

FROM THE EDITOR'S DESK

International Journal of Health & Productivity (IJHP) Expands its Thought Leadership Role

ARTICLES

EAP Works: Global Results from 24,363 Counseling Cases with Pre-Post Data on the Workplace Outcome Suite® (WOS)

Demonstrating Value: Measuring Outcome & Mitigating Risk: FOH EAP Study Utilizing the Workplace Outcome Suite®

Development and Validation of a Critical Incident Outcome Measure

Validation of the 5-item Short Form Version of the Workplace Outcome Suite®

Measuring Coaching Effectiveness: Validation of the Workplace Outcome Suite® for Coaching



Measuring Coaching Effectiveness: Validation of the Workplace Outcome Suite for Coaching

Richard D. Lennox, Ph.D; David Sharar, Ph.D; Francine Miller

ABSTRACT

KEY WORDS:

Workplace Outcomes, WOS,
Coaching Call Centers,
Disease Management,
Depression, Work Engagement,
Absenteeism, Presenteeism,
Workplace Distress,
Life Satisfaction

Chestnut Global Partners developed the coaching version of the Workplace Outcome Suite (WOS) to measure the effectiveness of coaching interventions in helping to address a range of workplace issues. For this purpose the original 25-item version of the WOS, developed in 2009, was slightly modified to provide outcome information on the coaching process. While the changes were modest, there was some concern that they might adversely affect the psychometric characteristics of the 25-item tool compromising its validity and reliability. Prudence requires the coaching version of the WOS be evaluated to assure results are equivalent with those of the original WOS. Data from 309 clients, provided with disease management coaching services for depression and diabetes, were analyzed. Various methods, including reliability analysis, confirmatory factor analysis, and correlational studies, were included in the study. The results of these analyses showed that the coaching tool performed slightly better than the original WOS. The instrument was shown to be extremely reliable, particularly for a short scale. The analysis revealed that the 25-item coaching version of the WOS can be used to test the efficacy and effectiveness of a coaching program without concern for significant measurement error. Results suggest that psychometric studies of the original 25-item WOS can also reflect clinical change from the workplace health programs coaching version even with samples as small as 50 clients.

INTRODUCTION

A growing literature offers promising support for the efficacy of coaching-based disease management (DM) programs for many chronic diseases, including heart disease, hypertension, depression, and diabetes.¹⁻² This research demonstrates that organizing EAP call centers for a case-management strategy can produce statistically positive clinical changes over time. Moreover, the use of randomized clinical trial (RCT) designs in many of these studies has provided convincing evidence that the structured intervention contributes to improvement in employees' clinical status. However, this type of research does have its critics. Lewis and Khanna assert these studies to be methodologically flawed and unable to claim that the DM approach is effective.³

Despite these concerns, recent corporate-based disease management programs do support coaching to help mitigate employer

costs of such chronic conditions such as depression, diabetes, and heart disease.⁴ While focusing on absenteeism, productivity and company-based health insurance costs, these programs use 24-hour call centers and case management behavioral support services to encourage medication compliance, treatment adherences, and other wellness activities. Studies have found that workplace disease management improves clinical status while also improving presenteeism.⁵ In addition long-term outcome studies find there are benefits in the area of reduced health care utilization.¹⁻²

Impact studies of coaching-based DM programs have explored specific effects within the health services systems.¹ For example, research has sought to establish the efficacy of DM programs in terms of their ability to improve specific clinical disease markers with less attention given to the

complimentary impact these programs have on workplace functioning and general health.¹ Often, this emphasis has been accompanied by the use of Randomized Controlled Trial (RCT) methodologies that place a high value on being able to assert casual direction at the cost of a broader, longer range analysis of workplace functioning and health care costs. With the proliferations of coaching in EAP programs, it would be prudent to examine various coaching models within the context of broader long-term effectiveness research. Thus far, only short-term efficacy models have been designed to isolate evidence that an intervention is capable of producing change in clinical markers.

There has been a growing interest in DM programs to treat depression and other mental health problems.² Recent RCT studies have reported that DM programs targeting depression resulted in reduced symptoms, improved job retention, increased work productivity, and reduced absenteeism.⁶⁻⁸ Although these results are very encouraging, little is known about the inter-relationships among improvements in health and improvements in workplace functioning. That is, are changes in depression symptoms associated with reductions in absenteeism and presenteeism, and are those changes associated with reductions in future health care costs?

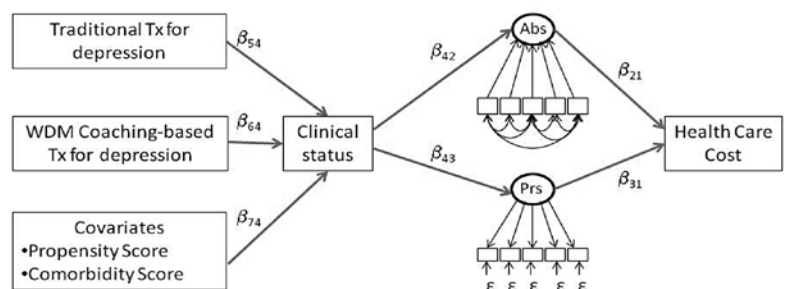
Most existing evaluations of coaching programs were not designed to also assess comprehensive workplace outcomes. Schneider et al.,⁹ conducted an evaluation of a coaching program for managing diabetes but the evaluation was limited to satisfaction with the program and clinical goal attainment. Unfortunately, it did not include a control group. Merrell et al.,¹⁰ designed a wellness program clinical outcome study, but it did not include a coaching component, workplace oriented functional measures, or a control group. Hawkes et al.,¹¹ conducted a randomized clinical trial of a scripted telephone coaching program for treating cardiovascular disorders. The study included measures of quality of life from the Short Form 36-item (SF36) Health Survey and level of physical activity but did not ask about workplace functioning or life satisfaction. Linden et al.,¹² employed an interview-based coaching approach comparing participants to

non-participants. They found higher levels of self-efficacy, patient activity, lifestyle change, perceived health status, and lower levels of risk over time. Another researcher tested the effectiveness of a Body Mass Index (BMI) coaching program, finding improvement in weight loss, dietary habits, and quality of life as measured with the SF36 and life satisfaction.¹³ Again, none of these studies examined the impact of these coaching changes on employees' workplace outcomes. Patja et al.,¹⁴ also tested the effectiveness of telephonic coaching for the self-care of hypertension disease in a randomized clinical trial. Results, however, failed to support the hypothesis when testing for primary clinical outcomes, quality of life, or work functioning measures.

While some studies were unable to find consistent support for coaching, the Margolis et al. study¹⁵ of hypertension treatment found an interesting frequency response effect indicating that the more coaching sessions a patient received, the better the outcomes. These studies suggest that the DM coaching field could benefit from standardized outcome measures with which to test for the efficacy and effectiveness of various health promotional activities.

Figure 1 illustrates a model for using structural equation modeling with latent variables to examine the impact of coaching for depression and its impact on future health care costs.¹⁶ The model shows how three exogenous variables related to coaching sessions and propensity scores capture pre-existing differences between a treatment group and a control group on clinical measures of depression and its associated effect on absenteeism and presenteeism. The various β terms in the figure illustrate the structural coefficients that can be isolated by

Figure 1: A Path Model of Coaching-Based Depression DM



the model.

For example, β_{21} and β_{31} , depict how absenteeism and presenteeism contribute to future health care costs and how mitigating the effects of depression by coaching assignments can impact health care costs through the other variables in the model. The primary objective of this type of model would be to test the effectiveness of DM on three classes of healthcare cost outcome measures:

- (1) **proximal** clinical outcomes,
- (2) **medial** workplace outcomes, and
- (3) **distal** future workplace outcomes.

The Structural Equation Model (SEM) presented in Figure 1 also illustrates the systemic impact that depression and diabetes DM programs can have on the workplace. The path model shows that treatment for depression contributes to reducing depression symptoms and impacting workplace functioning in terms of absenteeism and presenteeism.¹⁷

The Workplace Outcome Suite was designed to evaluate the effectiveness of EAP services. It was validated from studies on the effectiveness of traditional call center and other therapeutic intervention programs but not on coaching interventions per se.¹⁸

WHAT IS THE WORKPLACE OUTCOME SUITE?

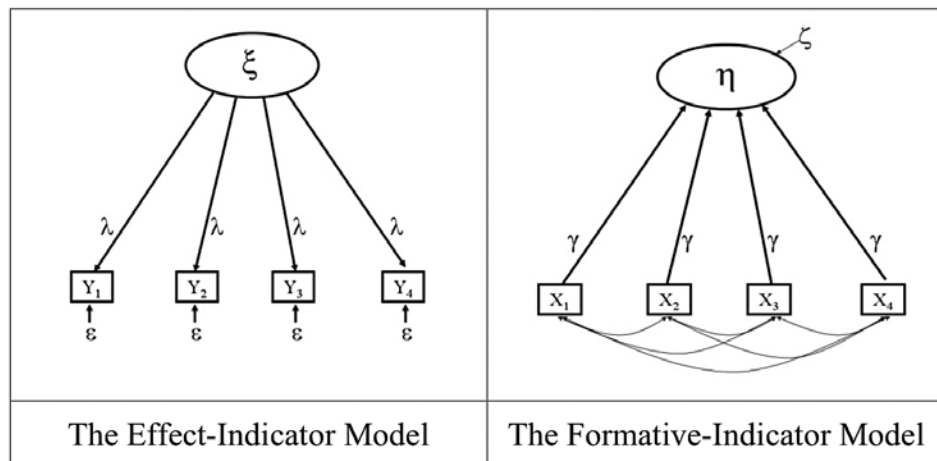
The original Workplace Outcome Suite (WOS) is an easy-to-administer tool developed by Chestnut Global Partners' (CGP) Division of Commercial Science. It

uses a short, precise, and easy-to-administer survey that collects EAP specific outcome data both before (pre – at start of the counseling) and after (post – usually after three months) EAP services. Thus, the WOS is a measure of change that examines five key aspects of workplace functioning: absenteeism, presenteeism, work engagement, workplace distress, and life satisfaction. Consult the 2017 WOS Report for a discussion of how it is administered, basic calculation methods for change over time, and the history of the development of the full 25-item, 9-item, and brief 5-item versions.¹⁹

Unlike most prior workplace outcome measures, the scales in the WOS were designed using a common theoretical framework built around more than 100 years of applied psychometric research and practice. The result is a set of 25-item, 9-item, and brief 5-item versions designed around the same measurement theory and validated using the same set of psychometric principles.²⁰⁻²¹ Starting with the explicit measurement model that Bollen prescribed for scales, a 5-item version was constructed to facilitate obtaining results on services provided in a short amount of time.¹⁶

The original WOS was constructed on the basis of two formal measurement models: an effect-indicator model and a formative measurement model.²² As shown in **Figure 2** in the effect-indicator model, the measured indicators are thought to emerge from a single underlying latent variable (illustrated by the ellipse). The small

Figure 2: Formal Measurement Models



arrows under the boxes represent the random errors in the items that are effectively removed by adding the items together. The mechanics of the effect indicator model can be characterized as: $Y_i = \lambda_{i1}\xi_1 + \epsilon_i$ [1]

Where Y_i is the i^{th} item in the scale, ξ_1 is the single latent construct presumed to underlie the covariance among the item set, λ_{i1} is the factor loading of the i^{th} item on the single latent factor, and ϵ_i is the random measurement error in the i^{th} item. The random measure in ϵ_i is at the heart of the effect-indicator model and provides the basis for offsetting random errors. Errors are expected to be half in the positive direction and half in the negative direction, thus offsetting one another in the sum or average of the items.

The formative measurement model uses different items that combine linearly to create a broad construct. The right side of Figure 3 illustrates the manner in which the items combine to form the latent construct. The formative model can be expressed as:

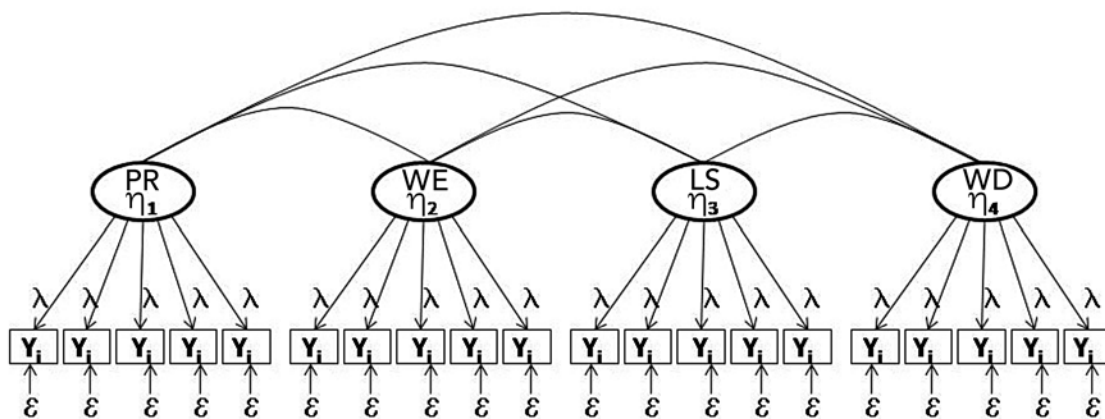
$$\eta_1 = \gamma_{11}X^1 + \gamma_{21}X^2 + \gamma_{31}X^3 \dots \gamma_{i1}X^i + \zeta_1$$
 [2]

Where η_1 is the latent construct and $\gamma_{11}X^1 + \gamma_{21}X^2 + \gamma_{31}X^3 \dots \gamma_{i1}X^i$ are coefficients that indicate the contribution each item makes to the formulation of the latent construct in a manner similar to multiple regression. The ζ_1 term reflects the variance in the latent construct not explained by the specific set of indicators. Notice that there is no random measurement error for the individual items. For this reason, coefficient alpha is not considered appropriate for

assessing reliability in this type of model.^{16,22,23}

Results of the psychometric analysis of the WOS provide initial support for its reliability, structural validity, and construct validity. Two separate validation studies of the suite, one completed using paper-and-pencil ($N = 220$) and the other conducted via a telephone interview ($N = 228$), tested the reliability of the scales, the structural validity of the items, and the construct validity of the unit-weighted scale scores. Note that the Absenteeism scale is based on a formative measurement model that does support coefficient alpha as an index of reliability or factor analytic procedures, so it was not included in these analyses. The effect-indicator scales were found to have moderate (coefficient alphas in the range of .75) to excellent (coefficient alphas in the range of .90) levels of internal consistency. All of the items produced factor loadings of .30 or higher in the paper-and-pencil study. Only two items in the Work Engagement scale produced low factor loadings in the telephone interview study. However, since these low loadings did not replicate in the paper-and-pencil study, they were considered spuriously low. Correlations of the scale scores with self-reported measures of relevant behavior and emotions provided limited evidence of construct validity for all five scales. The results suggest support for use of the WOS to evaluate EAP services and interventions. The coefficient alphas for the telephone interview study were similar to those of the Paper and Pencil Study. Item-total correlations for all

Figure 3: Theoretical Structure of the Effect-Indicator Measurement Models of the Presenteeism, Work Engagement, Life Satisfaction and Workplace Distress Scales



20 effect-indicator items were substantial and significant beyond $p < .01$. Analysis of the alpha coefficients created after iteratively removing a single item showed that there was very little improvement to be gained by removing any one item. There are some items in the Work Engagement scale that suggest some room for improvement. This point will be better explained in the context of the confirmatory factor analysis in the next section.

We conducted a confirmatory factor analysis on the two data sets to test the overall structure of the items as they relate to their respective latent constructs and one another. This analysis allowed us to evaluate the item-level internal consistency and discriminant validity of the items in terms of loading on their respective factor but not on other factors. In confirmatory factor analysis we test the hypothesis that the item defines its appropriate factor and is unrelated to the other factors. This is achieved by fixing the “off factor” loading to zero and testing the goodness-of-fit for the model of the four-correlated factors. **Figure 3** illustrates this factor structure with the four ellipses representing the respective latent factors for Presenteeism, Work engagement, Life Satisfaction, and Workplace Distress.

In a previous study we tested the model using Maximum Likelihood estimation. For the telephone interview sample of 228 respondents, modeling the observed covariance matrix to the hypothesized four-factor model yielded a Bentler-Bonnet fit index of .86, a Bollen fit index of .88, and a comparative fit index of .88. Similarly, modeling the observed covariance matrix from the paper-and-pencil data set ($N = 220$) yielded a Bentler-Bonnet Fit index of .83, a Bollen fit index of .86, and a comparative fit index of .85. Both models yield respectable fits to their respective covariance matrices.

The results support the general fit of the 25-item scale to the hypothesized structure of the four effect-indicator latent variables. This paper presents a validation of the slightly revised version of the original Workplace Outcome Suite,¹⁸ after being modified for use with coaching interventions. Changes were focused on instructions to the respondents as they relate to coaching rather than to EAP services. All 25 items and sub-scales

in the coaching version of the WOS and the validations is focused on the suitability of the revised instruction. The basic methodology is the same as was used in validating the original WOS.

This article tests the validity of using the previously validated Workplace Outcome Suite (WOS) slightly modified, for use with coaching interventions.¹⁸

METHOD

Data Sources

Data for the study were provided by two EAPs offering coaching services as part of their program. In Study I, Chestnut Global Partners provided data from 309 clients using a coaching service as part of the disease management program for depression or diabetes. In Study II, the Well Call Center provided data from 311 clients also using a coaching approach.

RESULTS AND DISCUSSION

Analysis

Descriptive statistics. Central tendency and dispersion in the individual items in the Workplace Outcome Suite will be examined for severe departure from normality in distributions that are expected to be normal, which might create problems for traditional statistical analysis. Variation around the mean will also be monitored for evidence of small variance and thus limited discriminability and/or large variance that might indicate confusion and cause excessive measurement error.

Reliability analysis. The basic structure of the effect-indicator measurement models will be assessed using an analysis of internal consistency. Under the assumption of the traditional unidimensional effect-indicator model, the reliability of a multi-item scale of parallel items can be used to assess the reliability of scale scores as indicated by the common variance reflected in the first factor. Coefficient alphas will be used to assess the internal consistency of the respective item-sets in the individual scale.²⁴ The alpha coefficient will be used as a standard for evaluating the random measurement error of the individual scales.

Confirmatory factor analysis. Scales measuring Presenteeism, Work Engagement, Life Satisfaction, and Workplace Distress are

based on the effect-indicator models and are expected to define a unidimensional factor that reflects their respective underlying construct. **Figure 3** illustrates the theoretical structure of the four effect-indicator latent variables. The figure shows that the individual scales are thought to emerge from their respective single underlying constructs, their unique variance is considered to be random measurement error, and the four latent constructs are expected to be intercorrelated with one another.

In contrast to the absenteeism that assesses the individual components of total time away from work, these four items contain highly intercorrelated elements that share a common factor reflected in the common variance. Confirmatory factor analysis is used to test the internal structure of the 20 items in these scales as being capable of defining four correlated, yet unique factors within their covariance matrix. Maximum likelihood estimations are used to fit a four-factor model to the covariance matrix, with a χ^2 test of goodness-of-fit used to test the residual covariance matrix against the model matrix. It is hypothesized that the four-correlated-factor model will account for the majority of the reliable variance in the covariance matrix. We also expect that standardized factor loading connecting the individual items to their respective latent variable will be significant and substantial.

Bivariate correlations of scale scores. To provide some limited construct validity we correlated scale scores with some self-reported pseudo-behavioral and emotional measures that are expected to correlate with the constructs we are measuring. For example, the item “I have a hard time getting out of bed” negatively correlates with work engagement because the engaged worker would be passionate about getting to work. A positive correlation or an insignificant correlation might indicate a problem with our work engagement measures. Similarly, the item “I often feel sad” would be expected to correlate positively with workplace distress because the upset employee would be expected to feel sad even when away from the job. The following items are used as pseudo-criterion related targets for the five individual outcome measures.

- *I have a hard time getting out of bed* (i.e.,

item should correlate negatively with Life Satisfaction).

- *I often feel sad* (i.e., item should correlate positively with Workplace Distress).
- *I keep falling behind schedule at work* (i.e., item should correlate negatively with Work Engagement).
- *I am rarely late for work* (i.e., should correlate negatively with Absenteeism)
- *I often get home late from work.* (i.e., should correlate positively with Work Engagement).

RESULTS

Descriptive statistics. **Table 1** shows the coaching version of the 25-item Workplace Outcome Suite.

Table 2 presents the means and standard deviation for the 25 items in the WOS across the board for the coaching version of the WOS.

Reliability analysis. **Table 3** presents the internal consistency for the four effect-indicator measures across the two studies.

Coefficient alpha for the paper-and-pencil study are as follows Presenteeism = .90; Work Engagement = .74; Life Satisfaction = .76, and Workplace Distress = .90. For the telephone interview study, Presenteeism = .92; Work Engagement = .63; Life Satisfaction = .78, and Workplace Distress = .88. Item-total correlations for all 20 effect indicator items are significant beyond the .01 level and substantial. Analysis of the alpha coefficients that would be created by removing a single item shows that there is very little improvement to be gained by removing any one item. There are some items in the workplace engagement that suggest some room for improvement, which will be better understood in the context of the confirmatory factor analysis in the next section.

Confirmatory factor analysis. We conducted a confirmatory factor analysis on the two data sets to test the overall structure of the items as they relate to their respective latent constructs and the others. This allows us to evaluate item level internal consistency and the discriminant validity of the items in terms of loading on their respective factor and not on other factors. In confirmatory factor analysis we test the hypothesis that the item defines its appropriate factor and

Table 1: The Coaching Version of the 25-item Workplace Outcome Suite

Workplace Outcome Suite (Coaching)		© Chestnut Global Partners				
General Instructions. Below is a series of statements that refer to aspects of your work and life experience that may be affected by the coaching health and lifestyle issue you want to address with the coaching program during the past 30 days. Please read each item carefully and answer as accurately as you can.						
Instruction for items 1-4. Please report for the period of the past thirty (30) days the total number of hours your coaching health and lifestyle issue:						Number of Hours
Absenteeism						
1.	caused you to miss work altogether.					
2.	made you late for work.					
3.	caused you to take off early.					
4.	pulled you away from your normal work location while still at work.					
5.	required you to be on the phone, e-mail or internet while at work.					
<i>Instruction for items 6-15. The following statements reflect what you may do or feel on the job or at home. Please indicate the degree to which you agree with each of the statements for the past thirty (30) days. Use the 1 - 5 response key to the right.</i>						
						Strongly Disagree
						Somewhat Disagree
						Neutral
						Somewhat Agree
						Strongly Agree
Presenteeism						
6.	I have a hard time doing my work because of my health and lifestyle issue.					1
7.	My health and lifestyle issue kept me from concentrating on my work.					2
8.	Because of my coaching health issue I was not able to enjoy my work.					3
9.	My health and lifestyle issue made me worry about completing my tasks.					4
10.	I could not do my job well because of my health and lifestyle issue.					5
Work Engagement						
11.	I feel stimulated by my work.					1
12.	I often think about work on my way to the work site.					2
13.	I feel passionate about my job.					3
14.	I am often eager to get to the work site to start the day.					4
15.	I often find myself thinking about my work at home.					5

is unrelated to the other factors. This is achieved by fixing the “off factor” loading to zero and testing the goodness-of-fit for the model of the four-correlated factors. **Figure 3** illustrates this factor structure with the four ellipses representing the respective latent factors for Presenteeism, Work Engagement, Life Satisfaction, and Workplace Distress.

The curved lines connecting the ellipses represent the correlations among the factors that are to be estimated. The arrows connecting each item to its respective ellipse represent the factor loading of the items. The small arrows at the bottom of the items represent the random measurement

error in each item that is removed from the latent factor.

We tested the model in Figure 3 using Maximum Likelihood estimation for the telephone interview sample involving 228 respondents. Modeling the observed covariance matrix to the hypothesized four-factor model yields a Bentler-Bonnet fit index of .86, a Bollen fit index of .88, and a comparative fit index of .88. Modeling the observed covariance matrix from the paper-and-pencil data set yields a Bentler-Bonnet fit index of .83, a Bollen fit index of .86, and a comparative fit index of .85. Both models yield respectable fits to their respective covariance matrices.

Table 2: Descriptive Statistics of the CGP Outcome Suite Candidate Items

Abbreviated Item Text		M	SD	N
1.	...caused you to miss work altogether.	.62	3.66	618
2.	...made you late for work.	.23	2.01	615
3.	...caused you to take off early.	.22	1.24	618
4.	...pulled you away from your normal work location while still at work.	.29	3.44	616
5.	...required you to be on the phone, e-mail, or internet while at work.	1.36	8.84	612
6.	I had a hard time doing my work because of my personal problems.	1.52	0.96	609
7.	My personal problems kept me from concentrating on my work.	1.54	0.96	610
8.	Because of my personal problems I was not able to enjoy my work.	1.58	1.00	610
9.	My personal problems made me worry about completing my tasks.	1.52	0.94	607
10.	I could not do my job well because of my personal problems.	1.40	0.85	605
11.	I feel stimulated by my work.	4.04	0.97	602
12.	I often think about work on my way to the work site.	4.03	1.07	599
13.	I feel passionate about my job.	4.15	0.98	603
14.	I am often eager to get to the work site to start the day.	3.70	1.09	601
15.	I often find myself thinking about my work at home.	3.76	1.14	603
16.	My life is nearly perfect.	3.44	1.05	610
17.	I am not very satisfied with my life as a whole.	1.96	1.07	609
18.	So far, my life seems to be going very well.	4.05	0.90	612
19.	There isn't anything about my life that I would change if I could.	2.64	1.23	606
20.	I am very disappointed about the way my life has turned out.	1.60	0.93	612
21.	I often feel anxious at work.	2.66	1.25	601
22.	Thinking about being at work makes me upset.	2.00	1.18	601
23.	I am unhappy most of time at work.	1.83	1.11	600
24.	I dread going into work.	1.85	1.15	596
25.	I can't wait to get away from work.	2.28	1.25	601
V1	I have a hard time getting out of bed.	2.04	1.16	290
V2	I often feel sad.	1.76	1.03	292
V3	I keep falling behind schedule at work.	1.97	1.20	291
V4	I am rarely late for work.	4.24	1.19	293
V5	I often get home late from work.	3.41	1.29	292

Table 4 contains the standardized loading for the items on their respective latent factor. The absenteeism items are not included in this analysis. Virtually all of the factor loadings for the telephone interview data were significant and substantial, suggesting that they are all appropriately related to their respective latent factor. The two exceptions are item 15, "I often find myself thinking about my work at home," and item 12, "I often think about work on my way to the work site," which are hypothesized to be core items of the work engagement scale. These loadings suggest that respondents affirming these items may not necessarily be highly engaged in their work and may

instead be distressed by their work. Still, all other items have high loading, and the two negatively worded life satisfaction items produced the expected negative loading on their respective factor. The standardized factor loading for the telephone interview data set produced significant and substantial fit for all items, including the two items in the Work Engagement scale. This suggests that the problem seen in the first sample may not replicate and it may be premature to consider modifying the scale for these items.

Table 5 presents the correlations among the latent factors as estimated in the confirmatory factor analysis. These factors are estimated by using the precise weights

Table 3: Reliability Analysis of the CGP Outcome Suite Coaching Items			
Abbreviated Item Text		Item-Total <i>r</i>	α if Deleted
1.	caused you to miss work altogether.	NA ¹	NA ¹
2.	made you late for work.	NA ¹	NA ¹
3.	caused you to take off early.	NA ¹	NA ¹
4.	pulled you away from your normal work location while still at work.	NA ¹	NA ¹
5.	required you to be on the phone, e-mail, or internet while at work.	NA ¹	NA ¹
6.	I had a hard time doing my work because of my personal problems.	0.84	0.92
7.	My personal problems kept me from concentrating on my work.	0.84	0.92
8.	Because of my personal problems I was not able to enjoy my work.	0.80	0.93
9.	My personal problems made me worry about completing my tasks.	0.85	0.92
10.	I could not do my job well because of my personal problems.	0.86	0.92
11.	I feel stimulated by my work.	0.64	0.77
12.	I often think about work on my way to the work site.	0.60	0.78
13.	I feel passionate about my job.	0.71	0.75
14.	I am often eager to get to the work site to start the day.	0.66	0.77
15.	I often find myself thinking about my work at home.	0.46	0.83
16.	My life is nearly perfect.	0.69	0.74
17.	I am not very satisfied with my life as a whole.	0.56	0.78
18.	So far, my life seems to be going very well.	0.68	0.75
19.	There isn't anything about my life that I would change if I could.	0.51	0.81
20.	I am very disappointed about the way my life has turned out.	0.58	0.78
21.	I often feel anxious at work.	0.58	0.90
22.	Thinking about being at work makes me upset.	0.79	0.85
23.	I am unhappy most of time at work.	0.76	0.86
24.	I dread going into work.	0.82	0.85
25.	I can't wait to get away from work.	0.73	0.87
Note: alpha coefficients for scales are: Presenteeism=.938; Work Engagement=.816; Life Satisfaction=.807; Workplace Distress=.889			

from the analysis and should not be confused with correlations with the unit-weighted scale scores. Correlations for the paper-and-pencil data set are presented below the diagonal and the telephone interview is presented above the diagonal. The results show a generally parallel pattern of positive and negative correlations with the difference in direction being attributed to the direction of scoring in the items that indicate which is associated with a high score. Workplace distress and work engagement produce a very high correlation in the telephone interview data set ($r = -.80$), but it is considerably lower in the paper-and-pencil data set ($r = -.57$) [Consult Table 1]. The two are not interchangeable, but it remains to be seen what level of differential prediction can be found with such highly correlated items. All other correlations are consistent with their respective theoretical constructs.

Taken together, these two sets of results support the structure of the 20 effect-indicator items in the WOS and measure four correlated latent constructs. When considered with the alpha coefficients, the results suggest that the scale can be expected to produce reliable measures of their respective constructs without excessive random measure error or overlap with other related constructs in the suite.

DISCUSSION

These studies support the basic psychometric properties of the Workplace Outcome Suite as a valid measurement tool of EAP outcomes in the workplace and for coaching support for disease management. The formal measurement models provide a transparent description of the manner in which the individual items impact the latent variable. The two potential problems

Table 4: Standardized Coefficients from the Four-factor Confirmatory Factor Analysis of the CGP Outcome Suite Coaching Items

Abbreviated Item Text		Standardized Loading
1.	caused you to miss work altogether.	NA ¹
2.	made you late for work.	NA ¹
3.	caused you to take off early.	NA ¹
4.	pulled you away from your normal work location while still at work.	NA ¹
5.	required you to be on the phone, e-mail, or internet while at work.	NA ¹
6.	I had a hard time doing my work because of my personal problems.	0.87
7.	My personal problems kept me from concentrating on my work.	0.87
8.	Because of my personal problems I was not able to enjoy my work.	0.83
9.	My personal problems made me worry about completing my tasks.	0.88
10.	I could not do my job well because of my personal problems.	0.90
11.	I feel stimulated by my work.	0.81
12.	I often think about work on my way to the work site.	0.48
13.	I feel passionate about my job.	0.90
14.	I am often eager to get to the work site to start the day.	0.90
15.	I often find myself thinking about my work at home.	0.35
16.	My life is nearly perfect.	0.77
17.	I am not very satisfied with my life as a whole.	0.65
18.	So far, my life seems to be going very well.	0.77
19.	There isn't anything about my life that I would change if I could.	0.58
20.	I am very disappointed about the way my life has turned out.	0.66
21.	I often feel anxious at work.	0.60
22.	Thinking about being at work makes me upset.	0.82
23.	I am unhappy most of time at work.	0.83
24.	I dread going into work.	0.90
25.	I can't wait to get away from work.	0.80

Table 5: Correlations Among the Four Latent Factors

	Presenteeism	Work Engagement	Life Satisfaction	Workplace Distress
Presenteeism	I			
Work Engagement	-0.27	I		
Life Satisfaction	-0.37	0.35	I	
Workplace Distress	0.32	-0.45	-0.33	I

in the Work Engagement scale found in the telephone interview studies were not found to be a problem in the paper-and-pencil study, suggesting that those smaller coefficients did not replicate, and as such should be considered spurious for the time-being. However, the items will be carefully examined in another study of the Internet modality that is currently underway. The items will be considered for replacement should the problem reappear.

Taken together, the results of the reliability analysis, the confirmatory factor analysis, and the correlational studies all support

the use of the WOS in its current form for coaching interventions. Alpha coefficient for the Presenteeism and the Workplace Distress scales are in the .90 range, showing them to be extremely reliable, especially for short scales. The mid-70s alpha for the Life Satisfaction scale in both studies shows to have an acceptable level of reliability. The one low alpha coefficient for the Work Engagement scale does not replicate in the paper-and-pencil studies, suggesting that the scale has acceptable reliability for a short scale in its early development.²⁰ The basic psychometric characteristics of the WOS show it capable

of measuring individual difference relevant to EAP interventions. However, further research on the construct validation, looking at actual behavior and records, will provide greater confidence on the value of the instrument.

One of the main purposes of creating the WOS was to provide EAP researchers with a set of short measures useful under a single validation umbrella. The evaluator does not need to search the literature for measures, nor comb through the proprietary catalogues, many of which do not include the syntax of the items before their purchase. The transparency of the measurement models coupled with the public nature of the validation material is designed to streamline the evaluation process and facilitate the empirical testing of EAP interventions. The standardized approach to the development and validation of the five constructs facilitate evaluation of the measures for specific use. Finally, Chestnut Global Partners, recently acquired by Morneau Shepell, makes the measures available for specific intervention testing at no cost to use the scientific method

for evaluation, comparison, and learning about the impact of EAP services.

AUTHORS

Richard D. Lennox PhD, is founder and President of Work Matters based in New York, U.S.A. Prior to Work Matters, Dr. Lennox was Chief Scientist and Vice President of Commercial Science at Chestnut Global Partners (CGP).

David Sharar, PhD, is CEO of Chestnut Health Systems and Director of Chestnut Global Partners, (a Morneau Shepell company) Division of Commercial Science, Bloomington, IL, U.S.A., is co-creator of the Workplace Outcome Suite® (WOS) and a leading provider of employee assistance services worldwide.

Francine Miller is Vice President of Wellness Divisions & Director of Clinical Services for Health Advocate headquartered in San Francisco, CA, U.S.A. Francine was Co-Founder of WellCall, which was acquired by Health Advocate.

AUTHORS' NOTES

Although the measures are copyrighted by Chestnut Global Partners, they are available for unqualified public use. We request that those interested contact us prior to using the measures so we can keep track of how the measures are being used by the community. Questions regarding the use of the measures should be directed to Drs. Richard Lennox (Rlennox@chesnut.org; (919) 942-0448) or David Sharar (dsharar@chesnut.org; (309) 820-3570). Correspondence should be sent to Dr. Richard Lennox, 2404 Western Park Lane, Hillsborough, North Carolina 27278.

REFERENCES

- Pelletier, K. R. (1999). A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1995-1998 update (IV). *American Journal of Health Promotion*, 13, 333-345.
- Pelletier, K. R. (2005). A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: Update VI 2000-2004. *Journal of Occupational and Environmental Medicine*, 47, 1051-1058.
- Lewis, A., & Khanna, V. (2014). *Surviving Workplace Wellness With Your Dignity, Finances and Major Organs Intact*. *The Health Care Blog*; 1 edition
- Petrak & Herpertz, (2009). Treatment of Depression in Diabetes: an update. *Current Opinion Psychiatry*, 22 (2) 211-217.
- Burton, Chen, Conti; Schultz & Edington, 2006; Johns, 2008; Burton, Chen, Conti, Schultz, Pransky, & Edington, 2005. The Association between Health Risk Change and Presenteeism Change. Wayne N. Burton, Chin-Yu Chen, Daniell. Conti, Alyssa B. Schultz, Dee W. Edington, *Journal of Occupational and Environmental Medicine*. 48 (3):252-263, 2006.
- Rost, K., Smith, J. L., & Dickinson, M. (2004). The effect of improving primary care depression management on employee absenteeism and productivity: A randomized trial. *Medical Care*, 42, 1202-1210.
- Wang, P. S., Simon, G. E., Avorn, J., Azocar, F., Ludman, E. J., McCulloch, J., Petukhova, M. Z., & Kessler, R. C., (2007). Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes. *Journal of the American Medical Association*, 298, 1401-1411.
- Wang, P. S., Simon, G. E., & Kessler, R. C. (2008). Making the business case for enhanced depression care: the National Institute of Mental Health-Harvard work outcomes research and cost-effectiveness study. *Journal of Occupational and Environmental Medicine*, 50, 468-475.
- Schneider, JI., Hashizume, J., Heak, S., Maetani, L., Rude Ozaki, R., & Leong Watanabe, D. (2011). Identifying challenges, goals and strategies for success for people with diabetes through life coaching. *Journal of Vocational Rehabilitation*, 34: 129-139.
- Merrill RM, Aldana SG, Vyhldal TP, Howe G, Anderson DR, Whitmer RW. *J Occup Environ Med*. 2011 Feb;53(2):127-31. doi:10.1097/JOM.0b03e318209c18b.
- Hawkes, AL., Atherton, J., Taylor, CB., Scuffham, P., Eady, K., Houston-Miller, N. & Oldenburg, B. (2009). Randomised controlled trial of a secondary prevention program for myocardial infarction patients ('ProActive Heart'): study protocol. *BMC, Cardiovascular Disorders*, 9:6.
- Linden, A., Butterworth, SW, & Prochaska, JO (2010). Motivational interviewing-based health coaching as a chronic care intervention. *Journal of Evaluation in Clinical Practice*, 16, 166-174.
- Aoun, S., Osseiran-Moisson, R., Shahid, S., Howat, P., & O'Connor, M. (2011). Telephone Lifestyle Coaching: Is it Feasible as a Behavioral Change Intervention for men? *J. Health Psychol*. Retrieved from <http://hpq.sagepub.com/content/early/2011/07/08/1359105311413480>
- Patja, K., Pilvikki, A., TOKula, K., Kyto, J., Kuronen, R., Ovaska, T., Harno, K., Nenonen, M., Wiklund, T., Kettunen, R., & Talja, M. (2012). Health coaching by telephony to support self-care in chronic diseases: clinical outcomes from The TERVA randomized controlled trial. *BMC Health Services Research*, 12:147.
- Margolius, D., Bodenheimer, T., Bennett, H., Wong, J., Ngo, V., Padilla, G., & Thom, DH. (2012). Health Coaching to Improve Hypertension Treatment in a Low-Income, Minority Population. *Annals of Family Medicine*, 10: 99-205.
- Bollen, KA (1989) *Structural Equations with Latent Variables*. New York: Wiley.
- Sanderson, Kristy, Tilse, Elizabeth, Nicholson, Jan, Oldenburg, Brian and Graves, Nick (2006). Which presenteeism measures are more sensitive to depression and anxiety?. *Journal of Affective Disorders* 101 (1-3) 65-74. <https://doi.org/10.1016/j.jad.2006.10.024>
- Lennox, R., Sharar, D., Schmitz, E., and Goehner, D. (2010). Development and validation of the Chestnut Workplace Outcome Suite. *Journal of Workplace Behavioral Health*, 25, 107-131.
- CGP Annual Report 2017 Link: <http://chestnutglobalpartners.org/Portals/cgp/Publications/Trends-Report-April2017.pdf>
- Nunnally, JC (1978). *Psychometric Theory*. (2nd Edition) New York: McGraw-Hill.
- Nunnally, JC, & Bernstein, IH., (1994) *Psychometric Theory*, (3rd Edition). New York: McGraw-Hill.
- Bollen, K. A., & Lennox, R. D. (1991). Conventional Wisdom on Measurement: A Structural Equations Perspective. *Psychological Bulletin*, 110:2, 305-314.
- Bollen, K. A. (1984). Multiple Indicators: Internal Consistency or no necessary relationship? *Quality and Quantity*, 18, 377-385.
- Cronbach, LJ. (1951). Coefficient alpha and the internal consistency of tests. *Psychometricka*, 16, 297-334.