Measuring Work Stress among Correctional Staff: A Rasch Measurement Approach

George E. Higgins, Ph.d.
Associate Professor
Department of Justice Administration
University of Louisville
Louisville, KY 40292
(502) 852-0331
e-mail: gehigg01@gwise.louisville.edu

Richard Tewksbury, Ph.d.
Professor
Department of Justice Administration
University of Louisville
Louisville, KY 40292

Andrew Denney
Department of Justice Administration
University of Louisville
Louisville, KY 40292

Abstract

Today, the amount of stress the correctional staff endures at work is an important issue. Research has addressed this issue, but has yielded no consensus as to a properly calibrated measure of perceptions of work stress for correctional staff. Using data from a non-random sample of correctional staff (n=228), the Rasch model was used to assess whether a specific measure of work stress would fit the model. Results suggest that three items accurately represented correctional staff perceptions of work stress.

Measuring Work Stress among Correctional Staff: A Rasch Measurement Approach

Few will argue that working in a correctional facility is stressful. The work stress generally arises from interactions with inmates, superior officers/administrators, and perceptions role conflict/ambiguity. While, researchers have shown that correctional staff experiences a substantial amount of work stress (Auerback, Quick, & Pegg, 2003; Carlson, Anson, & Thomas, 2003; Finn, 1998; Griffin, Hogan, Lambert, Tucker-Gail, & Baker, 2010; Paoline, Lambert, & Hogan, 2006; Shaufeli & Peeters, 2000). To obtain information about work stress, researchers have used a number of approaches of conceptualizing and operationalizing the concept of work stress. For instance, Auerback et al. (2003) used a self-reported survey design measuring 57 separate items adapted from Sandman (1992) to measure job stress, producing 9 separate subscales (i.e. control, supervisor competence, supervisor communication skills, uncertainty, bureaucratic tape, time constraints, co-worker competence, co-worker communication skills, physical danger). Another example is where Paoline et al. (2006) conducted seven 2-hr focus groups with approximately 50 employees from all administrative levels of a jail for the future development of a questionnaire that was administered to 1,062 employees at nine separate jail facilities, measuring both job stress and job satisfaction. A final example is where Carlson, Anson, and Thomas (2003) conducted an Ordinary Least Squares (OLS) regression on 227 surveys of correctional officers from two different prisons that were separated into the categories of demographic variables (i.e. gender, marital status, etc.), and facility characteristics (i.e. male or female only, military experience, and rank). One measure that is consistently used to operationalize work stress in the literature is Cullen, Link, Wolfe, and Frank's (1985) six-item measure. This measure is a subscale of a larger stress measure. To date, this measure has been used in 15 studies where researchers have used the same exact items, or the researchers have

used borrowed some of the items (Dial, Downey, & Goodlin, 2010; Lambert & Paoline, 2010, Worley & Worley, 2011). Of the 15 studies, of these studies, a few researchers have presented true test score internal consistency items (Crank, 1991; Dial et al., 2010; Lambert, Hogan, Altheimer, Jiang, & Stevenson, 2010; Lambert & Hogan, 2010; Moon & Maxwell, 2004a; Moon & Maxwell, 2004b; Robinson, Porporino, & Simourd, 1996; Sundt & Cullen, 2002), and only one study has provided any psychometric assessment of the scale. Dial, Downey, and Goodlin (2010) used principal components factor analysis not examine the structure of the measure, but to create a standardized score on the scale across respondents. This lack of psychometric examination means that little is known about the Cullen et al. (1985) measure.

Understanding work stress among correctional officers is important because it provides insights into production and safety. Researchers have shown that correctional staff that has lower levels of work stress are more productive and actively work toward maintaining safety inside the correctional facility (Lambert, Hogan, & Griffin, 2008; Owen, 2006; Paoline et al., 2006; Schaufeli & Peeters, 2000). Without a psychometrically sound measure of work stress, correctional administrators are not able to provide the necessary programming or policies that may be able to reduce instances of work stress. At the present time, a gap exists in the literature that makes it paramount to gain an understanding of the psychometric properties of this measure.

To fill this gap in the literature, the Rasch model may be used to provide some important psychometric information. The basis of Rasch modeling comes from Thurstone's (1925, 1927) ideas of fundamental measurement. A central component of the Rasch model is that the items of a scale are supposed to be focused on developing a unidimensional or single dominant trait. When the items represent a unidimensional trait--in this study, work stress--the responses to items indicate the location on that trait. The Rasch model is a one parameter model for each

individual and one parameter for each item, and one set of category thresholds that apply to all items. These parameters can be measured and are used to represent the location on a common scale. Overall, these estimates may be used to determine the probability of individual *n* scoring in category *k* for item *i*. An important feature of the Rasch model is that it differs from true score theory. True score theory requires an abundance of combinations of items, and use multiple samples to determine reliability and validity of a measure of work stress. Rasch modeling does not rely on a combination of items. In fact, Rasch modeling provides an examination of each individual work stress item. In addition, Rasch modeling uses separation indexes to calibrate the sample independently of the measurement of item difficulty. This allows for the estimation of the location of an item on a continuum as it relates to similar items. In the context of the present study, a well-constructed work stress measure should have items that represent the continuum of work stress so that it can be used with other samples of correctional staff. The calibration and estimation process allows researchers to have an equal-interval scale of work stress to use in the correctional environment. To date, no study of this type has been produced using Cullen et al.'s (1985) measure of work stress; even though, it is a staple measure in the criminal justice literature.

Purpose

The purpose of the present study is to fill a gap in the literature because no study has examined the Cullen et al. (1985) measure of work stress, especially in the context of correctional staff using Rasch measurement. This study contributes to the Rasch measurement and criminal justice literature in several unique ways. First, this study examines how efficiently the correctional staff uses the response categories for the work stress measure. Second, the study examines how well the items and correctional staff fit the Rasch model. Third, the study

examines how well the items and correctional staff lie on the continuum of work stress. Fourth, the study examines the dimensionality and separation of the work stress items. Fifth, if any of these issues are deemed unsatisfactory to generally agreed on standards, some effort will be given to find a set of items to do provide satisfactory results.

Method

Procedures and Sample

The data for this study came of self-report surveys from all of the correlational staff that were employed at two medium security prisons in Kentucky. After Institutional Review Board (IRB) approval, all of the correctional staff members that have their primary appointment in one of the two institutions received an explanatory cover letter and the survey along with their paychecks in the spring 2004. The cover letter explained that participation in the survey was voluntary, and the letter informed the correctional staff that their responses would be confidential and anonymous. The cover letter, and the survey, instructed the correctional staff to return the surveys in one of the locked collection boxes that was located inside the main entrance for each institution. The locked collection boxes were kept in place for 2 weeks. To improve the number of completed surveys, the wardens at each institution assisted with data collection by sending either written or e-mail encouragement and reminders to the correctional staff to complete the survey.

These procedures resulted in a sample of 228 correctional staff members returning completed surveys from the 650 that were distributed; thus, the response rate was 35 percent. Table 1 presents the descriptive statistics for this sample. The sample was 68.1 percent male; had an average age or 42.5; and had an average of 9 years and 2 months experience working in corrections. Thirty-three percent of the sample had at least a college education. In the sample

34.5 percent worked in programs (i.e., unit management, education, chapel, and industries), 23.3 percent worked in security (i.e., correctional officer or supervisor), 19.5 percent worked in either medical or mental health services, 10.7 percent worked in administration, 7.1 percent worked in support services, and 4.9 percent worked in classification, unit management, or other.[1]

Insert Table 1 about here

Measures

The measurement for this study is the six items from Cullen et al.'s (1985) work stress scale. Cullen et al. (1985) designed the scale to examine how anxious or pressured officers felt while they were on duty, table 2 presents the items. Correctional staff marked their responses to these six-items using a 5-point Likert-type scale that was anchored by "strongly disagree" (1) to "strongly agree" (5). Higher scores on the scale indicated higher levels of work stress.

Insert Table 2 about here

Results

The perceptions of correctional staff work stress responses were analyzed using WINSTEPS (Linacre, 2009). The analysis begins with a focus on the examination of the category functioning using the methods outlined from Linacre (2002). The mean square statistics (MNSQ) is used to determine goodness-of-fit. The MNSQ compares the expected score against the actual score, in effect; this is a comparison of the hypothetical model against the observed data. Bond and Fox recommended that the MNSQ Bond and Fox (2001) argued that the MNSQ has a range from 0 to infinity. Smith's (2001) recommendation is used to determine the proper infit or outfit of the measures (i.e., below 0.77 or above 1.30). Others have argued that ZSTD or the cube root of the MNSQ may be used for additional information (Smith,

[1]The manner that our data were collected does not allow for disaggregating into (clerical staff, correctional officers, or correctional supervisors).

Schumacker, & Bush, 1998). In addition, the category functioning is examined for monotonic movement across the categories via the calibrations. In other words, the correctional staff use of the categories occurs in one general way, and the correctional staff does not use them out of order. If this does not occur, the misuse of the categories may be taken as information that additional refinement is needed.

Table 3 presents the rating scale analysis. The MNSQs show that the first category has an outfit--1.71--that is beyond the 1.30 standard. The fifth category has an outfit of 1.33 that is slightly above the 1.30 standard. In addition, the calibrations appear to be disordered. These results suggest that refinement is necessary to the rating scale.

Insert Table 3 about here

Then, the focus turns to the fit of the items. The MNSQ is used to determine the infit and outfit of the items (i.e., perceptions of work stress). The item fit measures whether the individuals are consistently using the items (i.e., some items may not be used at all). When the items do not fit the Rasch model, additional refinement is necessary.

Table 4 shows that items 3 that only one item does not fit the Rasch model. Item has an infit MNSQ and an outfit MNSQ that exceed the 1.70 standard. These results suggest that refinement is necessary for the work stress measure.

Insert Table 4 about here

Next, the person-item map is presented. This is a map of how well each of the items is located with the individuals Left of the center line is the placement of the persons, and right of the center line is the placement of the items. The letter *M* stands for the mean for both persons and items. When the *M* for persons and items are perfectly aligned, then the items target the

persons. The *M*s for this analysis are within 0.5 of a logit, suggesting the items are a reasonable target of the items (Bond & Fox, 2011).

Figure 1 presents the person-item map. This map shows that the mean for the persons does not match the mean for the items. In other words, the measure in its current form does not reflect the use of the items by the person. Supporting the item fit information, item 3 is above 1 standard deviation unit away from the mean for the items and above the mean for persons. Unlike the item fit information, the person-item map shows that item 2 is one standard deviation below the mean for the items and well below the mean for the persons. In other words, this item does not seem to work very well for most of the people in the sample. Item 4 appears to be capturing similar information as item 6. Overall, these results suggest that refinement to the original measure is necessary.

Insert Figure 1 about here

Next, the analysis moves to the dimensionality and reliability of the measure. The dimensionality is a presentation of the principal components analysis of the residuals. According to Reskage (1979), proper dimensionality is found when the principal components analysis of the residuals accounts for a substantial portion of the variance, and this gets at the central issue of whether the measure is unidimensional. The reliability of the measure is examined as the consistency that persons and items can be differentiated on an interval scale. This means that persons and items can be divided into strata representing separation. Smith (2001) argued that separation indexes that were 2.0 and above were sufficient.

Table 5 shows that the items in the measure account for a substantial portion of variance in the residuals. This means that the measure may be unidimensional, but the other issues that have been mentioned above suggest refinement. The reliability of the measures are important

and suggest that the items are not a good representation of the people. The separation is below the 2.0 mark that Smith (2001) suggested was the standard for persons. This information is supportive the interpretation of the person-item map.

Insert Table 5 about here

Finally, there is a presentation of the differential item functioning (DIF). The DIF determines whether there are differences between two groups. In this study, the differences between the two groups are males and females. DIF analysis results in a t-coefficient, and when the t-coefficient is over 2.0 then there are differences across the groups. This is another indication that there are differences between persons and items and refinement is necessary.

Table 6 shows that the items functions the same for males and females. The DIF analysis does not show that there are any differences between males and females. The t-statistic for the DIF analysis does not show that there is any difference between males and females. Overall, these results suggest that refinement is necessary.

Our refinement effort begins with the response categories. Table 7 shows a four category rating scale operates properly. The responses are monotonically moving, and the infit and outfit statistics indicate that there is proper fit.

Insert Table 7 about here

Next, we move on to item fit. Table 8 with the four category rating scale, three items fit the Rasch model. That is, no items were below 0.77 or above 1.70. This is an indication that these three items are possibly capturing work stress the best.

Insert Table 8 about here

Then, the analysis turns to the person-item map. Figure 2 shows the person-item map where the mean for items is equivalent to the mean of persons. This shows that the items are

capturing the sentiment of the persons. The items spread from the first standard deviation unit. This is supporting information that the three items are the best model to capture these data.

Insert Figure 2 about here

The dimensionality and reliability is then examined. Table 9 shows that the three items explain over 50% of the variance in the residuals. This means that the three items are better for examining work stress. In addition, the reliability is better. The person separation is above 2.0 suggesting that proper reliability has been achieved.

Insert Table 9 about here

The differential item functioning (DIF) is examined next. Table 10 presents the DIF analysis that shows that the items function the same for males and females. This is further evidence that these three items are proper items to capture work stress.

Discussion

A major area of research in criminal justice is correctional staff work stress. The problem with this research is that it uses a handful of measures. One measure that appears in the research literature is Cullen et al. (1985) version of work stress. The original scale contains six-items where respondents indicate their level of agreement using a 5-point Likert-Type scale that is anchored by the answer choices--(1) strongly disagree to (5) strongly agree.

Overall, the results of this analysis indicate that the original Cullen et al. (1985) work stress scale does not satisfactorily fit the Rasch model. Specifically, the response categories are not being used properly. The response categories were disordered meaning that there were bunches of response categories that did not seem to make sense to the Rasch model. In other words, the response categories were deemed inefficient because they were viewed as being equi-distant by the correctional staff. The lack of fit of the response categories resulted in at least one

item shows problems fitting the model. This means, at least this item, did not seem to represent the data very well. Specifically, the misfitting item and the poor use of the response categories were highlighted by the items not representing the persons (i.e., correctional staff) very well. In other words, the poor person reliability did not assist in suggesting that the original scale may fit the Rasch model. This was supported by the person-item map that showed more items seem to misrepresent the persons (i.e., correctional staff). While the scale seems to unidimensional and operate properly across males and females, the other results suggest that refinement is necessary. The research that has used this scale in its original form is not invalid, but the researchers' results are clouded by noise or error. In other words, the researchers are not able to say that work stress is accounted for in their models completely. The additional items and poor response categories suggest that a substantial amount of noise or error is likely present in their measurement of work stress.

A refined version of the scale does fit the Rasch model. Three items seem to reach the epicenter of work stress using the scale (i.e., I am usually clam and at ease when I am working; I usually feel that I am under pressure when I am at work; There are a lot of aspects of my job that can make me pretty upset about things). Researchers seeking to use an efficient measure of work stress should consider these three items. Further, researchers using these three items should reduce their response categories from 5 to 4. This reduction created a better flow of responses to the items. The person-item map shows that the three items are less noisy or cleaner measurement where the items capture the sentiment of the correctional staff. This means that the three items targeted work stress better. The refined scale is unidimensional and has proper levels of separation. All of this suggests that the refined scale is valid and reliable via the standards of the Rasch model. In other words, researchers wishing to use a more efficient measure for work

stress among correctional officers are able to use the three items in this study. Further, the differential item functioning (DIF) shows that the measure works well for males and females. Overall, the refined 3-item scale seems to fix the problems that occurred with the 6-item scale.

The results of this study should be kept within the confines of its limits. First, the study does not address the changes in work stress by using longitudinal data. However, this study is the first to examine the psychometric properties of the Cullen et al. (1985) version of work stress. Second, the study sample is limited to data from correctional staff in one state. Regional variation may contribute to differential item function.

Despite the limits, the Cullen et al. (1985) 6-item work stress scale has provided an excellent starting point for understanding the issues that are relevant to correctional staff. Unfortunately, the results of this study show that the 6-item version of the scale that uses a 5-point Likert-Type scale does not fit the Rasch model very well. However, a refined 3-item work stress scale that uses a 4-point Likert-type scale does fit the Rasch model.

Table 1. Demographic Profile of the Sample

Measure	Mean or Percent	Standard Deviation
Male	68%	---
Age	42.5	9.16
Education		
High School Diploma	29.3%	---
Some College No Degree	2.1%	---
Associates Degree	8.0%	---
Bachelor Degree	10.3%	---
Graduate Degree	23.9%	---
Type of Position		
Security	23.3%	---
Programs	34.5%	---
Medical or Mental Health	19.5%	---
Administration	10.7%	---
Support Services	7.3%	---
Other	4.9%	---

Table 2: Items from the Work Stress Scale with Numerical Names

Item Number	Item
1	When I'm at work, I often feel tense or uptight.
2	A lot of the times, my job makes me very frustrated or angry.
3	Most of the time when I am at work, I don't feel that I have much to worry about.
4	I am usually calm and at ease when I am working.
5	I usually feel that I am under pressure when I am at work.
6	There are a lot of aspects of my job that can make me pretty upset about things.

5-point Likert-Type Scale: 1 = Strongly Disagree and 5 = Strongly Agree

Table 3. Summary of Rating Scale Analysis.

Label	Scale	Infit MNSQ	Outfit MNSQ	Structure Calibration
1	1	1.55	1.71	None
2	2	0.80	0.77	-2.57
3	3	0.77	0.84	0.59
4	4	0.90	0.89	-0.51
5	5	1.23	1.34	2.50
Missing				

Note: MNSQ = Mean Square--Standard 0.77 to 1.30

Table 4. Misfitting Items for the Work Stress Scale

Entry Number	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	PTMEA CORR
3	1.76	6.90	1.85	7.00	0.57
1	1.07	0.70	1.11	1.10	0.65
2	0.99	0.60	1.03	0.30	0.72
5	0.73	3.20	0.86	6.50	0.74
4	0.78	2.60	0.76	2.60	0.77
6	0.72	3.40	0.68	3.50	0.76

Note: MNSQ = Mean Squares; ZSTD = z-score; PTMEA CORR = Parameter Correlation

Figure 1. Person-Item Map for 6 Items from Work Stress Scale.

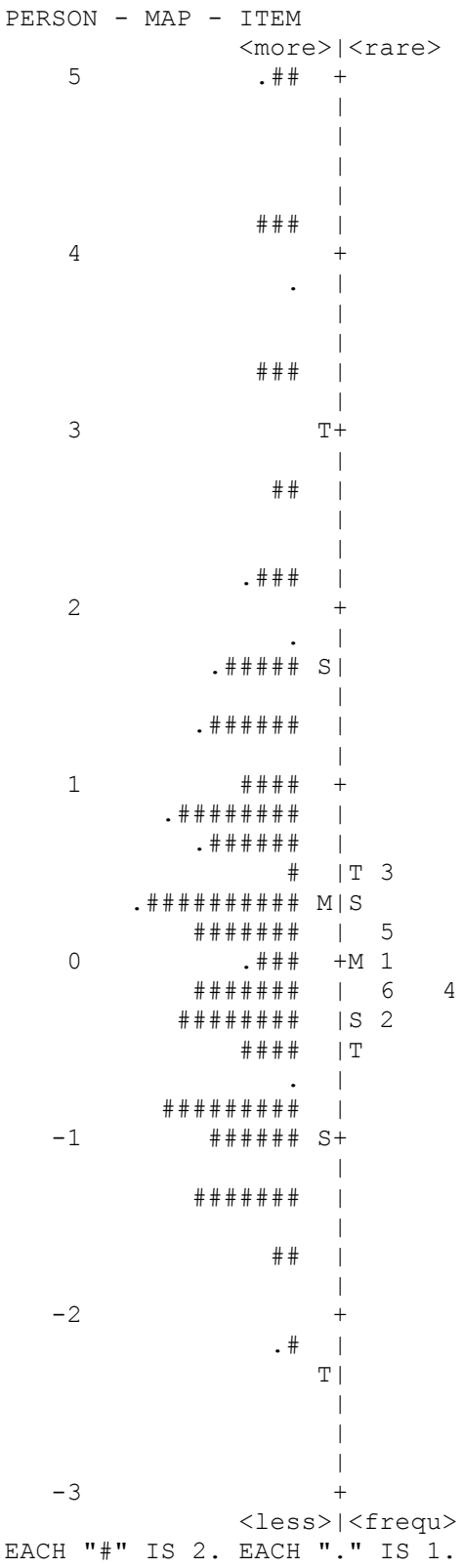

Table 5. Dimensionality, Separation, and Reliability.

Panel A:	Residual Variance Explained	
	48.30%	
Panel B:	Person Separation	Item Separation
	1.77	3.04
Panel C:	Person Reliability	Item Reliability
	0.76	0.90

Table 6. The Differential Item Functioning From the Rasch Model.

Person C #1	DIF Measure	DIF SE	Person C#2	DIF Measure	DIF SE	DIF Contrast	Joint SE	t	df	Item #
0.00	0.37	0.16	1.00	0.15	0.10	0.22	0.19	1.16	141	1
0.00	0.32	0.15	1.00	0.59	0.11	0.27	0.19	1.44	145	2
0.00	0.15	0.16	1.00	0.20	0.10	0.35	0.19	1.87	138	3
0.00	0.12	0.16	1.00	0.09	0.10	0.03	0.19	0.14	142	4
0.00	0.20	0.16	1.00	0.18	0.10	0.03	0.19	0.15	141	5
0.00	0.17	0.16	1.00	0.32	0.10	0.14	0.19	0.77	142	6

Note: DIF = Differential Item Functioning; *p<0.05

Table 7. Summary of Rating Scale Analysis for Modified Work Stress Scale.

Category Label	Scale	Infit MNSQ	Outfit MNSQ	Structure Calibration
2	2	1.01	0.93	None
3	3	0.95	0.93	-2.06
4	4	0.95	0.96	-1.54
5	5	1.12	0.90	3.60
Missing				

Table 8. Misfitting Items for the Modified Work Stress Scale

Entry Number	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	PTMEA CORR
5	1.14	1.30	1.03	0.30	0.84
6	0.92	0.80	0.86	1.00	0.87
4	0.91	0.80	0.90	0.70	0.88

Note: MNSQ = Mean Squares; ZSTD = z-score; PTMEA CORR = Parameter Correlation

Figure 2. Person-Item Map for the Modified Work Stress Scale.

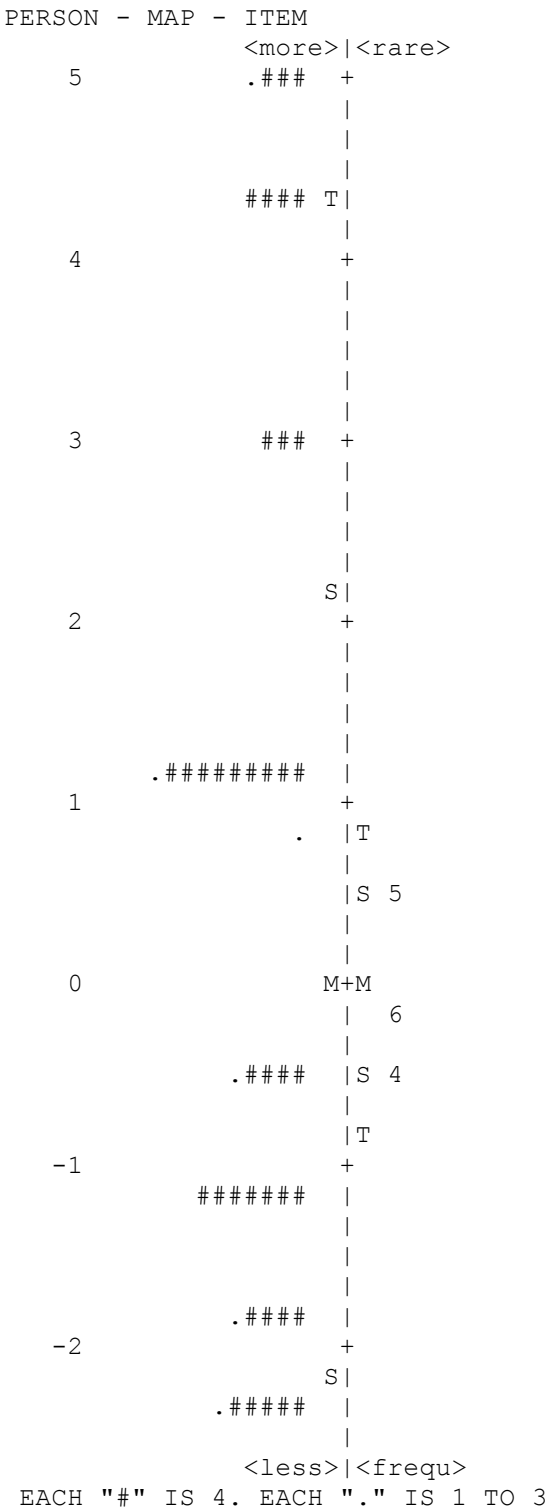

Table 9. Dimensionality, Separation, and Reliability for the Modified Work Stress Scale.

Panel A:	Residual Variance Explained	
	56.90%	
Panel B:	Person Separation	Item Separation
	2.04	2.67
Panel C:	Person Reliability	Item Reliability
	0.81	0.83

Table 10. The Differential Item Functioning From the Rasch Model.

Person C #1	DIF Measure	DIF SE	Person C#2	DIF Measure	DIF SE	DIF Contrast	Joint SE	t	df	Item #
0	0.24	0.25	1	0.05	0.16	0.19	0.30	0.64	100	4
0	0.55	0.25	1	0.58	0.16	0.03	0.29	0.29	101	5
0	0.30	0.25	1	0.54	0.16	0.24	0.30	0.30	102	6

Note: DIF = Differential Item Functioning; *p<0.05

References

Auerbach, S.M., Quick, B.G., & Pegg, P.O. (2003). General job stress and job-specific Stress in juvenile correctional officers. *Journal of Criminal Justice*, *31,* 25-36.

Carlson, J.R., Anson, R.H., & Thomas, G. (2003). Correctional officer burnout and stress: Does gender matter? *The Prison Journal, 83*(3), 277-288.

Crank, J.P. (1991). Work stress and job dissatisfaction in the public-sector: An examination of public safety directors. Social Science Journal, 28(1), 85-100.

Cullen, F., Link, B., Wolfe, N., & Frank, J. (1985). The social dimensions of correctional officer stress. *Justice Quarterly, 2,* 505-533.

Dial, K.C., Downey, R.A., & Goodlin, W.E. (2010). The job in the joint: The impact of generation and gender on work stress in prison. *Justice Quarterly, 38,* 609-615.

Finn, P. (1998). Correctional officer stress a cause for concern and additional help. Federal Probation, 62(2), 65-74.

Griffin, M.L., Hogan, N.L., Lambert, E.G., Tucker-Gail, K.A., & Baker, D.N. (2010). Job involvement, job stress, job satisfaction, and organizational commitment and the burnout of correctional staff. *Criminal Justice and Behavior, 37*(2), 239-255.

Lambert, E.G., & Hogan, N.L. (2010). Work-family conflict and job burnout among Correctional staff. Psychological Reports, 106(1), 19-26.

Lambert, E.G., Hogan, N.L., Altheimer, I., Jiang, S., & Stevenson, M.T. (2010). The relationship between burnout and support for punishment and treatment: A preliminary examination. *International Journal of Offender Therapy and Comparative Criminology*, *54*(6), 1004-1022.

Lambert, E.G., Hogan, N.L., & Griffin, M.L. (2008). Being the good soldier: Organizational citizenship behavior and commitment among correctional staff. *Criminal Justice and Behavior, 35*(56), 56-68.

Lambert, E.G., & Paoline, E.A. (2010). Take this job and shove it: An exploratory study of turnover intent among jail staff. Journal of Criminal Justice, 38, 139-148.

Moon, B., & Maxwell, S.R. (2004a). The sources and consequences of corrections officers' stress: A South Korean example. *Journal of Criminal Justice, 32,* 359-370.

Moon, B., & Maxwell, S.R. (2004b). Assessing the correctional orientation of corrections officers in South Korea. *International Journal of Offender Therapy and*

Comparative Criminology, 48(6), 729-743.

Owen, Stephen S. (2006). Occupational stress among correctional supervisors. *The Prison Journal, 86*(2), 164-181.

Robinson, D., Porporino, F.J., Simourd, L. (1996). Do different occupational groups vary on attitudes and work adjustment in corrections? Federal Probation, 60(3), 45-53.

Sandman, B. (1992). The measurement of job stress: Development of job stress index. In C. Crann, P. Smith, & E. Stone (Eds.), *Job satisfaction: How people feel about their jobs and how it affects their performance.* New York: Lexington Books.

Schaufeli, W.B., & Peeters, M.C.W. (2000). Job stress and burnout among correctional officers: A literature Review. *International Journal of Stress Management, 7*(1), 20-48.

Sundt, J.L., & Cullen, F.T. (2002). The correctional ideology of prison chaplains: A national survey. Journal of Criminal Justice, 30, 369-385.

Worley, R.M, & Worley, V.B. (2011). Guards gone wild: A self-report study of Correctional officer misconduct and the effect of institutional deviance on "care" within the Texas Prison System. *Deviant Behavior, 32*(4), 293-319.