.0	0.26	0.004	64.7
	PESNP02	0.60	0.002	272.1	0.53	0.002	246.1	0.27	0.004	70.2
	PESNP03	0.54	0.002	257.5	0.49	0.002	238.5	0.22	0.003	63.5

Table 4. Two-level CFA output for correlations among latent variables.

Within

	HA_W	QC_W	NM_W	SR_W	NP_W
HA_W	1	0.877	0.816	0.662	0.592
QC_W		1	0.734	0.666	0.643
NM_W			1	0.632	0.519
SR_W				1	0.484
NP_W					1

Between

	HA_B	QC_B	NM_B	SR_B	NP_B
HA_B	1	0.918	0.823	0.636	0.467
QC_B		1	0.799	0.718	0.527
NM_B			1	0.607	0.455
SR_B				1	0.406
NP_B					1