Assessing Team Effectiveness

P.K. Imbrie, Susan J. Maller, and Jason C. Immekus
Purdue University, West Lafayette, Indiana

Abstract

The continuation of the technology explosion of the second half of the 20th century requires the availability of a diverse and highly capable technical workforce. Current teaching pedagogies rely heavily on students collaborating, either informally or formally, in a team-like environment. Unfortunately, even with the increased emphasis on the use of student teams in academia there has been little-to-no effort to develop quantitative instruments to measure how successful the teaming experience is for participating students. Therefore, the goal of this study was to begin developing a self-assessment instrument, which would have evidence of reliability and validity, to facilitate the identification of effective student teams. For the purpose of the paper, students' perception of functionality/effectiveness has been operationalized in terms of a self-report, 24-item instrument requiring students to indicate the degree to which their team worked together across a range of domains, including: interdependency, learning, potency and goal-setting.

Although the instrument was conceptualized as a four factor model, results of this study indicate the current scale represents a single-factor model of team effectiveness. Subscale internal consistency reliability estimates, based on Cronbach's coefficient alpha, were: .96, .92, .96, and .94 for Interdependency, Learning, Potency, and Goal-Setting, respectively. Cronbach's coefficient alpha for the entire scale was .98.

Background

The NAE National Research Council Board on Engineering Education, NSF Engineering Education Coalition Program, and the Accreditation Board for Engineering and Technology [1] Engineering Criteria 2000 ushered in a movement to reshape the engineering curricula. To build on these pioneering initiatives, new educational pedagogies must be used to develop graduates as successful professional contributors and lifelong learners in global, multi-disciplinary markets; be flexible to support diverse career aspirations; be agile to rapidly transform in response to emerging social demands; and have a profound understanding of the importance of teamwork [2-3]. In response to this, education in engineering has seen a significant increase in the emphasis on design and on the wide range of teamwork skills that engineering students will need when they enter the workplace [4-8]. In the program outcomes, at the heart of Engineering Criteria 2000 accreditation guidelines, students are mandated to be able to function on multidisciplinary teams in addition to acquiring traditional engineering knowledge of mathematics, science, and engineering and gaining experience in engineering problem solving and system design [1].

Unfortunately, even with the increased emphasis on the use of student teams in academia there have been very few research studies that rigorously attempt to assess team effectiveness. Because this is an essential element of the overall collaborative experience, the work presented herein will endeavor to operationalize a definition of effective teaming, thus facilitating a

Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition

student's ability to readily identify attributes of good team performance. In addition, for the purpose of this paper, we shall define a student team using the work of Guzzo [9]: *A team is a group that consists of individuals who see themselves and are seen by others as a social entity, which is interdependent because of the tasks performed as members of a group. They are part of the educational process, performing tasks that affect both individual and group learning.* The key to student teams is that they are interdependent, and this is the major factor that distinguishes a "team" from a "group."

Although the focus of measuring team effectiveness has centered on "industry" type of teams, the pioneering work of a number of researchers will serve as the theoretical framework for the proposed research. Starting in the 1940s, work groups became a new focus of attention after the Hawthorne studies were published. In recent years the use of work teams in organizations has been increasing substantially, and this trend is expected to continue [10]. It is indicated that eighty percent of organizations with over 100 employees report 50% of their staff are on at least one team (Beyerlein & Harris, 1998). To remain competitive, it is important for organizations to create and maintain teams which are as effective as possible. ABET Criterion 3 reminds accredited engineering programs that working in multi-disciplinary teams is critical to pursuing an engineering career.

According to Campion, Medsker, and Higgs [11], team effectiveness can be defined in terms of productivity, employee and customer satisfaction, and manager judgments. Within this model, job design, interdependence, composition, context and potency (i.e., does the team think it can be successful) are attributes of effective teams. Campion et al. [11] have tested the above model in an empirical study and found that almost all of the characteristics of work groups (e.g., interdependence, potency) related to the three criteria of effectiveness (e.g., customer satisfaction). Potency was found to be the strongest predictor of all characteristics and related to all three effectiveness criteria, thus supporting those who assert that it is one of the most important characteristics of a work team [12].

Campion et al.'s [11] ideas can be traced to Richard Hackman [13], who defined effectiveness in terms of three dimensions: the group's output meeting quality standards, the group's ability to work interdependently in the future, and the growth and well-being of team members. Another model of effectiveness is presented by Guzzo [9], who conceptualizes it in the same general terms as Hackman. In his view, effectiveness is defined by measurable group produced outputs, consequences the group has for members, and the capability to perform well in the future. Collectively, these models provide a rich foundation for conceptualizing effective teams in academic settings.

Theoretical Model

Figure 1 shows the hypothesized four-factor model of the teamwork scale investigated in this study. As shown, the relationships among the 24 items are posited to be explained by the following four latent traits: *Interdependency*, *Learning*, *Potency*, and *Goal-Setting*. Within the figure, the observe variables (items) are represented by squares, and the unobserved, latent traits are represented by circles. The arrows connecting the observed and latent variables are factor loadings, or pattern coefficients, that refer to the strength of the relationship between the variables. The arrows corresponding to each observed variable indicate error variance, or the unique variance unaccounted for by the latent trait. The arrows connecting the latent traits indicate they are correlated.

Methods

Participants

The sample included 1,060 freshmen engineering students (190 females, 870 males) at a large Midwestern university. Students' ethnicity was as follows: 22 African Americans, 6 American Native, 86 Asian/Pacific Islanders, 27 Hispanics, 786 Whites, 17 others, 105 Internationals and 11 missing.

Students completed the scale following participation on student teams in a freshmen engineering course during the Fall 2004 academic semester. The scale was designed for the purposes of identifying whether students' perceived their team as effective. Item responses were recorded on a Likert scale (1=strongly disagree to 5=strongly agree).

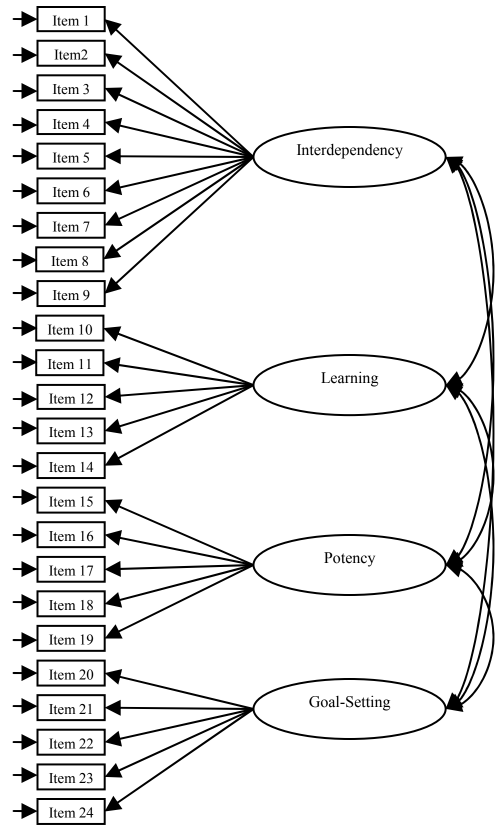

Figure 1. Hypothesized Four-Factor Model of the Teamwork Scale

Data Analysis

The team effectiveness measure's four-factor model was testing using methods of confirmatory factor analysis (CFA). CFA represents a theoretically-based approach to test an instrument's factor structure. CFA is based on *a priori* information specifying the relationships between observed and latent variables [14-15]. The ability of CFA to empirically test the correspondence between test scores and latent variables makes it particularly useful to document evidence of construct validity [16]. The degree to which the measurement model represents observed data is evaluated through various fit statistics. Based on findings, models sometimes are re-specified [14]. After evaluating the four-factor model, a one factor model, in which all items indicated an overall team effectiveness latent variable, subsequently was tested. Item means and reliability analyses also were reported.

LISREL 8.53 [17] was used to conduct the CFAs. Specifically, weighted least squares based on a polychoric matrix and corresponding asymptotic covariance matrix was used for parameter estimation. This procedure has been identified as producing stable parameter estimates with non-normal, ordinal data [18]. Various fit statistics included: chi-square statistic (χ^2), ratio of chi-square to degrees of freedom (χ^2/df), Root Mean Error of Approximations (*RMSEA*), Goodness of Fit (*GFI*), and Comparative Fit Index (*CFI*). The chi-square to degrees of freedom ratio suggests how much larger chi-square is than would be expected [14], with values less than 3.00 indicating good fit. The *RMSEA* provides a measure of the discrepancy between the actual and estimated variance-covariance matrix per degree of freedom [19-20], with values equal to or less than .05 indicating good model fit and values less then .08 indicating reasonable fit [19]. The *GFI* indicates the amount of variance and covariance in the sample covariance matrix accounted for by the implied model, and ranges between 0-1 with values exceeding .95 indicating acceptable fit. The *CFI* provides a measure of the discrepancy between a restricted and null model in relation to the fit of the null model [21-22]. Its values range between 0-1 with values above .95 suggesting adequate fit [23].

Results

Item Descriptive Statistics and Reliability

Table 1 shows the means and standard deviations of the instrument's 24 items. As shown, students generally agreed to all items. Subscale internal consistency reliability estimates, based on Cronbach's coefficient alpha, were: .96, .92, .96, and .94 for Interdependency, Learning, Potency, and Goal-Setting, respectively. Cronbach's coefficient alpha for the entire scale was .98. These estimates exceeded the desired criteria of .90 [24].

Table 1 *Means and Standard Deviations of Teamwork items.*

Subscale	Item	*M*	*SD*
Interdependency	1	4.00	1.16
	2	4.02	1.08
	3	3.73	1.23
	4	3.69	1.14
	5	3.82	1.10
	6	3.78	1.12
	7	3.71	1.08
	8	3.81	1.22
	9	3.98	1.12
Learning	10	3.96	1.20
	11	3.76	1.18
	12	3.58	1.08
	13	3.83	1.08
	14	3.77	1.14
Potency	15	4.01	1.12
	16	3.85	1.35
	17	3.86	1.15
	18	3.92	1.20
	19	3.96	1.10
Goal-Setting	20	3.82	1.14
	21	3.69	1.12
	22	3.81	1.08
	23	3.97	1.04
	24	3.85	1.17

Note. *N*=1,060. *M*=Mean, *SD*=Standard Deviation.

Confirmatory Factor Analysis

The four-factor model fit the data (χ^2[248]=312.24, *p*=.004, χ^2/df =1.26, *RMSEA*=.02, *CFI*=1.00, *GFI*=1.00). Based on the large sample size, it was expected that the model chi-square would be statistically significant, and other fit indices, including the χ^2/df, indicated acceptable fit. Pattern coefficients are regression coefficients that describe the linear relationship between the observed and latent variables, and ranged between .98 (Item 2) and 1.00 (Item 14). The correlations among the factors were 1.00. These strong, positive correlations between the factors indicated that the factors may not be indistinguishable. Therefore, a single-factor CFA model was investigated.

The single-factor model fit the data (χ^2[254]=316.15, *p*=.005, χ^2/df =1.24, *RMSEA*=.02, *CFI*=1.00, *GFI*=1.00). The chi-square is affected by large sample sizes, including the χ^2/df, and other fit indices indicated acceptable fit. Based on the parsimony of the single-factor model and little evidence suggesting that the scale is a multi-factor instrument, the single-factor model was accepted as the final measurement model. Table 2 provides the pattern coefficients, error variances, and explained variance for each item. Explained variance, R^2, indicates the amount of variance of the items explained by the underlying latent trait. As shown, the underlying factor accounted for over 80% of the variance across the observed variables.

Table 2 *Factor Pattern, Error Variance and Explained Variance Estimates for Four-Factor Model*

Item	Pattern	Error Variance	R^2
1[a]	1.00	.00	1.00
2	.98	.04	.96
3	.98	.05	.95
4	.97	.07	.93
5	.99	.01	.99
6	.99	.01	.98
7	.96	.02	.93
8	1.00	.07	1.00
9	1.00	.00	1.00
10	.98	.04	.96
11	.99	.02	.98
12	.90	.19	.81
13	1.00	.00	1.00
14	.99	.01	.99
15	1.00	.00	1.00
16	.99	.01	.99
17	1.00	.00	1.00
18	1.00	.01	.99
19	1.00	.01	.99
20	.99	.02	.98
21	.98	.04	.96
22	.99	.01	.99
23	1.00	.00	1.00
24	.99	.01	.99

Note. Completely standardized regression coefficients. All pattern coefficients were statistically significant ($p<.05$).

[a] Parameter set to 1.0.

Discussion and Conclusion

The purpose of this study was to test the factor structure of a teamwork effectiveness scale that was conceptualized based on theoretical considerations. Subscales represented key domains identified as encouraging effective teams [11]. The instrument is intended for identifying effective teams in engineering education settings.

Results indicated scores show impressive reliability evidence. Furthermore, a single-factor model best represented the measure of team effectiveness. Pattern coefficients indicated observed variables were strongly related to the underlying construct (latent trait) of team effectiveness, with small error variances. A large proportion of the variance in each item was explained by the latent trait. Therefore, the one factor model confirms that items can be summed to create a composite score, thus operationalizing a definition of effective teaming that is based on a measure with construct validity evidence.

BIBLIOGRAPHIC INFORMATION

1. ABET (2002). *Engineering Criteria 2002-2003*. Accreditation Board for Engineering and Technology, http://www.abet.org/criteria.html.
2. National Science Foundation Division of Undergraduate Education (1995), *Restructuring Engineering Education: A Focus on Change*, Arlington, VA..
3. McMaster, J. (2000). *Boeing List of "Desired Attributes of an Engineer,"* Industry-University-Government Roundtable for Enhancing Engineering Education.
4. ASEE Deans' Council and Corporate Roundtable (1994), *Engineering Education for a Changing World*, Washington, DC: American Society for Engineering Education, October.
5. ASEE (1995). Educating Tomorrow's Engineers. *ASEE Prism*, 11-15, May/June.
6. Dahir, M. (1993). Educating Engineers for the Real World. *Technology Review*, 14-16, Aug/Sept.
7. Hissey, T.W. (2000). Education and Careers 2000. *Proceedings of the IEEE, 88*(8), 1367-1370.
8. Valenti, M. (1996). Teaching Tomorrow's Engineers. *Mechanical Engineering Magazine, 118*(7).
9. Guzzo, R. A. (1986). Group Decision Making and Group Effectiveness. In Goodman, P. S. (Ed.). Designing Effective Work Groups, 34-71. San Francisco, CA: Jossey-Bass.
10. Katzenbach, Jon R. and Douglas K. Smith (1992) *Wisdom of Teams*, Harvard Business School Press.
11. Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology,* 46, 823-850.
12. Guzzo, R. A., Campbell, R. J. & Shea, G. P. (1993). "Potency in groups: Articulating a construct." British Journal of Social Psychology,31,87-106.
13. Hackman, J. R. (1990) Groups that work (and those that don't). San Francisco, CA: Jossey-Bass.
14. Bollen, K. A. (1989). *Structural equations with latent variables*. New York, New York: John Wiley & Sons
15. Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
16. Keith, T. Z. (1997). "Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests." In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 373-403). New York: Guilford.
17. Jöreskog, K., & D. Sörbom (2002). *LISREL 8.53* [Computer Program]. Lincolnwood, IL: Scientific Software International.
18. Flora, D. B., P. J. Curran (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*, 466-491.
19. Brown, M. W., & R. Cudeck (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.): *Testing Structural Equation Models.* Sage Publications.
20. Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling, 3*, 369-379.
21. Bentler, P. M. (1990). "Comparative fix indexes in structural models." *Psychological Bulletin, 107*, 238-246.
22. Rigdon, E. E. (1996). "CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling." *Structural Equation Modeling, 3*, 369-379.
23. Hu, L., & P. M. Benter (1999). "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives." *Structural Equation Modeling, 6*, 1-55.
24. Nunnally, J. (1978). *Psychometric Theory* (2^{nd} edition). New York: McGraw-Hill.

BIOGRAPHICAL INFORMATION

P.K. IMBRIE is an Assistant Professor in the Department of Engineering Education at Purdue University and is experienced in the design and development of new and innovative educational pedagogies. He was involved in the NSF funded Foundation Coalition as part of the local Management Team, while at Texas A&M University. He participated in the pilot programs for restructuring the freshman year by emphasizing content integration, an active/collaborative learning environment, efficient technology utilization and continuous change based on assessment and evaluation.

SUSAN J. MALLER is an Associate Professor of Educational Psychology (Measurement and Research Methodology) at Purdue University (1999-present), and is Co-Director of the Purdue University Psychometric Laboratory. She is Associate Editor of *Psychology in the Schools*, and is on the editorial boards of *Educational and Psychological Measurement, Journal of School Psychology* and *Journal of Psychoeducational Assessment*. She has received funding from The Spencer Foundation, NSF, and the Indiana Department of Education. Her publications concern the development and psychometric properties of the educational and psychological tests.

JASON C. IMMEKUS is an advanced doctoral student in the Educational Psychology program at Purdue University. His areas of specialization are applied measurement and research design. His research interests include test validity, structural equation modeling, and test development.