



How to Measure Empowerment

The Employee Empowerment Questionnaire can help.

by
Bob E. Hayes

THE EMPLOYEE EMPOWERMENT QUESTIONNAIRE (EEQ) allows companies to identify the extent of empowerment in their organizations and determine how empowerment is related to other organizational variables, such as perceived job characteristics, job satisfaction, and intentions to quit. The EEQ also enables a company to compare its level of empowerment with other companies' levels. Before discussing how to measure empowerment, however, several important issues on measuring attitudes must be addressed: constructs and observable indicators, reliability, and validity.

Constructs and observable indicators

A construct is a theoretical construction about the nature of human behavior.¹ Personality characteristics (such as anxiety and self-esteem), management styles (such as consideration of subordinates and initiating structure), and job attitudes (such as job satisfaction and social support at work) are examples of constructs. Constructs help people understand human behavior.

Since constructs can't be directly observed, they are usually discussed in abstract terms. But to have real meaning, they must be discussed using observable traits. This can be done by developing measures of the constructs, which are referred to as observable indicators.

People often use constructs in everyday language to describe a set of observable indicators. For example, Mary might draw the conclusion that Joe is happy (construct) because he is smiling and laughing. Likewise, Mary might think Joe is anxious (construct) because he is fidgeting and sweating. In both instances, Mary has inferred Joe's internal (unobservable) state by examining his observable indicators.

Inferences about underlying constructs can be made in a more systematic fashion using questionnaires. For example, if a company wants to measure the amount of social support employees perceive they have at work, it could use a questionnaire that asks several standardized questions,

such as "If you had too much work to do, how much would your supervisor be willing to help you if he/she had to go out of his/her way?" and "If you had unclear goals at work, how much could your co-workers be relied on to tell you what you should do about this problem?" The employees' responses would be the observable indicators that provide information about the construct (level of social support at work).

When using questionnaires to measure constructs, reliability and validity must be addressed in survey development and evaluation. Reliability allows the survey designers to determine the degree of systematic variance in the questionnaire, while validity allows the designers to, in a sense, label this systematic variance.

Reliability

Reliability is the degree to which measurements are free from random errors. Reliability can be thought of as the relationship between the true underlying score and the observable score. Random error decreases the measurement's reliability; that is, as random error is introduced into measurement, the observed score is not a good reflection of the true underlying score. For one to feel confident that a questionnaire's scores accurately reflect the underlying dimension, the questionnaire must have high reliability. Although many types of reliability exist, internal consistency reliability is vital to surveys.²

Internal consistency indicates the extent to which the items in the measurement are related to each other. The higher the interrelationship among the items, the higher the internal consistency. If a questionnaire is designed to measure one underlying construct, the items are expected to be related to each other—that is, people who respond in one way to an item are likely to respond the same way to the other items in the measure.³

There are several statistical indexes used to estimate the degree of internal consistency. The most commonly used index is Cronbach's coefficient alpha.⁴ Basically, this alpha coefficient indicates

Cronbach's Alpha Estimate

Cronbach's alpha estimate is used to indicate how highly the items in a questionnaire are interrelated. Cronbach's estimate is calculated using the variances of individual items and the covariances between items. This estimate, however, can also be calculated using the correlations between items. Given that items within a questionnaire use the same scale, both approaches give similar estimates. The latter approach, however, is easier to understand. The formula for Cronbach's reliability estimate is:

$$r_{xx'} = (K/[K - 1])(1 - [(\sum X_{ii})/(\sum X_{ii} + \sum X_{ij})]); \text{ where } i \neq j$$

The elements X_{ii} and X_{ij} are the elements of the covariance matrix or correlation matrix, and K is the number of items in the scale. The numerator $(\sum X_{ii})$ indicates that the elements in the diagonal of the covariance matrix be added together. The denominator $(\sum X_{ii} + \sum X_{ij})$ indicates that all the elements in the covariance (correlation) matrix be added together.

Table A represents a correlation matrix for a three-item scale. Using the formula just given, the internal consistency of this scale can be calculated:

$$r_{xx'} = (3/2)(1 - [3/(3 + 2.72)]) = 0.71$$

Table A. Correlations Among the Three Items

	Item No. 1	Item No. 2	Item No. 3
Item No. 1	1.0	0.45	0.38
Item No. 2	0.45	1.0	0.53
Item No. 3	0.38	0.53	1.0

the degree to which items are related to each other. This index can range from 0 to 1. A reliability of 0 indicates that the observed score is not related to the underlying true score; a reliability of 1 indicates that the observed score is a perfect indicator of the underlying true score. Generally, a reliability of 0.8 or greater is an acceptable level of reliability.⁵ (For more information on the index, see the sidebar "Cronbach's Alpha Estimate.")

Although reliability is an important ingredient in the evaluation of a questionnaire, it cannot solely determine the quality of the questionnaire. The questionnaire's validity must also be addressed.

Validity

Validity refers to the degree to which evidence supports the inferences made from scores derived from measurements, or the degree to which the scale measures what it is designed to measure. Unlike reliability, there is no single statistic that provides an overall index of the validity of inferences about the scores.

The methods for gathering evidence of validity can be grouped into three categories: content-related evidence, criterion-related evidence, and construct-related evidence.⁶ These methods should not be considered distinct types of validity. The labels simply enable people to discuss the types of information that might be considered when determining the validity of the inferences.

Content-related evidence is concerned with the degree to which the items in the questionnaire are representative of a

"defined universe" or "domain of content." The domain of content typically refers to all possible items that could have been used in the questionnaire. The goal of content-related validity is to have a set of items that best represents the defined universe.

Criterion-related evidence is concerned with examining the systematic relationship (usually in the form of a correlation coefficient) between the questionnaire and another measure, or criterion. In this case, what the criterion is and how it is measured are of central importance. The main question to be addressed in criterion-related validity is how well the survey can predict the criterion.

Construct-related evidence is concerned with the questionnaire as a measurement of an underlying construct. Unlike criterion-related validity, the primary focus is on the questionnaire itself rather than on what the survey predicts. Construct-related evidence is derived from both previous validity strategies. A high degree of correlation between the questionnaire and other scales that purportedly measure the same construct is evidence of construct-related validity. Construct-related validity can also be evidenced by a low correlation between the questionnaire and other scales that measure a different construct.⁷

Developing the attitude survey

The development of attitude questionnaires can be broken into three steps:

1. *Define the construct to be measured.* One way to define a construct is by using words or other constructs. Basically, the construct is defined in the way that words are defined in the dictionary. This type of definition is referred to as the constitutive definition.⁸ Definitions from applicable literature can be used for constructs.

2. *Generate items to measure the construct.* These items, which act as observable indicators, bring the construct into the observable world by specifying how it can be measured. The measurement of a construct is referred to as an operational definition. The items help define the construct in more specific terms, which in turn, help distinguish the construct from related constructs.

Items can be generated through brainstorming sessions. The number of items needed to measure a construct depends on the type of construct being measured. For example, a narrowly defined construct requires only a few items for adequate measurement. For example, if a survey were conducted to measure people's ages, only one question would be needed: "How old are you?" The same information would be obtained if other questions such as "What is your age?" or "When were you born?" were asked.

On the other hand, if the construct has a broad definition, many items might be needed to measure or capture the entire construct. For example, to measure empowerment, more than one item is needed because the construct of empowerment encompasses a much broader content domain.

3. *Evaluate the items.* The quality of the items generated in step 2 must be evaluated. The generated items should be compiled into a trial questionnaire, which is given to a sample of respondents to complete. It is important that these respondents be a representative sample of the population for which the questionnaire is targeted. For example, a trial empowerment questionnaire should not be given to fired or retired employees; their responses would not help in the evaluation. Instead, the trial sur-

vey should be given to full-time and part-time employees because they are the targeted population.

After the respondents complete the trial survey, the responses must be statistically analyzed to determine the quality of the items. Their quality is judged by several statistical criteria that indicate the degree to which the items are related to one another. Since all items are designed to be observable indicators of the same underlying construct, responses to items should be interrelated; that is, if all items are good items, employees who score high on one item should score high on the rest of the items and employees who score low on one item should score low on the rest of the items. Two commonly used statistical analyses are the correlational analysis and the factor analysis.

Correlational analysis determines the linear relationship between two variables. One statistical index often used is the corrected item-total correlation, which is the correlation between one item and the remaining items in the scale. To be considered good, an item's corrected item-total correlation should be relatively high. A low corrected item-total correlation indicates that the item is not related to the other items in the scale; in other words, the item does not measure the same thing that the other items are measuring.

Factor analysis is a more sophisticated method of determining relationships among items and underlying factors (constructs). Factor analysis indicates the number of factors that the set of items is measuring and which items are measuring which factors.

The results of the factor analysis are presented in tabular format, usually in a factor pattern matrix. This matrix contains regression coefficients (similar to correlation coefficients) that represent the degree of relationship between the items and the factors. These regression coefficients are referred to as factor pattern loadings. The interpretation of factor pattern loadings is basically one of identifying which items represent which factors. Basically, items that have high factor pattern loadings are good indicators of the construct. Items that have low loadings are poor indicators of the construct.

The results of the correlational and factor analyses are used to select those items that are good observable indicators and thus should be used in the questionnaire. In most cases, some items are dropped from the original list compiled in step 2.

After items are dropped from the trial survey, additional validity evidence is collected and evaluated. Typically, criterion-related evidence is collected and analyzed. This evaluation process helps define what the new survey is assessing.

These three steps have been explained in general terms. The bibliography contains literature that discusses questionnaire development in detail.

How the EEQ was constructed

Following the three-step process just outlined, here is how the EEQ was developed:

1. *Defining the construct.* Since the construct is empowerment, empowerment must be defined. Literature contains many constitutive definitions of empowerment. For example, the article "The Empowerment Process: Integrating Theory and Practice," defines empowerment as the "...process of enhancing feelings of self-efficacy among organizational members through the identification of conditions that foster powerlessness and through their removal by both formal organizational practices

and informal techniques of providing efficacy information."⁹ Another article, "Cognitive Elements of Empowerment: An 'Interpretive' Model of Intrinsic Task Motivation," defines empowerment as increased intrinsic motivation.¹⁰ The Malcolm Baldrige National Quality Award criteria link empowerment to "enhanced employee authority to act...such as when quality standards may be compromised."¹¹

These definitions examine employees' perception of empowerment. In other words, empowerment should be determined by asking employees about their perception of the work environment, their level of self-efficacy, or their perception of authority to act to increase quality.

The definition of empowerment outlined in the Baldrige Award criteria was selected as the constitutive definition, although it must be noted that it overlaps with the other defini-

Figure 1. Factor Pattern Matrix of the Factor Analysis of the EEQ Items (After Rotation)

Item	Factor I	Factor II
1	0.61*	-0.03
2	0.10	0.76*
3	0.63*	-0.01
4	0.78*	0.09
5	0.62*	0.26
6	0.67*	0.00
7	0.67*	0.02
8	0.44	0.18
9	0.72*	-0.15
10	0.53*	-0.06
11	0.26	0.05
12	0.50	0.36
13	0.31	0.93*
14	0.43	-0.21

*Loaded

tions. This definition provides a starting point from which to develop a measure for empowerment.

2. *Generating the items.* Originally, 14 items (declarative statements) were generated for the EEQ. The sample of respondents indicated the degree to which they agreed or disagreed with each of the following statements using a scale of 1 (strongly disagree) to 5 (strongly agree):

1. I am allowed to do almost anything to do a high-quality job.
2. I would like a job that would allow me more authority.
3. I have the authority to correct problems when they occur.
4. I am allowed to be creative when I deal with problems at work.
5. I do not have to go through a lot of red tape to change things.
6. I have a lot of control over how I do my job.
7. I do not need to get management's approval before I handle problems.
8. I have a lot of responsibility in my job.
9. I am encouraged to handle job-related problems by myself.

10. I can make changes on my job whenever I want.
11. I have to follow procedures closely in my job.
12. I have to go through a lot of red tape to get things done around here.
13. I wish management would give me more authority.
14. I can take charge of problems that require immediate attention.

Items No. 2, No. 11, No. 12, and No. 13 are called reverse coded items. To facilitate data interpretation, the scores of negatively phrased items are reversed before analyses are conducted on the EEQ. For example, if a person indicated a score of 5 on item 2, the score would actually be recorded as a score of 1 for the purpose of data analysis; likewise, if a person indicated a score of 4 on item 2, the score would be recorded as a score of 2. A score of 3 would remain the same.

3. *Evaluating the items.* Two studies were conducted to evaluate the items in the EEQ. Multiple studies are needed to ensure that the findings from one sample can be generalized to other samples.

The first study

To examine the quality of the items, a factor analysis was conducted using a heterogeneous sample of employees from five separate organizations. The respondents consisted of 111 full-time employees whose occupations ranged from manufacturing to professional to technical. Completion of the questionnaires was voluntary at each location. The age of respondents ranged from 21 to 53, and 42% of the respondents were male. Ninety-six percent of the respondents reported that they graduated from high school, 57% reported that they received a bachelor's degree, and 22% reported that they held a master's degree.

The results of the factor analysis suggested that the 14 items of the EEQ were measuring two different constructs. Figure 1 contains the factor pattern matrix, which indicates that many of the items are loaded on the first factor and some items are loaded on the second factor.

The items loading on factor I represent empowerment as defined by the Baldrige Award criteria (i.e., the extent to which employees believe that they have the authority to act on their own to increase quality). Some items that loaded highly on factor I were "I am allowed to be creative when I deal with problems at work," "I am encouraged to handle job-related problems by myself," and "I am allowed to do almost anything to do a high-quality job."

The items loading on factor II seem to reflect the extent to which employees desire to have more authority in their present jobs. Due to a limited number of items loading highly on factor II, the focus of the remaining analysis is on the items loading clearly and primarily on factor I. These are items No. 1, No. 3, No. 4, No. 5, No. 6, No. 7, No. 9, and No. 10; they represent the EEQ. The score for the EEQ is calculated by averaging these eight items. The reliability (Cronbach's alpha) for the items was 0.85. Considering the relatively low number of items in the EEQ, the scale had a fairly high reliability estimate.

Relationship of the EEQ scales to other variables

In addition to identifying the extent of empowerment in an organization, the EEQ was also developed to help companies determine how empowerment is related to other variables, such as perceived job characteristics, job satisfaction, job stress, and

intentions to quit. To do this, an expanded employee questionnaire was developed. Questions to measure other organizational variables were added, including:

- Task variety—the extent to which the job allows an employee to work on a variety of tasks
- Task autonomy—the extent to which an employee has a major say in the scheduling of his or her work
- Task identity—the extent to which an employee performs an entire piece of work
- Task importance—the extent to which an employee's job has a major impact on others' jobs (either inside or outside the organization)
- Task feedback—the extent to which an employee receives information on how well he or she is performing the job
- Participation—the extent to which an employee is allowed to participate in decision making
- Organization-based self-esteem—the extent to which an employee believes he or she is a valuable, worthwhile, effective member of the company¹²
- Management's commitment to quality—the extent to which upper management emphasizes quality improvement in its practices (seven-item scale)
- Supervisors' commitment to quality—the extent to which supervisors demonstrate quality improvement practices (eight-item scale)
- Co-workers' commitment to quality—the extent to which co-workers are committed to quality (six-item scale)

To determine job satisfaction, respondents were asked to indicate how satisfying their jobs are using a scale of 1 (very dissatisfying) to 10 (very satisfying). To determine job stress, respondents were asked to indicate the amount of stress in their jobs using a scale of 1 (the least stressful job imaginable) to 10 (the most stressful job imaginable). In addition, respondents were asked to indicate how frequently they thought about leaving their jobs using a scale of 1 (never) to 5 (always) and their company tenure.

Table 1 shows the means, standard deviations, and intercorrelation of the variables. As the table shows, the EEQ is positively related to job satisfaction, organization-based self-esteem, task variety, task importance, and participation. Those respondents who reported higher levels of empowerment also were more satisfied with their jobs, felt more valued, and perceived their jobs as having higher levels of variety, importance, and participation compared to those who reported that they were less empowered in their jobs. In addition, the EEQ was significantly related to supervisors' commitment to quality. Respondents who indicated that their supervisors were committed to quality also reported higher levels of empowerment.

The commitment-to-quality scales showed acceptable psychometric properties. Cronbach's alpha estimates of reliability for each of these scales were 0.82 for management's commitment to quality, 0.83 for supervisors' commitment to quality, and 0.88 for co-workers' commitment to quality. (The scores for each scale were calculated by averaging the items within the respective scales.)

Supervisors' commitment to quality and management's commitment to quality were positively related to the amount of task feedback. Perhaps upper management (and consequently supervisors), through commitment-to-quality actions such as providing employee training, continually looking for ways to improve

quality, and encouraging high-quality work, were able to create work environments in which a great deal of job performance feedback is given. Also, management's and supervisors' commitment to quality were positively related to job satisfaction; that is, employees who reported that management and their immediate supervisors were committed to quality also reported higher levels of job satisfaction than employees who reported that management and their immediate supervisors were not committed to quality.

Only supervisors' commitment to quality was related to empowerment, task importance, and participation. Perhaps empowerment is not greatly fostered through the behavior of management and co-workers. Rather, as the pattern of correlations suggests, empowerment is better fostered through job redesign, which might be one important responsibility of the immediate supervisor. Jobs that require more autonomy, have more task variety, and are perceived as important, then, might lead to feelings of empowerment.

The second study

A new sample of employees was surveyed for the next set of analyses. Administering the survey to a new sample provides a means of determining the generalizability of the first study's results. The process of determining the stability of the results in a new sample is referred to as cross-validation.

For this study, a questionnaire, which included the EEQ, was administered to a sample of 2,000 employees working for a federal government agency. A total of 647 surveys were returned

for a response rate of 32%. Other information was assessed for this study, including supervisory status and company tenure. There were two levels of supervisory status: supervisor (indicated by 17% of the respondents) and nonsupervisor (83%). Company tenure was divided into four categories: zero to 5 years (indicated by 19% of the respondents), 6 to 10 years (33%), 11 to 15 years (19%), and 16 years or more (29%).

A factor analysis of the EEQ portion revealed a clear one-factor solution: All items had factor loading greater than 0.77. In addition, the reliability (internal consistency estimate) was 0.94, suggesting that the items in the EEQ were measuring the same construct.

The results showed that employees who were in supervisory positions tended to have greater company tenure than those who were not. Thus, hierarchical regression analysis was conducted to determine how well the two organization status variables predicted EEQ scores. Hierarchical regression analysis determines the relative strength of each variable in predicting EEQ scores. The results indicated that only supervisor status significantly predicted EEQ scores. Supervisors had higher EEQ scores (mean = 3.71) than nonsupervisors (mean = 3.04). Company tenure did not add significantly to the prediction of EEQ scores.

Empowerment can be measured

In the past, empowerment has been discussed at great length without regard to measurement. But if quality professionals want to advance the understanding of empowerment, they should take time to develop and use measures that will reliably

Table 1. Means, Standard Deviations, and Intercorrelations of the Study Variables

Variables	Mean	Standard deviation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Empowerment	3.43	0.62	(0.85)														
2. Organization-based self-esteem	3.90	0.43	0.49**	(0.82)													
3. Task variety	3.52	0.75	0.35*	0.11	(0.76)												
4. Task feedback	3.50	0.73	0.26	0.35**	0.22	(0.81)											
5. Task identity	3.34	0.95	0.16	0.14	0.37**	0.24	(0.81)										
6. Task importance	3.72	0.77	0.30*	0.60**	0.19	0.44**	0.03	(0.72)									
7. Participation	3.14	0.79	0.60**	0.48**	0.31*	0.27*	0.41**	0.28*	(0.74)								
8. Task autonomy	3.42	0.61	0.42**	0.24	0.27	0.16	0.34**	0.16	0.41**	(0.59)							
9. Management's commitment to quality	3.57	0.61	0.05	0.27*	-0.10	0.40**	0.14	0.26	0.02	0.11	(0.82)						
10. Supervisors' commitment to quality	3.65	0.53	0.21*	0.44**	0.01	0.50**	0.19	0.27*	0.27*	0.14	0.78**	(0.83)					
11. Co-workers' commitment to quality	3.63	0.58	0.09	0.08	0.10	0.17	0.21	0.24	0.11	0.11	0.40**	0.20*	(0.88)				
12. Job satisfaction	7.00	1.62	0.34**	0.38**	0.43**	0.49**	0.52**	0.33*	0.34**	0.24	0.34*	0.33**	0.13	-			
13. Job stress	5.92	1.84	-0.02	-0.01	0.26	-0.10	0.10	0.13	-0.04	-0.08	-0.06	-0.09	0.02	-0.11	-		
14. Intent to quit	2.65	1.11	-0.11	-0.22	-0.02	-0.36**	-0.18	-0.19	-0.21	-0.07	-0.37**	-0.31*	-0.15	-0.64**	0.19	-	
15. Company tenure	3.13	3.89	0.18	-0.05	0.13	-0.27	-0.09	0.11	0.28	0.15	-0.41**	-0.33*	-0.07	-0.04	0.08	0.16	-

Note: N = 51 to 111. Cronbach's alpha (reliability estimate) is located in the diagonal.

*p < 0.05

**p < 0.01

