

ENHANCING DoD CIVILIAN LEADER DEVELOPMENT WITH THE CIVILIAN LEADER IMPROVEMENT BATTERY (CLIMB)

Robert Kilcullen*

U.S. Army Research Institute for the Behavioral and Social Sciences
Arlington, VA 22202

ABSTRACT

The U.S. Army Research Institute (ARI) created a leader development tool called the Civilian Leader Improvement Battery (CLIMB) to help incumbent and upcoming civilian leaders reach their full performance potential in leadership positions. The CLIMB assesses individual strengths and weaknesses with respect to critical supervisory competencies and also provides test-takers with clear feedback regarding their strengths and weaknesses, plus recommendations for improving identified areas of concern. Because the CLIMB scales have been demonstrably related to civilian leader job performance, the CLIMB assessment and feedback system offers the potential for civilian leaders to make substantive improvements in their job performance.

1. INTRODUCTION

The Department of Defense (DoD) is implementing in stages a new personnel system for civilian employees called the National Security Personnel System (NSPS). Although NSPS has many potential benefits, it will likely expand the scope and responsibility of civilian supervisors to manage work and monitor subordinate performance.

To help civilian supervisors meet the challenges of the new personnel system, the U.S. Army Research Institute (ARI) created a leader development tool named the Civilian Leader Improvement Battery (CLIMB). The CLIMB is a self-report instrument that measures motivational constructs related to leader performance. The CLIMB also provides test-takers clear and meaningful feedback on their strengths and weaknesses with respect to the constructs, and practical strategies for improving job performance based on identified areas of concern.

A distinctive feature of the CLIMB is the empirical validity of its scales for predicting the job performance of Department of the Army (DA) civilian leaders. This, along with the ability to translate scale scores into developmental feedback and the overall ease of use of the test, were critical factors in selecting the CLIMB for operational use.

The history of the development of the CLIMB includes two parts: (1) the development of the operational CLIMB measuring instrument, which began over a decade ago, with recent modifications to adapt to the needs of the NSPS system, and (2) the development of the feedback package for NSPS, which was completed recently. This paper focuses primarily on the development of the measuring instrument, but also provides an overview of this instrument's place in the overall CLIMB assessment and feedback system.

2. METHOD

2.1 Subjects & Procedure

A total of 1,775 first-line civilian supervisors in the Department of the Army served as participants in the research that produced the operational CLIMB measuring instrument. A wide variety of occupations and grade levels were represented in the sample. The demographic composition of the subjects was as follows: 30 percent were female, 74 percent were White, and 17 percent were Blacks. Nearly all of the subjects completed high school, 43 percent had some college experience, 23 percent graduated from college, and 10 percent had graduate or professional degrees. The participants were distributed across 54 locations throughout the Continental United States.

The participants were briefed about the purpose of this research and were then administered a large battery of tests, including an In-Basket exercise and self-report items, some of which were eventually incorporated into the CLIMB. The order of test administration was randomized.

Matching of predictor and criteria scores resulted in effective sample sizes between 1,665 and 1,769 cases.

2.2 The CLIMB Test

The CLIMB is a rational biodata test. Rational biodata tests measure temperament characteristics by asking multiple-choice questions about the test-taker's past behavior and reactions to life events. Creating these tests involves identifying motivational constructs (e.g.,

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 01 NOV 2006		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Enhancing Dod Civilian Leader Development With The Civilian Leader Improvement Battery (Climb)				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences Arlington, VA 22202				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADM002075.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 4	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Work Motivation) likely to predict the criterion of interest and then writing items that sample behavior believed to be manifestations of these attributes.

In this research, each member of a panel of psychologists generated 10 to 15 items referring to past behaviors and life events thought to be indicative of each targeted construct. Candidate items were reviewed by the panel for construct relevance, response variability, relevance to the civilian population, readability, non-intrusiveness, and neutrality with respect to social desirability. A consensus decision was reached concerning the best items for each construct. The surviving items were then pilot-tested and revised based on item analysis.

The biodata items were designed to measure job-related manifestations of the targeted constructs. For example, questions targeting Cognitive Flexibility focused on trying new ways of accomplishing work, rather than trying new hobbies or new places to eat. The idea is to tap into the aspects of the construct most relevant to the performance area of interest and, thereby, to improve the validity of the test for this performance area (Kilcullen, White, Mumford, & Mack, 1995). It also makes feedback more meaningful to participants in the workforce in terms of diagnosing work-related problems and providing useful strategies for improving on-the-job performance.

Item responses were scored based upon their presumed relationship to the construct, and item scores were summed to form scale scores having substantive meaning. Previous research with rational biodata scales suggests that these assessments can reliably and validly measure their intended constructs and may be less fakable than traditional personality measures (Kilcullen, White, Mumford, & Mack, 1995).

This design process produced over 400 biodata items for measuring 21 temperament characteristics thought to be relevant to supervisory job performance. These temperament characteristics were based on a job analysis of supervisory positions and a literature review of individual difference predictors of leader job performance and advancement.

It would not be practical to use a 400-item test in an operational DoD civilian leader assessment and feedback program. Therefore, the research sought to identify the subset of temperament characteristics and the items measuring them for inclusion in an operational version of the CLIMB. The goal was to create a shorter, more practical 25-minute test consisting of the most promising scales and items.

To create this shortened form, the validities of the original scales were assessed against two job performance criteria to determine the most promising scales for predicting job performance in supervisory positions. Based on these analyses, scales measuring six attributes were targeted for inclusion in the operational version of the CLIMB. These were:

- *Work Motivation.* Willingness to give one's best effort and to work hard to achieve work objectives.
- *Cognitive Flexibility.* Willing to entertain new approaches to solving problems. Enjoys formulating new plans and ideas. Initiates and accepts change and innovation.
- *Peer Leadership.* Willingness to assume the role of leading one's peers. Seeks positions of authority and influence. Comfortable being the person responsible for the group's performance.
- *Stress Tolerance.* Being calm and worry-free. Able to maintain composure under pressure.
- *Social Perceptiveness.* Being perceptive about the feelings and agendas of others, and taking this information into account to work more smoothly with coworkers.
- *Team Orientation.* Being tactful and diplomatic. Willingness to work well with others. Able to establish supporting and trusting relationships with coworkers.

In addition to the six scales just described, a "Response Distortion" scale was also included in the CLIMB to detect individuals who appear to describe themselves more favorably than they likely are. Research has shown that the validities of self-report temperament tests improve when response distortion is taken into account in validation analyses (White & Kilcullen, 1998).

The full array of items for each of the six predictor scales still resulted in a test requiring approximately 45 minutes to complete. Therefore, for each scale, "best bet" items for use in the CLIMB were identified based on psychometric analyses, item validities, and low item correlations with the Response Distortion. The final iteration of the CLIMB consisted of 72 items, which took 25 minutes to complete.

2.3 Performance Criteria

The validities of the CLIMB scales were assessed against two job performance criteria. One criterion measure consisted of confidential, off-the-record ratings of the participants' job performance made by their supervisors. These ratings were collected in separate sessions while the participants completed the test battery. Participants were evaluated on 19 performance dimension

scales. Six scales tapped generic job performance dimensions such as Work Quantity, Accuracy, and Job Knowledge. Thirteen scales were based on supervisory performance dimensions identified in a job analysis. These included scales pertaining to maintaining employee morale, resolving conflicts, providing personal/career counseling, and maintaining standards and discipline. The alpha reliability of the 19-scale instrument was .963.

The second criterion consisted of a self-report measure of administrative records reflecting job performance over the past four years. Included were items relating to the frequency of awards, letters of commendation, performance-based pay raises, disciplinary actions as well as several other verifiable indicators of performance. This measure was administered concurrently with the predictor measures. Responses were summed to form a composite score, with an alpha reliability of .65. Previous research in the U.S. Army (Campbell, 1987) indicated that self-report administrative data were more accurate than official records due to errors in processing personnel actions and delays in updating personnel files.

3. RESULTS

Table 1 presents descriptive statistics and internal consistency reliability coefficients for the CLIMB scales. Reliability coefficients over .70 were achieved for all of the scales, indicating a sufficient level of internal consistency in the scales. The median correlation between the CLIMB scales was $r = .25$ ($p > .01$), reflecting a relatively modest level of overlap in scale variance.

Table 1. CLIMB scales descriptive statistics and reliabilities ($n=1,770-1,775$)

Predictor	Mean	s	α
1. Cognitive Flexibility	3.09	0.45	.70
2. Stress Tolerance	3.14	0.50	.72
3. Social Perceptiveness	3.43	0.46	.70
4. Work Motivation	3.47	0.51	.73
5. Team Orientation	3.17	0.47	.72
6. Peer Leadership	3.46	0.58	.72

Zero-order correlations were computed to assess the criterion-related validities of the CLIMB scales (see Table 2). Each CLIMB scale predicted at least one of the job performance criteria, and only two out of the 12 validity coefficients were not statistically significant in the expected direction. Supervisory ratings of performance was best predicted by Social Perceptiveness ($r = .21$, $p < .01$), Peer Leadership ($r = .20$, $p < .01$), and Work Motivation ($r = .15$, $p < .01$). The administrative performance criterion was best predicted by Peer

Leadership ($r = .23$, $p < .01$), Work Motivation ($r = .19$, $p < .01$), and Team Orientation ($r = .18$, $p < .01$).

Table 2. Zero-order correlations between CLIMB predictor scales and job performance criteria ($n=1,665-1,769$)

Predictor	Supervisory	Administrative
	Ratings	Criteria
1. Cognitive Flexibility	.06*	.15**
2. Stress Tolerance	.12**	-.04*
3. Social Perceptiveness	.21**	-.01
4. Work Motivation	.15**	.19**
5. Team Orientation	.05*	.18**
6. Peer Leadership	.20**	.23**

Note: * $p < .05$; ** $p < .01$

Overall, the pattern of results was such that each CLIMB scale significantly predicted one or both indices of job performance. In contrast to more senior levels of leadership, supervisors typically have frequent and direct contact with their subordinates (Jacobs & Jaques, 1991). Therefore, it is not surprising that the CLIMB scales pertaining to interacting well with others, as well as the willingness to work hard and the desire to lead, were the primary predictors of supervisory job performance

4. DISCUSSION

This developmental research showed that the CLIMB scales were demonstrably related to DoD civilian supervisory job performance. Because these assessments are job-relevant, they have the potential to provide to supervisors, and those aspiring to supervisory positions, meaningful feedback on their own strengths and weaknesses with respect to critical competencies.

Follow-up research by Kilcullen, White, Zaccaro, & Parker (2000) investigated the validity of the CLIMB scales for predicting leader performance across organizational levels. The job performance of approximately 340 supervisors (GS-13 and below), managers (GS-14 and GS-15), and Senior Executive Service (SES) leaders was assessed using confidential ratings of the participants' job performance made by their superiors. The results revealed that the CLIMB scales predicted job performance at each level of leadership. At the supervisory level, job performance was best predicted by Work Motivation ($r = .32$) and Peer Leadership ($r = .26$), each $p < .01$. Performance at the managerial level was best predicted by Peer Leadership ($r = .23$), Cognitive Flexibility ($r = .21$), and Work Motivation ($r = .19$), each $p < .01$. Performance at the SES level was best predicted by Peer Leadership ($r = .29$, $p < .01$).

In a subsequent longitudinal study using a different population, a subset of the CLIMB scales was administered to 169 Lieutenant Colonels. The career progression of these officers was tracked for a number of years to identify predictors of advancement to General Officer rank. The results showed that the CLIMB scales of Peer Leadership ($r = .22$) and Work Motivation ($r = .20$) significantly predicted promotion to General Officer (both $p < .05$).

Taken together, the results from the initial study and follow-up research indicate that the characteristics measured by the CLIMB scales are broadly applicable across a variety of leadership levels and occupations.

To facilitate DoD civilian leader development, a web-based version of the CLIMB assessment and feedback system was created and is available to all Department of the Army and Department of the Navy civilian employees. Web-based administration of the CLIMB is quick, convenient, and cost-effective. The web test can be accessed at any time. The test is scored immediately after the last question has been answered, and test-takers have immediate access to their CLIMB feedback package.

The CLIMB feedback package consists of three parts. The first part is a brief introduction to the CLIMB – what it measures and how it has been validated. The second part presents graphs illustrating the individual's relative standing on each CLIMB scale based on the normative data collected in this research.

The third part consists of narratives that explain the results presented in the graph. The tailored narratives describe the individual's strengths and weaknesses with respect to the motivational constructs measured by the CLIMB. To guide future self-development activities, the narratives also provide practical recommendations for improvement in each area of concern identified by the CLIMB.

By providing job-relevant assessment and feedback, this system will help civilian leaders to focus their self-development activities in areas that are likely to have payoff in terms of improved job performance and advancement potential.

REFERENCES

Campbell, J.P. (1987). Improving the selection, classification, and utilization of Army enlisted personnel: Annual report, 1985 fiscal year. (ARI Technical Report 745). (AD A193 343)

Jacobs, T.O., and Jaques, E. "Executive leadership." In R. Gal, A.D. Mangelsdorff, et al. (eds.) *Handbook of Military Psychology*, pp. 431-447. Chichester, England, UK: John Wiley & Sons. 1991.

Kilcullen, R.N., White, L.A., Mumford, M.D., and Mack, H. "Assessing the Construct Validity of Rational Biodata Items," *Military Psychology*, Vol 7(1), pp. 17-28, 1995.

Kilcullen, R.N., White, L.A., Zaccaro, S., & Parker, C. "Predicting Managerial and Executive Performance". Paper presented at the 15th annual Society for Industrial and Organizational Psychology conference, New Orleans, LA, April 2000

White, L. A., and Kilcullen, R. N. "Effects of socially desirable responding on the criterion-related validity of temperament measures." Paper presented at the 14th Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX, April 1998.