



Published in final edited form as:

Am J Health Promot. 2008 ; 22(5): 359–367.

Assessing Management Support for Worksite Health Promotion: Psychometric Analysis of the Leading by Example (LBE) Instrument

Lindsay J. Della Ph.D.¹, David M. DeJoy Ph.D.^{1,*}, Ron Z. Goetzel Ph.D.², Ronald J. Ozminkowski Ph.D.², and Mark G. Wilson H.S.D.¹

¹Department of Health Promotion and Behavior College of Public Health University of Georgia

²Institute for Health and Productivity Studies Cornell University Thomson Medstat

Abstract

Objective—This paper describes the development of the Leading by Example (LBE) instrument.

Methods—Exploratory factor analysis was used to obtain an initial factor structure. Factor validity was evaluated using confirmatory factor analysis methods. Cronbach’s alpha and item-total correlations provided information on the reliability of the factor subscales.

Results—Four subscales were identified: business alignment with health promotion objectives; awareness of the health-productivity link; worksite support for health promotion; leadership support for health promotion. Factor by group comparisons revealed that the initial factor structure is effective in detecting differences in organizational support for health promotion across different employee groups

Conclusions—Management support for health promotion can be assessed using the LBE, a brief, self-report questionnaire. Researchers can use the LBE to diagnose, track, and evaluate worksite health promotion programs.

Introduction

Traditionally, workplace health promotion programs have focused almost exclusively on changing individual employee health behavior. Relatively little attention has been given to environmental factors impacting healthy lifestyle, neither in terms of the physical environment of the workplace nor the broader operational-organizational context including leadership support for such initiatives. The narrow focus on individual behavior is somewhat surprising given the acknowledged importance of environmental factors in most conceptualizations of health promotion. Green and Kreuter, for example, define health promotion as: “...the combination of educational and environmental supports for actions and conditions of living conducive to health.” (p. 4).¹ O’Donnell also acknowledges the interaction of behavioral and environmental factors, and further argues that the environment is likely to be the most important influence in producing sustained changes in health practices.² Even *Healthy People 2010* references “comprehensive programs” when setting objectives for worksite health promotion.³ Within this framework, supportive social and physical environments are considered essential aspects of comprehensive programs.

*Corresponding Author and Reprint Requests Address: 315 Ramsey Center, University of Georgia, Athens, GA 30602, Fax: 706-542-4956, E-mail: dmdejoy@uga.edu.

The focus on individual behavior change research in the workplace, however, is beginning to shift toward more inclusive approaches. One of the factors responsible for this shift is a growing interest in socio-ecological models of health promotion and the use of multi-level interventions that involve combinations of individually- and environmentally-focused programs. Another force behind this shift can be found in the burgeoning practice of translating community-based capacity building concepts to a workplace environment. In the literature, Stokols advocates for expanding the health promotive capacity of environments, and DeJoy and Wilson discuss the merits of organizational health promotion.^{4,5} Additionally, the National Institute for Occupational Safety and Health (NIOSH) recently commissioned two position papers on the integration of the occupational safety and health and worksite health promotion as part of its *Steps to a Healthier Workforce* initiative.^{6,7} These papers further highlight the environment-behavior interface in terms of employee health and well-being and the importance of maximizing human capital through health promotion and health protection initiatives.

Finally, the obesity epidemic has been a catalyst in shifting worksite health promotion research toward more inclusive approaches that consider environmental and ecological interventions. The increasing proportion of Americans who are overweight or obese has raised questions about the relative contribution of environmental factors to this worsening public health problem. Concerns about obesity have spilled over into the workplace, and the National Heart, Lung, and Blood Institute (NHLBI) recently funded seven workplace studies to examine the contribution of environmental factors to overweight and obesity-related health and financial outcomes. Our participation as one of the study sites in this initiative formed the basis for the research reported in this article.

The Role of Business Leaders in Health Promotion Efforts

Management support is typically viewed as critical to the success of workplace health promotion programming.^{8,9,10,11} As programming efforts seek to modify work environments and socio-organizational systems, the issue of management support becomes even more critical. While leadership support is needed to institute and sustain individually focused health promotion programs, higher levels of support are required for programs that aim to make changes to the physical environment and introduce changes to operational and organizational policies and procedures related to worker health. Although management support has been widely discussed in the workplace health promotion literature, there have been surprisingly few attempts to describe or measure it.

For our research, we sought to develop a tool that could be used to assess the level of organizational support and management engagement in health promotion. Our focus on management support led us to the broad realm of organizational climate research. Our review of the literature revealed a number of studies and instruments assessing the climate for workplace safety, but a dearth of research focused on assessing the climate for health promotion.^{12,13,14}

Ribisl and Reischl's health climate questionnaire represents one of the few attempts to assess health-related climate factors within work organizations.¹⁵ Their instrument features 12 subscales. One of the subscales, "employer health orientation", provides a global assessment of management support for health promotion. More recently, Barnett and colleagues developed an organizational leadership scale as part of the Alberta Health Project in Canada.¹⁶ This scale follows an organizational learning perspective, but is oriented more towards communities than workplaces. For our research, we sought a scale or set of subscales that measure the various facets of management support for health, and that have some diagnostic value for use in intervention studies and program evaluations.

Our quest to identify a tool that could be used to measure specific elements of management support for health promotion led us to the “Leading by Example” (LBE) questionnaire, developed by Partnership for Prevention.¹⁷ The questionnaire had been used as a descriptive/educational tool as part of Partnership’s broader Leading by Example initiative. With permission from Partnership, we adopted their tool as the foundation for the current instrument. We modified the survey items in an effort to develop a more robust tool for diagnosing management issues and challenges, and tracking management support over time.

In this paper, we describe a psychometric analysis of the modified LBE. Our primary interest was to confirm that questionnaire items successfully operationalized different facets of management commitment/engagement with respect to health promotion (instrument validity), as well as to determine whether the items yielded consistent measurements of each factor (instrument reliability).

A variety of analytical tools were used to assess reliability and validity. Reliability, or measurement consistency, was estimated using single-test procedures. Specifically, Cronbach’s alpha and item-total correlations served as evidence for internal consistency within LBE subscales.¹⁸ Three types of validity assessments were employed to gauge the extent to which operationalized survey items measured the phenomena they were designed to measure: content, construct, and discriminant validity.

Methods

Instrument Development: Assessing Content Validity

The starting point for instrument development was the original LBE assessment developed by Partnership for Prevention in Washington, D.C. and The WorkCare Group, Inc. in Charlottesville, VA. This first version of the LBE was developed primarily as an awareness/educational tool to be used by human resources professionals, health promotion managers, and corporate managers and executives. It includes 19 items grouped into six labeled categories: mission, data management, benefit design, programming, corporate environment, and evaluation. In the original version, the response format requests a “Yes” – “No” – “Don’t Know” answer.

The first version of the LBE was reviewed in 2004 by the team of researchers involved in the NHLBI-funded research project. This team included Ph.D. level specialists in health promotion, public health, applied psychology, statistics, economics, and communications. The reviewers also included health promotion, medical, and human resources professionals from the participating corporation. The original developers of the LBE were also consulted. We were principally interested in selected portions of the assessment that focused on management support, commitment, and engagement.

The original LBE provided a core of seven items directly related to management support, commitment, and engagement. Subsequently, new items were generated, critiqued, and revised by the team through a series of team meetings and conference calls. The new items that emerged from this iterative process addressed topics such as health promotion goal setting and alignment, leadership training, communication, culture building, and financial and other supports for health promotion. All items, both old and new, were edited and simplified for use with employees representing a range of educational levels.

Additionally, modifications were made to some of the items in an effort to match the terminology used by the partnering organization. The response format for all items was changed to a five-point Likert scale, with a neutral midpoint (1 = “strongly disagree” to 5 = “strongly agree”). A pilot version of the questionnaire was then tested in spring of 2005, using one of

the control sites participating in the larger obesity-related intervention study. The pilot test was used to identify any ambiguous or confusing items. The top portion of Table 1 contains the 15-item LBE that emerged from this review and development process during 2004–2005.

Respondents

As part of formative research activities for the larger intervention study, the draft LBE was administered to groups of employees at 11 of the 12 sites participating in the study in 2005 (the 12th site had been used for the pilot test). Questionnaires were distributed to three groups of employees at nine of the 11 sites: site leadership (management team members), health services staff, and members of the cross-discipline team that served as employee advisory committees. Because cross-discipline teams were not appointed at the control sites (two of the 11 sites), questionnaires were only administered to leadership and health services at the control sites. Participants were sent an electronic copy of the LBE questionnaire. All potential respondents were requested to return their responses electronically or via fax. Completion of the questionnaire was voluntary and respondents were not compensated.

The average response rate across sites was 56.7%, resulting in an initial sample size of 136. Descriptive statistics identified one outlier across nearly half of the fifteen items assessed via the LBE. Based on the pervasiveness of this outlier, this respondent was excluded, reducing the total sample size to 135 for all subsequent statistical procedures. This sample size reflects a nine-to-one ratio of respondents to variables, a ratio generally deemed acceptable for multivariate statistical analyses.^{19,20} Of the 135 respondents, half (51.1%) classified themselves as site leadership; one quarter (24.4%) as members of health services; and one quarter (24.4%) as cross-discipline team members.

Initial tests of sampling adequacy confirmed that factor analysis procedures could be performed on these data. Bartlett's test of sphericity was statistically significant ($p = .000$), indicating that the correlation matrix was not an identity matrix and that at least some of the items were correlated (a prerequisite for factor analysis to produce interpretable solutions).¹⁹ Additionally, measures of sampling adequacy confirmed that sufficient correlations existed among variables at this sample size for conducting factor analyses. The Kaiser-Meyer Olkin summary test yielded an index of .873, and all of the anti-image correlations (i.e., negative partial correlations) were low.¹⁹

A second sample was collected from all 12 worksites one year later, in 2006, yielding 178 responses. These data were reviewed for outliers and subsequently used to validate the factor structure generated from the 2005 sample. Validating factor structure results in a second data set can help ensure that the final model is not overfitted to the development data set.¹⁹

In general, the 2005 and 2006 samples were similar. In 2006, however, all three groups of employees at all 12 sites were asked to respond to the survey. As a result, the 2006 sample is broader: It contains responses from cross-disciplinary team members at control sites.

Analysis: Assessing Reliability & Validity

The initial sample ($N = 135$ responses) was subjected to two separate factor analyses. In general, factor analysis produces two primary outputs. First, it generates a set of factors: Factors are groups of variables that are highly correlated with one another, but relatively uncorrelated with other variables in the dataset. Second, it estimates factor loadings: Factor loadings are approximations of the strength of association each variable holds with each factor.²⁰ An item's factor loading is an indication of its contribution to the explanation of the extracted abstract factor.

In this study, LBE responses were first entered into exploratory factor analyses (EFA), and subsequently into confirmatory factor analyses (CFA). Confirmatory factor analyses were employed to assess the extent to which the factors extracted in the EFAs exhibited discriminant validity, a requisite criterion for confirming a specific factor structure.

Because LBE items had been selected based on substantive merit but had been drawn from different sources, it was determined that the data should first be entered into an EFA. We reasoned that the items included in the LBE might actually be separate subscales tapping different factors related to management's role in health promotion. Because EFA makes no *a priori* assumptions about the structure of the subscales, it was identified as the most appropriate statistical procedure to test this assertion. Based on the factor structure that emerged from the EFA (evidence of construct validity), we tested the hypothesis that the multi-factor EFA structure fit the data better than a general one-factor structure (i.e., the items are assumed to load on only one general factor). This hypothesis was tested via confirmatory factor analysis, and served as a tentative indication of discriminant validity among the subscales.^{21,22} Finally, a strict CFA was used to validate the factor structure findings in the second sample.

The EFAs employed the principle components method (PCA) of factor estimation, and were run using an oblique (Oblimin) factor rotation in SPSS, version 13.0 (Chicago, IL; SPSS, Inc., 2004). The principle components method of factor estimation (as opposed to common factor methods) was employed because it is less likely to suffer from factor indeterminacy. And, in general, researchers have found that PCA returns factor scores with negligible differences from those generated through common factor techniques.²³ Furthermore, an oblique factor rotation was deemed as the optimal rotation for these data because it allows extracted factors to covary. We reasoned that if two factors exhibited an orthogonal relationship, the orthogonal nature of the relationship could still be identified via small, non-significant correlations in the component correlation matrix.

During the EFA process, decisions about the number of factors to retain were based on the convergence of several different factor retention criteria²⁰: eigenvalues greater than 1.0²⁴, Cattell's scree plot²⁵, parallel analysis²⁶, and theoretical interpretability of the final factor structure. The eigenvalues greater than 1.0 rule, also referred to as the Kaiser rule, suggests that researchers retain all factors that extract at least as much variance as one of the original variables. Cattell's scree test plots factor eigenvalues in descending order. Researchers typically look for an "elbow" in this plot (i.e., the last significant drop in variance explained before the plot evens out).²⁷ Unlike the first two approaches, parallel analysis proposes a less subjective approach to factor retention decisions. Parallel analysis compares the factor eigenvalues to eigenvalues obtained from random data. Using parallel analysis as one's factor retention criteria dictates preserving only those factors that produce eigenvalues greater than their associated, randomly-generated factor eigenvalues.²⁷

After obtaining a final EFA factor structure, Cronbach's alpha¹⁸ and item-total correlations were assessed to gauge internal consistency (i.e., reliability) within each subscale.²⁸ Internal consistency estimates provide an assessment of intercorrelations among a set or subset of measured scale items, while item-total correlations measure the degree of association between an individual item in a scale and the overall scale score. Researchers can generally claim strong internal consistency with high Cronbach alphas and significant item-total correlations. Determining that participants' responses are consistent across scale items is generally viewed as preliminary evidence that the scale items represent one underlying content domain (construct validity), but in practice should not be cited as evidence of factor unidimensionality.²⁹

Confirmatory factor analyses can be used as a more rigorous test of factor unidimensionality (i.e., construct validity) and factor distinction (i.e., discriminant validity). In this study, all

CFAs were run using LISREL 8.7 (Lincolnwood, IL; Scientific Software International, 2004). The chi-square statistic was used to assess model fit.³⁰ The chi-square statistics is an absolute measure of how well the hypothesized model fits the variation observed in the data (i.e., how well the fitted covariance matrix matches the sample covariance matrix). To conclude that the hypothesized model fit the data well, the fitted covariance matrix should be essentially equivalent to the sample covariance matrix. In this context, a non-significant chi-square value indicates strong model fit.

The chi-square statistic, however, can be influenced by sample size and should not be used as a stand-alone measure of model fitness. Other indices such as the Tucker Lewis Index (TLI), the Standardized Root Mean Square Residual (SRMR), the Comparative Fit Index (CFI), and the Root Mean Square Error of Approximation (RMSEA) index, which are based on differing theoretical foundations, were also used to assess model fit. Hu and Bentler recommend evaluating models against the following cutoff values: values below a RMSEA of .06 and a SRMR of .08, and values above a TLI of .95 and a CFI of .95.³⁰

Additionally, models were compared with one another based on indices of statistical fit (ΔX^2) as well as practical fit (ΔCFI). If one or more restricted CFA models are nested within a freely estimated model, the difference between the two chi-square values (ΔX^2) can be assessed for statistical significance (using the chi-square distribution).³¹ Similarly, the difference between the comparative fit indices (CFI), a measure of incremental model fit, can be used to compare nested models. Cheung and Rensvold suggest that a .01 CFI increase in a freer model indicates a significant improvement in fit over the more restricted more parsimonious model.³²

Results

Principle Components Analysis

The initial EFA yielded three, possibly four main factors. Cattell's scree plot suggested three to four primary factors, the eigenvalues greater than 1.0 criteria suggested three factors, and theoretical interpretability allowed for four factors. Parallel analysis, however, suggested only one strong factor existed with eigenvalues greater than those that could be expected by random variables. The total variance explained by the first four possible factors is detailed in the top portion of Table 2. Overall, the four factors extracted in this initial EFA accounted for nearly two-thirds of the variation in the data (64.8%)

According to the factor pattern matrix, items 2 and 9 posted loadings less than .40 across all four factors. Nunnally suggests researchers strive to interpret loadings above .40, stating that .30 should serve as an absolute lower bound for interpretability.³³ In this study, .40 was used as the item retention cut-off because it indicates that an item represents about 15% of the variation in the factor. Guadagnoli and Velicer also suggest that .40 is a minimum loading for acceptable interpretability, stating that it represents a weakly defined component at best.³⁴ They advocate striving for factor patterns that possess moderate component saturation, loadings of at least .60. Using the .40 cut-off, items 2 and 9 were eliminated from the final EFA. Additionally, items 2 and 15 posted factor loadings barely above Nunally's .40 interpretation rule-of-thumb (.437 and .417, respectively) and below Guadagnoli and Velicer's .60 moderate saturation recommendation. These two items also posted cross-loadings (i.e., loadings on a second factor) above .30 (.324 and .334, respectively). As a result, items 2 and 15 were also eliminated from further analyses. All other parameters were retained for the final EFA. The final EFA returned similar results as the initial analysis, despite the fact that the low-loading items were excluded, indicating that items 2, 8, 9, and 15 minimally impacted factor extraction. Rerunning the EFA procedure without items 2, 8, 9, and 15 returned a factor structure in which all items' factor loadings exceeded .60 (except for item 1 that loaded at .

595), which was well above Nunnally's .30 "rule of thumb" cut-off criterion for item retention and which satisfied Guadagnoli and Velicer's .60 suggestion for moderate saturation.^{33,34} Table 3 presents items' loadings for the final EFA analysis. As shown in the bottom portion of Table 2, the four factor solution, after refining the analysis, explained more than two-thirds of the variation in the data (69.8%).

The interfactor correlation matrix produced under the oblique rotation was also examined. This matrix indicated that the factor defined by items 4 and 5 (factor 2) was nearly orthogonal to the other factors. That is, factor 2 exhibited very low (almost) zero correlations with all of the other factors (.169 with factor 1; .180 with factor 3; -.243 with factor 4). Conversely, factors 1, 3, and 4 were moderately correlated with one another, with correlations ranging between -.306 and -.439.

Reliability/Internal Consistency of Resulting Factor Structure

The Cronbach's alpha reliability coefficient for the first factor was the highest, .82. Other Cronbach alpha reliability coefficient values were generally acceptable for scales in the early stages of development: .61 (Factor 2), .65 (Factor 3), and .77 (Factor 4).³³ Average item-total correlations also met acceptable levels for exploratory analyses.³⁵ The average item-total correlations for factors 1 through 4 were .68, .45, and .60, respectively. These results suggest that the scale items relating to each factor exhibit adequate measurement consistency, meeting a necessary (but not sufficient) prerequisite for construct validity (i.e., factor unidimensionality).

Factor 2 posted the lowest Cronbach's alpha reliability coefficient as well as the lowest average item-total correlation.²⁸ This low internal consistency is most likely due to the fact that only two items loaded on factor 2. Two additional items were added to the survey in 2006. The items were designed to measure the same facet of management commitment to health promotion as questions four and five, and were an effort to increase subscale internal consistency (see bottom portion of Table 1, Q.17 & 18). The utility of adding these additional items was explored via the validation analyses conducted in 2006.

Confirmatory Factor Analysis

While most EFA factor retention criteria suggested that one to three factors should be retained, internal consistency analysis found that a possible fourth factor had fairly strong reliability/consistency coefficients for an exploratory analysis (.77 and .60, respectively). Furthermore, the factor loadings for the fourth-factor items met acceptable retention levels (i.e., they all exceeded .60). Fabrigar, Wegener, MacCallum and Strahan argue that overfactoring is a less severe error than underfactoring.²⁷ Fabringar and colleagues, in making their argument, cite research by Fava and Velicer (1992) and Wood et al., (1996) that provides empirical support for the fact that overfactoring "introduces much less error to factor loading estimates than underfactoring" (p. 278).²⁷ Thus, the four factor model was used for subsequent confirmation and factor validation analyses.

As an omnibus test of discriminant validity, the fit of a four-factor model was compared with the fit of a one-factor model using CFA procedures (see Table 4). As might be expected, the four factor model fit the data well, yielding a nonsignificant chi-square value $X^2(38, N = 135) = 39.81, p = .39$. All item loadings were statistically significant at the $p = .05$ level, lending additional evidence of construct validity for each of the subscales assessed in the analysis. Significant correlations, however, were present at the latent factor level. These correlations were slightly stronger than those observed as a result of the EFA.

The one-factor model, on the other hand, fit the data poorly $X^2 (44, N = 135) = 119.04, p = .000$. Based on the ΔX^2 and ΔCFI tests, the one-factor model fit the data significantly worse than the four-factor model. Cheung and Rensvold, suggest the ΔCFI as a complement to the ΔX^2 when comparing model fit.³² Indices of practical fit also worsened for the one-factor model (see Table 4). Thus, model comparison results suggested better model fit for the more complex four-factor model, and we tentatively concluded that the four factors extracted during the EFA exhibited some level of discriminant validity.

Given the ability to assume a general level of discriminant validity, the next step involved testing the distinctness of each pair of factors individually. To accomplish this, the four-factor model was compared with a model in which two latent factor correlations were set to equal 1.0. All possible factor-pair correlations were successively set to equal 1.0 and then compared back to the four-factor model. In this way, we tested the discriminant validity of each possible latent factor.²² Chi-square difference tests between the four-factor model and the more constrained models indicated that discriminant validity held between each of the four factors individually (see Table 5): ΔCFI tests corroborated this conclusion.

Validation

The four-factor structure was then tested again in a second sample ($N = 178$), which was collected one year after the initial factor analyses in 2006. For this second application, confirmatory factor analysis was also used to determine model fit. In this separate sample, the model structure continued to show viability. While the X^2 was statistically significant in this sample ($p = .003$), the RMSEA, TLI, CFI and SRMR met or exceeded accepted levels of fit (.069, .98, .98, and .043, respectively). The two items that were added to bolster the reliability and internal consistency of factor 2 yielded significant and strong factor loadings (.92 and .99, respectively).

Factor Descriptions

Based on the factor structure from the EFA analysis (see Table 3), factor names were derived to describe each factor. Items with higher factor loadings were assigned more weight in the interpretation of the factor's meaning. Specifically, items loading on the first factor dealt with how well a site's business activities align with its health objectives. This first factor was labeled "*business alignment with health promotion objectives*." Items loading on the second factor asked about levels of education and training regarding the link between health and employee productivity. This factor was named "*awareness of the link between health and worker productivity*." Items loading on the third factor tapped into the concept of employees' perceptions that the worksite supports healthy behavior. This factor was labeled "*worksite support for health promotion*." Finally, the items that comprised the fourth factor assessed employees' perceptions of leadership support for health promotion in the workplace. Thus, this factor was labeled "*leadership support for health promotion*."

Upon completing the naming process, we were interested in interpreting our results from an intervention design and implementation point of view. In particular, we wanted to know whether the three groups of employees sampled in this study held different perceptions with regard to each of the four facets of management support for health promotion identified via the analyses. To explore this question, we created a weighted item composite to represent each factor. We then ran an analysis of variance (ANOVA) for each factor composite.

Group means were assessed for site leaders, cross disciplinary team members, and health services staff. For factors 1 (*business alignment with health objectives*) and 3 (*leadership support for health promotion*), health services staff rated the worksites significantly higher than site leaders and cross disciplinary team members in one-way ANOVA comparisons (see

Table 6). For factors 2 (*awareness of the link between health and worker productivity*) and 4 (*leadership support for health promotion*) no significant group differences materialized. Group differences across factors 1 and 3 suggest that intervention elements may need to be tailored to different worksite audiences or subpopulations.

Discussion

The results of this research suggest that management support and engagement for health promotion can be reliably assessed using the LBE, a brief, self-report questionnaire. A combination of the exploratory and confirmatory factor analyses were used to extract factors and demonstrate the validity of a four factor model containing 13 items. The use of confirmatory analyses provided a more rigorous test of factor unidimensionality and distinctiveness (i.e., construct and discriminant validity, respectively). The resulting four factors or subscales were labeled: 1.) business alignment with health promotion objectives, 2.) awareness of the link between health and worker productivity, 3.) worksite support for health promotion, and 4.) leadership support for health promotion.

Our goal was to develop a brief instrument that could be used at baseline as a diagnostic tool to assess organizational support and management engagement in health promotion. Also, we sought to develop an instrument that could be readministered at critical milestones after an intervention had been put in place. The idea being that it could be used to assess shifts in the environment over time. We were particularly interested in securing the ability to assess management support for health improvement over time.

Rather than relying on a simple global or overall assessment, we sought to develop a tool that assessed different aspects of management support and the organization's health promotion climate. Below, we list the items that comprise each of the four subscales on the LBE:

Factor #1: *Business alignment with health promotion objectives:*

- Our site health promotion programs are aligned with our business goals.
- Our site goals and plans advocate for the improvement of employee health.
- Site objectives for health improvement are set annually.

Factor #2: *Awareness of the link between health and worker productivity:*

- Employees at all levels are educated about the true cost of health care and its effects on business success.
- All levels of employees are educated about the impact a healthy workforce can have on productivity and cost management.
- Site leadership shares information with employees about the effect of employee health on overall business success.
- All levels of management are educated regarding the link between employee health and productivity and cost management.

Factor #3: *Worksite support for health promotion:*

- This site offers incentives for employees to stay healthy, reduce their high risk behaviors, and/or practice healthy life styles.
- Our health benefits and insurance programs support prevention and health promotion.
- Our work teams provide support for participation in health promotion programs.

Factor #4: *Leadership support for health promotion:*

- The organization provides our site leadership training on the importance of employee health.
- Our leaders view the level of employee health and well-being as one important indicator of the site's business success.
- Our site leadership is committed to health promotion as an important investment in human capital.

Given these factors, the LBE can be used as part of preliminary or formative research activities, exposing specific areas where an organization's health promotion climate might hinder intervention fidelity and effectiveness. The questionnaire can also be a valuable tool for tracking and monitoring changes in management support that result from comprehensive worksite health interventions or other health-related programmatic activities.

Our analyses of the LBE indicate that it can be effective in identifying differences in health climate perceptions across employee groups. Specifically, the one-way ANOVA comparisons of weighted group-by-factor means points to health services staff as a segment of employees who may perceive awareness of the health-productivity link and worksite support for organizational health promotion programs differently from their counterparts. These differences in perception may simply be self-serving on the part of health services staff or they may reflect actual differences in how these particular employees process and interpret the words and actions of management. In any event, the opinions of health services staff may not yield the most valid or useful assessment of management support and organizational climate for health promotion activities. Based on the results of this research, we suggest that researchers strive to obtain formative data from a variety of audience segments within an organization, including employees themselves.

While we still advocate for conferring and consulting with health services staff when developing programs, as these individuals are generally the strongest internal champions of health promotion interventions, we also suggest that researchers seek input and engage employees at all levels of an organization and across all job roles, including mid-level and top-level leadership. Obtaining feedback and opinions from a variety of internal audiences will provide researchers with more detailed information about the health promotion climate within an organization and across organizational audiences. It will also aid researchers in quantifying the constructs of worksite health promotion climate and management support, thus highlighting potential challenges and hurdles that could affect intervention success.

Not only do we propose that the LBE may be valuable in the formative research process, but we also feel that it could become an important element of intervention evaluation. Because different internal audiences (e.g., leadership, human resources, and health services) may possess different perceptions of alignment, awareness, and support for health promotion at intervention baseline, tracking group changes over time using the LBE should help researchers identify incremental changes in health climate constructs. Likewise, tracking each LBE factor over the course of the intervention could help pinpoint support or awareness problems during intervention implementation, when adjustments are still feasible. Finally, the LBE factors could be used to support assertions of intervention effectiveness, i.e., if factor means increase significantly over time from baseline estimates.

Limitations

Despite the utility of initial findings, a few limitations exist with regard to the analyses outlined above. First, response rate was only adequate in both 2005 and 2006. Guadagnoli and Velicer state that a minimum of 150 responses should be analyzed for proper factor structure

determination in EFAs when components possess moderate saturation (i.e., loadings of .60).³⁴ Boomsma recommends at least 200 data points for proper model estimation in a CFA context.³⁶ The samples utilized in these analyses were slightly below the EFA and CFA sample size recommendations.

Second, the model development sample and validation sample were both collected from the same organization and the same 12 worksites. As a result, we cannot generalize the factor structure to different types of organizations or other economic situations. Additional validation research is needed to further confirm the factor structure identified in this study and to help establish the LBE's value across various application circumstances.

Implications for Future Research

To build on our findings, further research should involve obtaining a larger number of responses from additional independent samples of a variety of organizations. Subsequently, these samples should be subjected to confirmatory factor analyses. Similar target-model versus one-factor model omnibus comparisons should be made to assess discriminant validity, and factor-restricted model comparisons should be conducted to assess individual factor distinctness. Test-retest reliability analysis should also be conducted to confirm subscale consistency. These types of additional analyses would help solidify the underlying factor structure and the reliability of the revised LBE for future instrument applications, perhaps allowing the LBE to be shortened from 18 items to the 13 subscale items identified in these analyses. Researchers could also begin testing the predictive validity of these health promotion climate constructs, assessing how the constructs differentially impact intervention and health promotion program success in worksites.

While much research is still needed to assess reliability and validity across various settings and samples, the results of this study introduce a solid base for future exploration of the worksite health promotion climate milieu. What is more, the four factors identified through these initial analyses appear to be useful for intervention planning and evaluation. As a result, we would encourage other researchers and practitioners to use the LBE as a diagnostic tool for intervention planning, but also as a tool for tracking intervention effectiveness over time. Only with repetition, multiple applications and administrations, will we be able to garner a full understanding of the reliability, validity, and value of the LBE.

ACKNOWLEDGMENT OF FUNDING

Funding for this study was provided by the National Heart, Lung and Blood Institute, (Grant # R01 HL79546). However, its contents are the sole responsibility of the authors and do not necessarily represent the official views of NHLBI.

Reference

1. Green, LW.; Kreuter, MW. *Health Promotion Planning: An Educational and Ecological Approach*. Vol. 3rd ed. Mountain View, CA: Mayfield; 1999.
2. O'Donnell MP. Definition of health promotion: part III: expanding the definition. *Am J Health Promot* 1989;3:5. [PubMed: 10292073]
3. U.S. Department of Health and Human Services. *Healthy People 2010*. Vol. 2nd ed. Washington, DC: U.S. Government Printing Office; 2000 Nov. With Understanding and Improving Health and Objectives for Improving Health.
4. Stokols D. Establishing and maintaining healthy environments: toward a social ecology of health promotion. *Am Psychol* 1992;47:6–22. [PubMed: 1539925]
5. DeJoy DM, Wilson MG. Organizational health promotion: broadening the horizon of workplace health promotion. *Am J of Health Promot* 2003;17:337–341. [PubMed: 12769047]

6. Goetzel, RZ. Examining the Value of Integrating Occupational Health and Safety and Health Promotion Programs in the Workplace. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health; 2004.
7. Sorensen, G.; Barbeau, E. Steps to a Healthier U.S. Workforce: Integrating Occupational Health and Safety and Worksite Health Promotion: State of the Science. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health; 2004.
8. Goetzel RZ. Essential building blocks for successful worksite health promotion programs. *Manage Employee Health Benef* 1997;6:1.
9. O'Donnell M, Bishop C, Kaplan K. Benchmarking best practices in workplace health promotion. *The Art of Health Promot* 1997;1:12.
10. Goetzel, RZ.; Guindon, A.; Humphries, L.; Newton, P.; Turshen, J.; Webb, R. Health and Productivity Management: Consortium Benchmarking Study Best Practice Report. Houston, TX: American Productivity and Quality Center International Benchmarking Clearinghouse; 1998.
11. Goetzel RZ, Ozminkowski RJ, Ascitutto AJ, Chouinard P, Barrett M. Survey of Koop Award winners: life-cycle insights. *The Art of Health Promot* 2001;5:2.
12. Flin R, Mearns K, O'Connor P, Bryden R. Measuring safety climate: identifying the common features. *Safety Sci* 2000;34:177–192.
13. Hopkins, A. Making Safety Work: Getting Management Commitment to Occupational Health and Safety. St. Leonards, Australia: Allen & Unwin; 1995.
14. Zohar D. Safety climate in industrial organizations: theoretical and applied implications. *J Appl Psychol* 1980;65:96–102. [PubMed: 7364709]
15. Ribisl KM, Reischl TM. Measuring the climate for health in organizations: development of the worksite health climate scales. *J Occup Med* 1993;35:812–824. [PubMed: 8229333]
16. Barrett L, Plotnikoff RC, Raine K, Anderson D. Development of measures of organizational leadership for health promotion. *Health Educ Behav* 2005;32:195–207. [PubMed: 15749966]
17. Leading by Example: Improving the Bottom Line Through a High Performance, Less Costly Workforce. Washington, D.C.: Partnership for Prevention; 2004. Partnership for Prevention.
18. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
19. Hair, JF.; Anderson, RE.; Tatham, RL.; Black, WC. Multivariate Data Analysis. Upper Saddle River, NJ: Prentice Hall; 1998.
20. Gorsuch, RL. Factor Analysis. Hillsdale, NJ: Erlbaum; 1983.
21. Hinkin TR. A review of scale development practices in the study of organizations. *J Manage* 1995;21:967–988.
22. Mallard AGC, Lance CE. Development and evaluation of a parent-employee interrole conflict scale. *Soc Indic Res* 1998;45:343–370.
23. Velicer WF, Jackson DN. Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. *Multivar Behav Res* 1990;25:1–28.
24. Kaiser HF. The application of electronic computers to factor analysis. *Educ and Psychol Meas* 1960;20:141–151.
25. Cattell RB. The scree test for the number of factors. *Multivar Behav Res* 1966;1:245–276.
26. Hayton JC, Allen DG, Scarpello V. Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. *Organ Res Methods* 2004;7:191–205.
27. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods* 1999;4:272–299.
28. Crocker, L.; Algina, J. Introduction to Classical and Modern Test Theory. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers; 1986.
29. Green SB, Lissitz R, Mulaik SA. Limitations of coefficient alpha as an index of test unidimensionality. *Educ Psychol Meas* 1977;37:827–838.
30. Hu LT, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol Methods* 1998;3:424–453.

31. Steiger JH, Shapiro A, Browne MW. On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika* 1985;50:253–264.
32. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling* 2002;9:233–255.
33. Nunnally, JC. *Psychometric Theory*. Vol. 2nd ed. New York: McGraw-Hill; 1978.
34. Guadagnoli E, Velicer WF. Relation of sample size to the stability of component patterns. *Psychol Bull* 1988;103:265–275. [PubMed: 3363047]
35. Streiner, DL.; Norman, GL. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Vol. 2nd ed. New York: Oxford University Press; 1995.
36. Boomsma, A. The robustness of LISREL against small sample sizes in factor analysis models. In: Jöreskog, KG.; Wold, H., editors. *Systems Under Indirect Observation: Causality, Structure, Prediction*. Amsterdam:: North-Holland; 1982. p. 149-173.

Table 1

Leading by Example questionnaire items

Items Analyzed in 2005

1. Our site leadership is committed to health promotion as an important investment in human capital.
2. Our site leadership provides adequate financial support for health promotion.
3. Our site health promotion programs are aligned with our business goals.
4. All levels of management are educated regarding the link between employee health and productivity and cost management.
5. Employees at all levels are educated about the true cost of health care and its effects on business success.
6. Our site goals and plans advocate for the improvement of employee health.
7. Site objectives for health improvement are set annually.
8. Our site provides management support for health promotion by issuing messages from the site leader about the importance of employee health to the site.
9. Our site provides support for participation in health promotion programs.
10. Our work teams provide support for participation in health promotion programs.
11. Dow provides our site leadership training on the importance of employee health.
12. Our health benefits and insurance programs support prevention and health promotion.
13. This site offers incentives for employees to stay healthy, reduce their high risk behaviors, and/or practice healthy life styles.
14. Our leaders view the level of employee health and well-being as one important indicator of the site's business success.
15. Overall, our site promotes a culture of health and well being.

Items Added in 2006

16. The effectiveness of our health promotion programs are evaluated based upon accepted definitions of success.
17. Site leadership shares information with employees about the effect of employee health on overall business success.
18. All levels of employees are educated about the impact a healthy workforce can have on productivity and cost management.

Table 2
Initial and final exploratory factor analyses, total variance explained

Factor	Eigenvalues	% of Variance Explained	Cumulative % of Variance Explained	Random Variable Eigenvalues (parallel Analysis)
<i>Initial Exploratory Factor Analysis</i>				
1	6.443	42.955	42.955	1.691
2	1.343	8.955	51.910	1.445
3	1.041	6.940	58.850	1.363
4	.894	5.957	64.806	1.147
<i>Final Exploratory Factor Analysis</i>				
1	4.497	40.886	40.886	1.555
2	1.276	11.601	52.487	1.390
3	1.034	9.397	61.884	1.241
4	.875	7.951	69.835	1.103

Table 3

Final solution, pattern matrix factor loadings

	Factor			
	1	2	3	4
Q6. site goals advocate for improving employee health	.869	-.063	.076	.005
Q3. health programs aligned with business goals	.845	.060	-.106	-.006
Q7. site objectives for health improvement set annually	.776	.020	.079	-.079
Q5. employees educated re: true cost of health care	.027	.852	.233	.180
Q4. levels of mgt educated re: link b/w healthy and productivity	.039	.762	-.150	-.323
Q13. site offers incentives to employees to stay healthy	-.100	.007	.804	-.162
Q12. health benefits/insurance programs support prevention	.164	.013	.709	.121
Q10. work teams support participation in health programs	.052	.125	.611	-.157
Q11. Site provides site leaders training on importance of employee health	.038	.170	-.068	-.818
Q14. leaders view the level of employee health as one important indicator of success	.050	-.032	.202	-.757
Q1. health important investment in human capital	.212	-.130	.185	-.595

Factor 1: Business alignment with health promotion objectives

Factor 2: Awareness of the link between health and worker productivity

Factor 3: Worksite support for health promotion

Factor 4: Leadership support for health promotion

Table 4

Model goodness-of-fit, omnibus test of discriminant validity

	df	χ^2	RMSEA	TLI	CFI	$\Delta\chi^2$	ΔCFI
4 factor model	38	39.81	.025	1.00	1.00		
1 factor model	44	119.04	.12	.89	.91	79.23 ^{***}	.09

*** Statistically significant at the p=.01 level

Table 5

Tests of individual factor distinctness

Assumption	df	X ²	RMSEA	TLI	CFI	ΔX ²	ACFI
Target model vs. model in which Factor 1 = Factor 2	39	55.37	.060	.97	.98	15.56 ^{***}	.02 [*]
Target model vs. model in which Factor 1 = Factor 3	39	65.62	.074	.95	.97	25.51 ^{***}	.03 [*]
Target model vs. model in which Factor 1 = Factor 4	39	83.77	.099	.92	.95	43.96 ^{***}	.05 [*]
Target model vs. model in which Factor 2 = Factor 3	39	49.80	.051	.98	.99	9.99 ^{**}	.01 [*]
Target model vs. model in which Factor 2 = Factor 4	39	49.22	.047	.98	.99	9.41 ^{**}	.01 [*]
Target model vs. model in which Factor 3 = Factor 4	39	60.98	.069	.96	.97	21.17 ^{***}	.03 [*]

* Statistically significant at the p=.05 level

** Statistically significant at the p=.01 level

*** Statistically significant at the p=.001 level

Table 6

Group comparisons by factor

TeamFactor/Total Responses	Site Leaders(69)	Cross Disc. (32)	Health Services (33)
1. Business alignment with health objectives	2.969 ^A	2.997 ^A	3.618 ^B
2. Awareness of the economics of health and productivity	2.742	2.520	2.707
3. Worksite support for health promotion	2.905 ^A	2.928 ^A	3.336 ^B
4. Leadership support for health promotion	3.309	3.109	3.255

Ratings based on a scale of 1 – 5, where 1 = “strongly disagree” and 5 = “strongly agree”

Shaded rows indicate that overall F value in ANOVA is significant at .05 level.

Superscript letters indicate significant group differences at the .05 level using LSD post-hoc contrasts in a one-way factor ANOVA