IJHP

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

Richard D. Lennox, Ph.D; David Sharar, Ph.D, Eileen Schmitz, **MA**, **LPC**; David B. Goehner, LCSW

ABSTRACT

This article reports on the validation of the WOS-5 an abbreviated version of the Workplace Outcome Suite© (WOS). For reasons of efficiency and ease of delivery the field was looking for an abbreviated version of the original WOS tool. In this new abbreviated version four of the 5-items correspond to latent variable measures of presenteeism, work engagement, life-satisfaction and workplace distress. These items were selected based on highest factor loading from the original confirmatory factor analysis in the 25-item WOS development study. The fifth item is the single measure of absenteeism created using a formative measures model to count total hours missed by collapsing the total and partial days absent from work. Correlation evidence indicates the 5-item WOS to be a good measurement representation of the 25-item version. Test of sensitivity for three versions of the WOS (WOS-5, WOS-9, and WOS-25) showed the 5-item version to provide comparable sensitivity to change from various EAP service interventions from our pooled dataset. The newly constructed single absenteeism measure for the 5-item scale was shown to be the most sensitive of the various measures, even outperforming the 25-item version. The advantage, while small, was statistically reliable. Together, these results suggest that the 5-item WOS can be used to approximate the 25-item version without excessive loss of reliability, validity or sensitivity.

INTRODUCTION

The Workplace Outcome Suite (WOS) was originally designed to provide standardized outcome measures for evaluating the efficacy and effectiveness of Employee Assistance Programs on the following dimensions: absenteeism, presenteeism, work engagement, life satisfaction and workplace distress (Lennox, Sharar, Schmidt, Goehner, 2010). Such an outcomes measure must provide the "sharpest pencil" possible, reliably detecting small effect sizes given the variety in both degree of severity and breath of different problems presenting to current EAP providers. There is also an expected variance in the types of services offered and the quality of the intervention provided by counselors, especially when call center affiliates are involved.

The 25-item WOS has demonstrated a robust degree of statistical sensitivity to change in such EAP evaluations with as few as 50 observations. Although short when compared to other measurement tools, many EAPs consider the 25-item WOS as too long for regular use in routine outcome monitoring. As a response to these comments, we developed a 5-item version of the WOS that takes one question from each of the original 25-item WOS scales of Presenteeism, Work Engagement, Life Satisfaction and Workplace Distress. Traditional scaling techniques allowed us to select the best representation of each latent variable using the confirmatory factor analysis reported in the original 25-item WOS validation (Lennox, et al., 2010).

Since factor loading cannot be used to reduce the Absenteeism scale, the WOS-5 includes a new single item based more on the semantic meanings of language than on earlier empirical findings. For that reason, the original validation analysis could not be confidently extended to the WOS-5 so entirely new data was collected for validation purposes. Psychometric theory leads one to expect a loss of reliability and subsequent

predictability when the Presenteeism, Work Engagement, Life Satisfaction, and Workplace Distress scales are reduced from 5 items down to a representative one. There is also reason to expect a loss of validity in the original five Absenteeism scales collapsing into a single new item. This article examines the validity and sensitivity of the WOS-5 5-item version of the tool as compared with the validation findings of the original 25-item WOS scale.

ORIGINAL WOS MEASUREMENT MODELS

The original measurement models for the 25-item WOS (Lennox, et al. 2010) were derived from traditional psychometric theory and practice. Bollen & Lennox (1991) provide two illustrations of the models in terms of the items in the measures and their expected directional relation to the underlying latent variables. As seen in the left section of the figure below, the multiple effect model assumes that it is the underlying construct that causes the variance in the items *(Y)*. In psychology these items would be referred to as symptoms and manifestations of the latent construct. On the right side of the figure, the direction of cause goes the other way where formative items *(X)* combine to create or cause the latent variables. They are not manifestations or symptoms of the latent construct but the building block for a broader construct. For example, in our absenteeism scale we use five different ways

in which an employee may be away from his or her job – missing full days, being late, leaving early, and being actively engaged in activities at the workplace related to his or her problem rather than on their job. The measure combines these different responses in a summed assessment of total absenteeism. In this way, the total absenteeism score is more comprehensive than the particular response given.

Figure 1 depicts the creation of the 5-item WOS scale from the original 25-item version and sets up the soundness of the short version.

While the pictures look similar, it is the directionality of the arrows linking the items to the underlying construct that differ, as well as the scoring mechanics used. Connecting multiple effects to their underlying constructs involves adding together similar items that vary in their random measurement error. The statements used in the tool have just slightly different wording aimed at the same key point. This kind of model approach provides a basis for selecting the best single item to represent the latent construct from what are essentially highly similar measures.

The approach followed in abbreviating the 25 items came in two distinct steps. The first involved creating single measures for the four effect-indicator models: presenteeism, work engagement, life satisfaction and workplace distress. This involved the relatively straightforward process of selecting the item from each sub-scale that had the

Figure 1: Original Measurement Models for the Workplace Outcome Suite
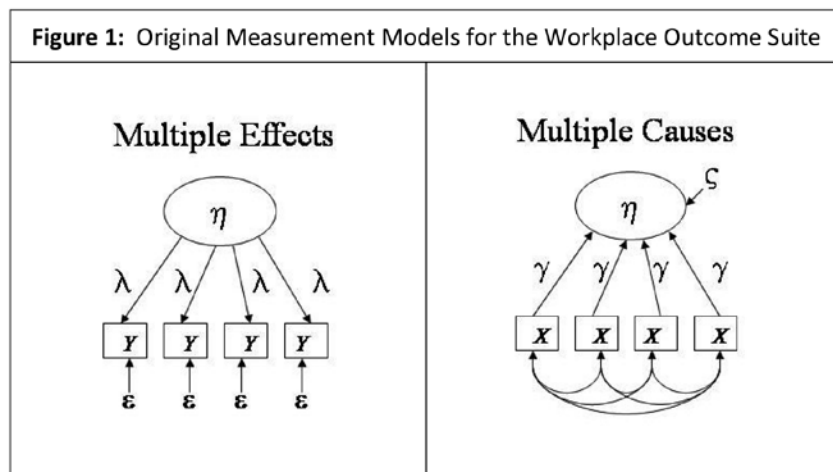
highest loading on their respective latent variables from the two samples in the original factor analysis (Lennox, et al., 2010). This is a common approach to selecting single indicators and can be shown empirically to be the best representation chosen from the item set. They can also be evaluated in terms of item-total correlations in a reliability analysis.

The second step followed in the abbreviation effort involved selecting one of the five absenteeism scale items. Selecting a single item from a cause indicator model can be challenging since the items are different and selecting the wrong one could lead to a serious bias in the final measure. Creating a single item absenteeism scale from the original 5-item scale required forming a completely different type of interim measurement tool to test.

A 9-item version of the WOS was created that retained the original five items of the absenteeism scale along with the single items for presenteeism, work engagement, life satisfaction and workplace distress. The 9-item WOS instrument detected change results similar to the full 25-item version using a pre-treatment/90-day post-treatment EAP interventions follow-up evaluation.

ABBREVIATING THE 5-ITEMS ABSENTEEISM SCALE

The original construction of absenteeism was based on a cause-indicator measurement model that could not be abbreviated by selecting the best single indicator due to the fact that the items were not designed to parallel one another. The original 5-item absenteeism scale is presented below:

Please report for the period of the past thirty (30) days the total number of hours your personal problem:

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.

The first step involved removing from consideration the last two items that referred to actually being at work, but away from one's regular job-duties while attending to one's personal problems. We reasoned that these two ideas were more closely aligned with presenteeism than absenteeism and whatever variance there was in these items was probably picked up by the presenteeism scale. One can certainly argue that the two items removed may be considered "bridge items" that span the absenteeism and presenteeism constructs and, as such, are part of both. The positive correlation between absenteeism and presenteeism reported in Lennox, et al. (2010) supports this view.

That left three items each of which measures time away from the job in full days, hours leaving early, or hours arriving late. This effectively defines "absent from work" as "not physically at the workplace". This definition collapses the first three items to simply asking about the total hours absent.

Regarding the two items removed, these may be considered "bridge items" that span the absenteeism and presenteeism constructs and as such are part of both. Yet, to define them as measures of presenteeism, then they would be the only presenteeism items measured in terms of "hours" instead of a Likert format.

CHALLENGES:

As stated previously, the purpose of this article is to investigate the validity of the 5-item version of the WOS as a time-saving and reasonable alternative to the full 25-item scale. Classical psychometrics predicts that the 5-item version will be less reliable than the full scale and more likely to be affected by greater random measurement error. The study began by examining the extent to which the 5-item construct captured the variance found in the 25-item version.

An additional goal of the investigation is to explore the degree to which any decrease in instrument reliability poses threats to the outcomes of certain types of research studies. For example, the effect of an increase in random measurement error may be offset by an inferential test of the WOS hypotheses by simply increasing the sample size. Prudence would then dictate that the 5-item version, with an expected decrease in reliability, might not be a good candidate for use in efficacy tests of small sample studies. Such research efforts would be better served by using the fine grain details from the original 25-item scale or be cautious when interpreting any non-significant finding due to the potential lack of statistical power.

METHOD

The validation investigation method followed is based on a quasi-experimental or correlational approach. Unlike a randomized clinical trial, no attempt is made to manipulate an independent variable to observe its effect on a dependent variable. The approach used here measures the constructs and assesses the covariance among scores from the same subject's responses. The 5-item WOS was evaluated for its correspondence to the 25-item and 9-items versions in terms of simple correlations and the similarity with which the various scale versions correlate with external criteria.

DATA SOURCES

Sample 1: Data from the paper-and-pencil study came from 200 clients served by Personal Assistance Services of St. Louis Missouri. Personal Assistance Services (PAS) began providing Employee Assistance Program (EAP) services in 1982 and has been a pioneer in transforming the traditional EAP model into a progressive risk management strategy to help employees negotiate and face many life events. PAS specializes in customizing innovative service designs that help keep employees healthy and productive on the job. Subjects were recruited by PAS as they arrived for routine in-clinic visits. To capture the broadest range of employee responses, no exclusion criteria were used. Subjects were simply asked to provide answers to the 25-item WOS tool on the questionnaire form. Responses were then transferred to an excel spreadsheet for compilation and analysis.

Sample 2: Data for the telephone interview modality was provided by 210 clients of Empathia, Inc. Empathia provides behavioral health solutions aimed at improving the well-being, safety and productivity of organizations and individuals. The company collaborates with private and public-sector entities ranging from Fortune 500

corporations to small businesses addressing a range of services. This range of services includes: employee assistance, disaster response and planning, workforce well-being, employee relations, leadership development, training, and benefits support.

Subjects were recruited as they participated in routine telephone contact with an Empathia clinician. The interviewer read the instructions and then asked the 25-item WOS questions in the order they appeared on the paper-and-pencil version, recording each response on the hard-copy questionnaire. Completed questionnaires were mailed to Chestnut Global Partners and entered into an excel spread sheet for analysis. During the pilot stage of the data collection at Empathia, respondents expressed frustration at being asked to recall their absenteeism for the previous 30 days. They claimed to simply be unable to remember the exact details. A modification was then made to switch to the Food Drug

Administration's (FDA) standard recall period of 7 days for this site while leaving the paper-and-pencil sample site at 30 days to provide a basis for comparison.

RESULTS AND DISCUSSION

Although the methods and analysis included several WOS versions to capture employee responses, the paper's results focus on what was learned from exploring the validity of the 5-item WOS. First, the paper addresses the performance results using the single absenteeism item in the 5-item scale constructed by collapsing "away-from-work" to asking if the employee was absent a full day, arrived late or left early. The remaining analysis then focuses on the cross-validation results derived from the original 25-item WOS validation study (Lennox, et al., 2010).

Descriptive Statistics: Table 1 presents the descriptive statistics for the 5-item version of the WOS. The top portion of

Table 1: Descriptive Statistics					
Empathia Sample					
Pre-test	N	Minimum	Maximum	Mean	Standard Deviation
Absenteeism	210	0	160	10.91	20.423
Presenteeism	210	1	5	3.48	1.349
Work Engagement	210	1	5	2.89	1.360
Life Satisfaction	210	1	5	2.48	1.223
Workplace Distress	210	1	5	2.51	1.439
Valid N	210				
PAS Sample					
Pre-test	N	Minimum	Maximum	Mean	Standard Deviation
Absenteeism	204	0	128	5.03	14.531
Presenteeism	200	1	5	2.79	1.340
Work Engagement	201	1	5	3.11	1.242
Life Satisfaction	201	1	5	2.91	1.156
Workplace Distress	199	1	5	2.18	1.298
Valid N	210				

Table 1 presents the Empathia sample.

Table 1 shows the 210 Empathia respondents with their minimum and maximum scores reflecting the single–item structure of the scale. The range on the singe absenteeism items is from 0 to 160 hours. The remaining four items have a range of 1 to 5 adhering to the Likert scale responses. The means and standard deviations for the four Likert scales show the means to approximate the center of the distributions and the standard deviations to reflect some amount of variability around the measures. The mean and standard deviation of the absenteeism question suggests the typical skewed distribution toward the lower end of hours missed. The maximum 160 hours is an extreme case. This skewed distribution is typical of some health behavior measures that capture data on workplace absenteeism and health care utilization.

Bivariate Correlations: The bottom portion of Table 1 show the same results for the PAS sample. The sample size of between 199 to 204 reflects a small amount of data missing from the responses. The means and standard deviation show a pattern similar to the Empathia sample on the four Likert items, with a slightly lower mean and standard deviation for the absenteeism measure. Taken together, the descriptive statistics show the expected central tendency and dispersions levels, including the skewed distribution on the absenteeism measure found in both samples.

The top of **Table 2** presents the bivariate correlations between and among the 5–item short scale, and the 25–item full scales of the WOS. It serves as the basic test of the capacity of the 5–item scales to actually stand–in for the 25–item version when a short scale is used. The correlations presented on the main diagonal of the matrix present the association between each of the sub–scales measures and the two different methods of measuring responses. As expected all of the correlations are statistically significant beyond the .000 level as they measure the same thing. Even so, they are expected to be affected by some degree of random measurement error. In Sample 1, all correlations exceed .80 except work engagement which is greater than .6. This strong correspondence shows the measures can be used interchangeably

without losing too much precision. Even the work–engagement question in the 5–item scale does not appear to lose significant precision. For the four effect–indicator measures, the correlation results are not surprising. Each of the four measures come from the parent measure so the correlation is actually the correlations of the specific items and the respective total of the items in the 25–items sub–scale.

The absenteeism item for the 5–item scale is entirely new and is not affected by common item variance. Therefore, the resulting correlation of .927 is a bit surprising given that some of the items were removed and two items collapsed. The high correlation suggests the single collapsed and truncated absenteeism items are an adequate proxy for the entire 5 items contained in the absenteeism section of the full WOS.

The bottom of Table 2 presents the bivariate correlations between the full WOS and the individual items of the 5–item WOS for the PAS sample. In general the pattern of correlations in the PAS sample parallel those in the Empathia sample. Although there are some differences in strength of correlation, the pattern of statistical significance in the two samples is virtually identical. Not only did the two samples produce parallel results of statistical significance, they also produced parallel results of non–significance. The results for the single absenteeism items are also parallel across the two samples. The one slight difference is in the correlations between the 5–item absenteeism scale from the 25–item WOS and the newly constructed single absenteeism item of the 5–item WOS scale. Specifically, the correlations for the Empathia sample was a very strong at .927 (p>.000) while slightly less at .626 (p<.000) in the PAS sample. While the absenteeism item showed some shrinkage from the Empathia to, the PAS sample, it still produced a significant and strong correlation. The difference does not pose a significant threat to the validity of the newly constructed absenteeism items.

The last line of the sample reports the efficiency measures for each of the five-item scales relative to their respective longer 25-item scale. Efficiency = (the number of short-form items, in this case the number of items in the full scale) divided by the diagonal

Table 2: Correlations between the 24- and 5-item version of WOS					
Empathia (N = 210)					
5-item WOS Scale					
Full WOS	Absenteeism	Presenteeism	Work Engagement	Life Satisfaction	Workplace Distress
Absenteeism	.927***	.287***	-.178**	-.174**	.182**
Presenteeism	.286***	.863***	-.389***	-.380***	.422***
Work Engagement	-0.85	-.260***	.683***	.214**	-.391***
Life Satisfaction	-.123	-.431***	.159*	.823***	-.212**
Workplace Distress	.169*	.405***	-.594***	-.262***	.895***
Efficiency	0.22	0.23	0.29	0.24	0.22
PAS (N = 200)					
5-item WOS Scale					
Full WOS	Absenteeism	Presenteeism	Work Engagement	Life Satisfaction	Workplace Distress
Absenteeism	.626***	.312***	-.059	-.124*	.037
Presenteeism	.199**	.897***	-.346***	-.348***	.371***
Work Engagement	.019	-.219**	.754***	.259***	-.339***
Life Satisfaction	-.042	-.413***	.310***	.724***	-.246***
Workplace Distress	.114*	.373***	-.519***	-.185**	.929***
Efficiency	0.32	0.22	0.27	0.28	0.22
Note: *** indicates $p < .000$; ** indicates $p < .01$; * indicates $p < .05$					

correlation between the two measures. The lower the result number, the better the efficiency since fewer items are associated with a higher correlation. The efficiency of the measures also supported the 5-item scale with all measures exceeding .20 suggesting that the items capture the meaning of the individual construct. As can be seen on the last rows of the individual matrices in Table 2, all efficiency values are greater than .20 indicating a good level of efficacy for all single-item scales. The one exception is in the absenteeism scale which exceeds .30 in

the PAS sample. This efficiency measure is not particularly high and it does not cross-validate to the other sample.

Table 3 presents the correlations between the individual scales of the 5-item WOS with several external criterion measures. The criterion measures represent behaviors and feelings that are expected to be differentially related to the five single-item measures. The Empathia sample shows significant correlations with all five WOS measures for trouble getting out of bed, feeling sad and falling behind at work. These three

Table 3: Pearson Correlations of the 5-item WOS with External Criterion Variables					
Empathia (N = 210)					
Criterion Variables	Absenteeism	Presenteeism	Work Engagement	Life Satisfaction	Workplace Distress
I. Getting out of bed	**.283*****	**.312*****	**-.235*****	**-.287*****	**.327*****
2. Feel sad	**.204*****	**.369*****	**-.241*****	**-.456*****	**.246*****
3. Falling behind	**.232*****	**.372*****	**-.376*****	**-.197****	**.373*****
4. Rarely late for work	**-.231*****	**-.199****	.099	.127	-.049
5. Working after hours	0.78	.055	**-.132****	.058	.107
PAS (N = 200)					
	Absenteeism	Presenteeism	Work Engagement	Life Satisfaction	Workplace Distress
I. Getting out of bed	.063	**.333*****	**-.250*****	-.092	**.438*****
2. Feel sad	**.183****	**.506*****	**-.196****	**-.354*****	**.374*****
3. Falling behind	.026	**.417*****	**-.311*****	-.126	**.506*****
4. Rarely late for work	-.121	-.137	-.001	.012	-.109
5. Working after hours	.099	**.157****	.082	-.048	**.152****
Note: *** indicates p <.000; ** indicates p <.01; * indicates p <.05					

behaviors are measures of constructs that may well translate into problems at work. All are positively associated with absenteeism, presenteeism, and workplace distress and are negatively associated with work engagement and life satisfaction.

The rarely-late-for-work and working-after-hours responses show significance for the absenteeism and presenteeism scale but fail to correlate with work engagement, life satisfaction or workplace distress. The working-after-hours questions only correlated significantly with work engagement. The fact that these two items do not correlate with all measures offers some support for the discriminant validity of the scale.

An attenuated pattern of significant correlations with these three criterion measures is found in the PAS sample. All correlations are in the expected direction,

but some fail to reach statistical significance. This pattern offers some limited support for the construct validity for all of the single item measures in the 5-item WOS. It is important to point out that all items are self-reported and subject to some level of measurement bias. The research reported thus far focuses on the correlations of the 5-item WOS scales to other forms of the WOS and some self-reported criteria. In the next section of the report we shift to considering the relative sensitivity of the measures to change across types of EAP interventions.

Table 4 presents the results of pooled comparison of the WOS version across EAP interventions.

Tables 4a − 4d presents the results of a sensitivity analysis of the three versions of the WOS. The analysis is based on a pooled collection of the results from several pre-treatment and post-treatment comparisons of

EAP interventions. The quasi-experimental analysis does not use randomized assignment or control groups. The analysis simply points to each version of the WOS to detect change regardless of it actual cause.

Table 4a shows the sensitivity analysis results for the 25-item WOS compiled from several studies. However, not all studies reported chose to use all the scales which helps explain the different sample sizes noted in the analysis. The samples approximate 1,000 study responses which should provide substantial statistical power for detecting the comparison between pre-treatment post-treatment change. In fact, for the 25-item version, all WOS scales produced statistical significance change from the baseline except for work engagement. It did produce change however in the proper direction and the difference reached the traditional level of statistical significance. All other scales showed statistical change at the .000 level.

Table 4b presents the results for the 9-item WOS version. Again, the 9-items version contains all five of the original absenteeism items and a single item from each of the remaining four Likert scales selected on their factor loading in the original validation study (Lennox, et al., 2010). The 9-item scale results were based on an analysis of approximately 3,300 responses. This large sample produced statistically significant change for the five-item absenteeism scales and for the single item measure for the remaining scales. Work engagement was statistically significant for

all measures at the .000 level.

Table 4c presents the results for the 5-item WOS scale based on a study sample of approximately 4,400 responses. The rationale for the modification of the absenteeism scale into a single item was presented earlier. They used this new collapsed measure and the same four single items of presenteeism, work engagement, life satisfaction and workplace distress as used in the 9-item version. The results mirrored the 9-item WOS scale, producing statistically significance change scores for all scales at the .000 level.

Finally, **Table 4d** shows the various definitions combined into a single pool with analysis extracting items from their different scale scores where possible. Results show the measures to be able to detect statistically significant change at .000 levels for all WOS scales.

It is also useful to consider the effect size of the intervention in addition to looking at the statistical significance of the various comparisons by examining the size of test statistic itself. As described earlier, the single absenteeism item used in the 5-item WOS performed well.

In summary Tables 4a – 4d shows the expected difference at pre-score, post-score, and the raw difference score. The percentage difference flips around to show a bigger percent improvement for the single item a difference that remains after using the z score. The data suggest that the 1-item version is more sensitive to change than

Table 4a. Results for Workplace Outcome Suite 25-Item Pre and Post-Test Scores							
Wos Scale	Pre Score	Post Score	N	Raw Difference Score	t^a	p-value	Difference Percentage
Absenteeism*[b]	**12.89**	**6.81**	**950**	**-6.08**	**-8.83**	**0.000**	**-47%**
Presenteeism*	**14.13**	**11.20**	**1,292**	**-2.93**	**-16.99**	**0.000**	**-21%**
Work Engagement**[b]	17.94	17.78	932	-0.16	-1.19	0.235	-1%
Life Satisfaction**	**12.27**	**13.64**	**1,288**	**1.37**	**-12.92**	**0.000**	**11%**
Workplace Distress*	**13.16**	**11.96**	**1,287**	**-1.20**	**-8.45**	**0.000**	**-9%**

Notes: *Lower scores are a better outcome; **Higher scores are a better outcome. Significant results are bold.
Presenteeism, work engagement, life satisfaction and workplace distress are the single item scores across all 3 versions of the WOS. [a]Wilcoxon signed rank test used to test change in absenteeism. The Z statistic is reported.

Table 4b. Results for Workplace Outcome Suite 9-Item Pre and Post Test Scores

Wos Scale	Pre Score	Post Score	N	Raw Difference Score	t^a	p-value	Difference Percentage
Absenteeism*	11.70	6.30	3,316	-5.4	-22.34	0.000	-46%
Presenteeism*	3.51	2.53	3,312	-0.98	-41.21	0.000	-28%
Work Engagement**	3.19	3.50	3,312	0.31	16.34	0.000	10%
Life Satisfaction**	2.73	3.56	3,312	0.83	39.76	0.000	30%
Workplace Distress*	2.40	2.03	3,309	-0.37	-18.38	0.000	-15%

Table 4c. Results for Workplace Outcome Suite 5-Item Pre and Post Test Scores

Wos Scale	Pre Score	Post Score	N	Raw Difference Score	t^a	p-value	Difference Percentage
Absenteeism*	5.30	2.38	4,333	-2.92	-25.43	0.000	-55%
Presenteeism*	3.25	2.31	4,453	-0.94	-40.37	0.000	-29%
Work Engagement**	3.25	3.45	4,448	0.20	9.43	0.000	6%
Life Satisfaction**	3.05	3.75	4,449	0.70	33.92	0.000	23%
Workplace Distress*	2.05	1.75	4,441	-0.30	-16.26	0.000	-15%

Table 4d. Results for Workplace Outcome Suite Pre and Post Test Scores Pooled Across Versions

Wos Scale	Pre Score	Post Score	N	Raw Difference Score	t^a	p-value	Difference Percentage
Absenteeism 9 and 25-item WOS versions*	11.97	6.42	4,266	-5.55	-23.87	0.000	-46%
Absenteeism 5-item WOS versions*	5.30	2.38	4,333	-2.92	-25.43	0.000	-55%
Presenteeism*	3.31	2.40	9,056	-0.91	-58.78	0.000	-27%
Work Engagement**	3.22	3.44	8,689	0.22	16.14	0.000	7%
Life Satisfaction**	2.94	3.64	9,041	0.70	51.70	0.000	24%

Notes: *Lower scores are a better outcome; **Higher scores are a better outcome. Significant results are bold. Presenteeism, work engagement, life satisfaction and workplace distress are the single item scores across all 3 versions of the WOS. [a]Wilcoxon signed rank test used to test change in absenteeism. The Z statistic is reported.

the 5-items, and thus probably more valid. However, it would be a mistake to rely on the either the raw pre or post scores alone as a basis for effectiveness for the single item measure because of the differences in scale or the difference in the hours missed. These differences can be corrected by using a change score standardized as a t statistic.

CONCLUSIONS

This paper reports on the validation results for an abbreviated and modified version of the Workplace Outcomes Suite. Measures of presenteeism, work engagement, life satisfaction, and workplace distress were selected from the initial confirmatory factor analysis that best represented their respective factor with the effect-indicator measurement model. Absenteeism was redesigned as a collapsed version of the original three items and now captures data on full and partial days taken off from work due to personal problems. The WOS-5 also removed two other scale items regarded as more in line with presenteeism than absenteeism.

Results from the correlational analysis show that the single absenteeism scale correlates highly with the other five absenteeism scales. The pattern of correlations is also consistent with the 25-item and 9-item versions of the scale. Tests of sensitivity of the various measures for evaluating EAP intervention effects show the 5-item WOS to parallel results from the longer tools. The new collapsed absenteeism items works slightly better than other absenteeism measures, especially when considering the effect size and statistical significance.

Finally for a straightforward test of the effectiveness of an EAP program, either the WOS-5 or the full 25-item WOS work well with the latter being more sensitive for use with smaller sample groups. Although the shortened scales would be suggested by psychometric theory to be less reliable, comparisons of the pooled data suggest they work very well. Taken together the correlational analysis and the pretreatment-post-treatment comparison of the pooled data show the 5-item WOS to be a reliable and valid measure when testing for the outcome constructs.

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.

Eileen Schmitz, MA, LPC, is Vice President at Personal Assistance Services (PAS), an Employee Assistance Program serving organizations globally located in St Louis, Missouri, U.S.A.

David B. Goehner, LCSW, is Director of Privacy & Compliance for Empathia, Inc. headquartered in Waukesha, WI, U.S.A.

BIBLIOGRAPHY

Bollen, K, A., & Lennox, R, D. (1991). Conventional Wisdom on Measurement: A Structural Equations Perspective. *Psychological Bulletin,* 110:2, 305-314.

Burgess, K., Lennox, RD., Sharar, D, & Shtoulman, A (2015). A Substance Abuse Intervention Program at a

Large Russian Manufacturing Worksite (2015)., *Journal of Workplace Behavioral Health,* 30:138–153.

Sharar, D., Pompe, J. & Lennox, R. (2012). Evaluating the workplace effects of EAP counseling: *Journal of Health and Productivity.* 6:2 5-14.

Lennox, R., Sharar, D., and Burke, J. (2010). Conducting an EAP evaluation using the Workplace Outcome Suite. *Journal of Employee Assistance,* Volume 40, No. 4, 24-27.

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.

Lennox, R., Sharar, D., Burke, J. (2009). Measuring workplace outcomes in EAPs. *Journal of Employee Assistance,* Volume 39,

Number 4, 18-19.

Sharar, D., Lennox, R. (2009). A new measure of EAP success: A workplace outcomes-based survey model provides validation. *Society for Human Resource Management Online,* 9/24/09.

APPENDIX

CGP WORKPLACE OUTCOME SUITE (WOS)					
GENERAL INSTRUCTIONS Below is a series of statements that refer to aspects of your work and life experience that may be affected by the personal problems you want to address at the EAP during the past 30 days. Please read each item carefully and answer as accurately as you can.					

	INSTRUCTIONS FOR ITEMS 1 - 5 Please report for the period of the last 30 days the total number of hours your personal problems.	NUMBER OF HOURS				
ABSENTEEISM	1. Caused you to miss work altogether.	1	2	3	4	5
	2. Made you late for work.	1	2	3	4	5
	3. Caused you to take off early.	1	2	3	4	5
	4. Pulled you away from your normal work location.	1	2	3	4	5
	5. Required you to be on the phone, e-mail or internet while at work.	1	2	4	4	5

	INSTRUCTIONS FOR ITEMS 6 - 25 The following statements reflect what you may do or feel on he job or at home. Please indicate the degree to which you agree with each of the statements for the past 30 days. Use the 1 - 5 response key to the right.	STRONGLY DISAGREE	SOMEWHAT DISAGREE	NEUTRAL	SOMEWHAT AGREE	STRONGLY AGREE
PRESENTEEISM	6. I had a hard time doing my work because of my personal problems.	1	2	3	4	5
	7. My personal problems kept me from concentrating on my work.	1	2	3	4	5
	8. Because of my personal problems I was not able to enjoy my work.	1	2	3	4	5
	9. My personal problems made me worry about completing my tasks.	1	2	3	4	5
	10. I could not do my job well because of my personal problems.	1	2	3	4	5
WORK ENGAGEMENT	11. I feel stimulated by my work.	1	2	3	4	5
	12. I often think about work on my way to the work site.	1	2	3	4	5
	13. I feel passionate about my job.	1	2	3	4	5
	14. I am often eager to get to the work site to start the day.	1	2	3	4	5
	15. I often find myself thinking about work at home.	1	2	3	4	5
LIFE SATISFACTION	16. My life is nearly perfect.	1	2	3	4	5
	17. I am not very satisfied with my life as a whole.	1	2	3	4	5
	18. So far, my life is going very well.	1	2	3	4	5
	19. There isn't anything I would change about my life if I could.	1	2	3	4	5
	20. I am very disappointed about the way my life has turned out.	1	2	3	4	5
WORKPLACE DISTRESS	21. I often feel anxious at work.	1	2	3	4	5
	22. Thinking about being at work makes me upset.	1	2	3	4	5
	23. I am unhappy most of the time at work.	1	2	3	4	5
	24. I dread going into work.	1	2	3	4	5
	25. I can't wait to get away from work.	1	2	3	4	5

APPENDIX

		WORKPLACE OUTCOME SUITE - 5 ITEM VERSION					
		GENERAL INSTRUCTIONS Below is a series of statements that refer to aspects of your work and life experience that may be affected by the personal problems you want to address at the EAP during the past 30 days. Please read each item carefully and answer as accurately as you can.					
			NUMBER OF HOURS				
AB	1.	For the period of the last 30 days, please total the number of hours your personal concern caused you to miss work, Include complete eight-hour days and partial days when you came in late or left early.					
		INSTRUCTIONS FOR ITEMS 2 - 5 The following statements reflect what you may do or feel on he job or at home. Please indicate the degree to which you agree with each of the statements for the past 30 days. Use the 1 - 5 response key to the right.	STRONGLY DISAGREE	SOMEWHAT DISAGREE	NEUTRAL	SOMEWHAT AGREE	STRONGLY AGREE
PR	2.	My personal problems kept me from concentrating on my work.	1	2	3	4	5
WE	3.	I am often eager to get to the work site to start the day.	1	2	3	4	5
LS	4.	So far, my life seems to be going very well.	1	2	3	4	5
WD	5.	I dread going to work.	1	2	3	4	5
		Copyright © Chestnut Global Partners 2013.07.02					