IJQRM
27,8

938

Received October 2008
Revised April 2009
Accepted May 2009

Validation of reliability capability evaluation model using a quantitative assessment process

Sanjay Tiku
Microsoft Corporation, Sammamish, Washington, USA, and
Michael Pecht
CALCE Electronic Products and Systems Center, Engineering Laboratory Building, University of Maryland, College Park, Maryland, USA

Abstract

Purpose – Reliability capability is a measure of the practices within an organization that contributes to the reliability of the final product and the effectiveness of these practices in meeting the reliability requirements of customers. The purpose of this paper is to propose a model for evaluating the reliability capability of electronics manufacturers.

Design/methodology/approach – A survey methodology and statistical methods based on multivariate correlational analysis were used to validate the model theoretically.

Findings – The result of the analysis is a list and ranking of tasks that are critical to the development of reliable electronics products.

Originality/value – The paper presents a generic model for evaluating both in-house reliability practices and those of suppliers to identify areas for improvement and for evaluating improvements over time.

Keywords Product liability, Supplier relations, Benchmarking, Electronic equipment and components

Paper type Research paper

1. Introduction

The maturity approach to determine organizational abilities has roots in quality management. Crosby (1996) described the typical behavior of a company at five levels of quality maturity using the quality management maturity grid. Since then, maturity models have been proposed for a wide range of activities including software development (Paulk *et al.*, 1993), supplier relationships (Macbeth and Fergusson, 1994), research and development effectiveness (Szakonyi, 1994a, b), product development (McGrath, 1996), innovation (Chiesa *et al.*, 1996), product development collaboration (Fraser and Gregory, 2002), product design (Fraser *et al.*, 2001), sub-sea reliability management (Strutt, 2001) and reliability capability of electronics companies (Tiku and Pecht, 2003a, b; Tiku *et al.*, 2007).

Maturity models for organizational abilities must have empirical validation. In management and marketing research, even though a relatively large number of abstract theoretical variables are used to explore the relationship among different organizational phenomenon, it has been reported that a serious shortcoming of most of these theoretical measuring instruments is that they lack validation (Jacoby, 1978; Schriesheim *et al.*, 1993).

International Journal of Quality & Reliability Management
Vol. 27 No. 8, 2010
pp. 938-952

0265-671X
DOI 10.1108/02656711011075125

Quantitative techniques, called psychometric methods, are available for validating lists of measurement items. Psychometric methods are rigorous statistical tools that are used to construct theoretical instruments which measure abstract organizational variables (Benson *et al.*, 1991; Churchill, 1979; Kuei *et al.*, 2001; Du Plessis, 2003; Saraph *et al.*, 1989; Zhang, 2001). These methods, which are based on statistical multivariate co-relational analysis, can also be used to validate the theoretical measurement model proposed for reliability capability. Figure 1 compares the steps in physical experimental research process and the empirical psychometric research process (Bryman and Cramer, 2001). In the former case, the test vehicle is a physical specimen; in the latter, the test vehicle is a survey questionnaire. In the former, the test results constitute the output data; in the latter, the scores or ratings from respondents constitute the output data.

The fundamental objective of any measuring instrument is to produce observable scores that approximate the true scores. The process of measurement involves rules for assigning numbers to objects to represent quantities of attributes (Nunnally, 1978). However, the attributes of objects as opposed to the objects themselves are measured. The measures are inferences, and the quality of the inferences depends on the procedures that are used to develop the measures and the evidence supporting the "goodness" of these measures (Churchill, 1979). The "goodness" is typically specified using the indices for internal consistency and validity. Figure 2 (Churchill, 1979; Hinkin, 1995; Saraph *et al.*, 1989) shows the process steps in the development and validation of the reliability capability model.

The first step in generating measurement items is exploratory research including the literature research and feedback from experienced professional (Churchill, 1979). In a previous study, the authors identified eight key reliability practices for reliability capability evaluation (Tiku and Pecht, 2003a). An evaluation questionnaire was then created, and as a pre-test, reliability audits were conducted for three electronics companies (Tiku and Pecht, 2003b). Based on these activities (the first three steps of the

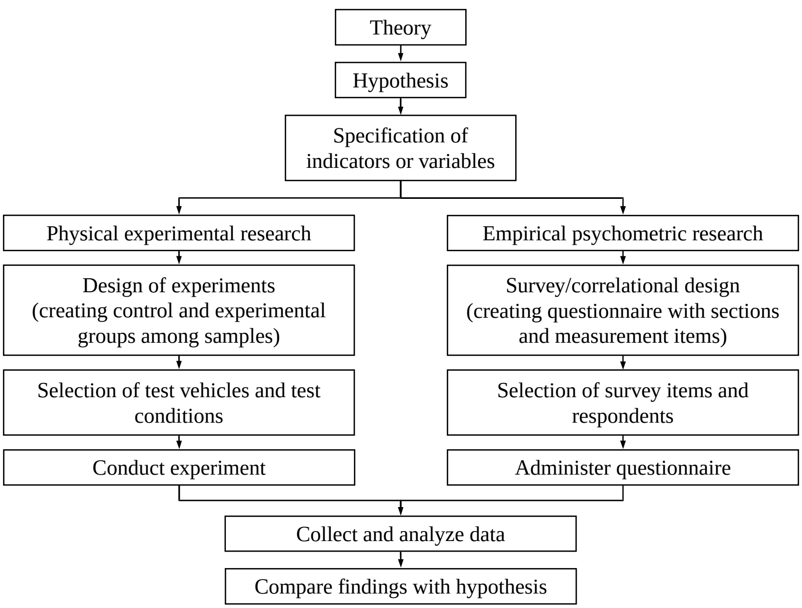

Figure 1.
Difference between physical experimental research and empirical psychometric research

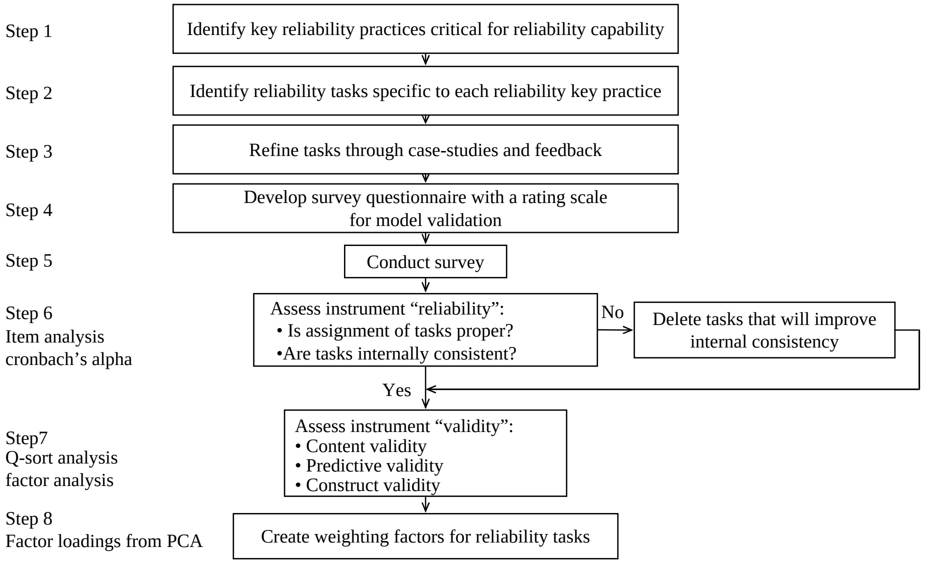

Figure 2.
Reliability capability model development and validation process

development process), 91 reliability tasks (Appendix, Table AI) were identified for measuring reliability capability.

In this study, a survey questionnaire, containing 91 reliability tasks, was created as a scientific instrument (Berdie and Anderson, 1974). The statement of each task was reviewed by researchers and reliability professionals to make them concise and unambiguous. In the survey, the respondents were required to grade each task on a five-point interval rating scale ("negligible" to "very high") in terms of the relevance of the task in ensuring or improving the reliability of an electronics product. The respondents were also asked to grade the collective relevance of each practice on the same rating scale.

The respondents to the survey questionnaire were chosen such that they would represent those who will eventually use or interpret the results of the instrument (Nunnally, 1978). In all, 211 responses were obtained from industry professionals, consultants and researchers associated with electronic reliability. These people also represent organizations of various sizes. The obtained data were analyzed using the Statistical Package for Social Sciences version 13.0 (Antonius, 2003; Bryman and Cramer, 2001; Foster, 1998; Kerr *et al.*, 2002) to evaluate the internal consistency and validity indices in Steps 6 and 7 and weighting factors in Step 8.

2. Assessing internal consistency

Internal consistency (also called "reliability" in psychometric parlance) refers to the stability or reproducibility of a test score based on a theoretical instrument (Nunnally, 1978). A measure is considered internally consistent if it will give the same results when the measurement is repeated. However, internal consistency is only a necessary and not a sufficient condition for validity.

Item analysis was first used to evaluate the appropriateness of the assignment of tasks to key practices (Nunnally, 1978; Saraph *et al.*, 1989), by considering the correlation of each task rating to the average rating for each key practice. A task is eliminated if it correlates more with some other key practice than the one to which it is assigned. The analysis was completed for all the 91 tasks. Tasks 1-04 and 2-10 showed close correlations with two key practice scores. However, they have the maximum correlations with their assigned practices. On the other hand, 8-11 shows better correlation with TAD (0.61), compared with RIMP (0.59), and hence was excluded from further analysis.

Within each key practice, one of the most commonly used coefficients that can be used for measuring internal consistency of a list of tasks is Cronbach's alpha (Cronbach, 1951; Nunnally, 1978; Saraph *et al.*, 1989). Mathematically, Cronbach's alpha is the average of correlations between all possible split-half estimates within a key practice. The value of Cronbach's alpha for a key practice containing "k" reliability tasks is given by (Cronbach, 1951, 2004):

$$\alpha = \left[\frac{k}{k-1}\right]\left[1 - \frac{\sum s_i^2}{s_{sum}^2}\right] \tag{3.1}$$

where s_i^2 is variance of each task rating and s_{sum}^2 is the variance of the average key practice rating. For each key practices, the square root of the coefficient α value is the correlation between the score that companies will get on the selected tasks (sample score) to the score that companies would get if all possible tasks corresponding to the key practice are included for evaluation (true score). Typically, an α value of 0.7 or more is considered adequate (Nunnally, 1978).

Table I shows the Cronbach's alpha values determined for eight different key reliability practices for this study. All the key practices have more than adequate α values, and deletion of any one task does not substantially improve the α value for the key practice. Hence, the tasks listed under different key practices used for measuring reliability capability demonstrate internal consistency in psychometric terms. In other words, coefficient α values show that the specified tasks are necessary and sufficient to describe each key practice.

For the entire reliability capability measuring instrument, which is a linear combination of measures of different key practices, the internal consistency can be estimated by the knowledge of Cronbach's alpha coefficients of key practices and the covariance among their average ratings. The internal consistency co-efficient for this linear combination is given by (Nunnally, 1978):

$$r_{RCM} = 1 - \frac{\sum \sigma_i^2 - \sum \alpha_i \sigma_i^2}{\sigma_y^2} \tag{3.2}$$

Key practice	Symbol	Number of tasks	"α" value
Reliability requirements and planning	RRP	12	0.779
Training and development	TAD	10	0.827
Reliability analysis	RA	11	0.838
Reliability testing	RTST	13	0.851
Supply chain management	SCM	15	0.897
Failure data tracking and analysis	FDTA	11	0.899
Verification and validation	VAV	08	0.871
Reliability improvements	RIMP	11	0.856

Table I. Cronbach's alpha values for different key practices where σ_i^2 is the variance in the rating for the ith key practice, α_i is the value of the coefficient α for the ith key practice and σ_y^2 is the sum of all elements in the covariance matrix of average key practice ratings. Using the values of Cronbach's alpha values for each key practice, and the values of correlation between average ratings, the internal consistency coefficient of the entire reliability capability model was found to be 0.972. This value indicates that the key reliability practices and included tasks are significantly necessary and sufficient to evaluate reliability capability.

3. Assessing validity

The validity of a measure refers to an extent to which it measures what it is intended to measure (Nunnally, 1978). It is also the extent to which the differences in scores based on the instrument reflect the true differences among organizations on the characteristic that the instrument is supposed to measure and nothing else (Churchill, 1979). Thus, a measurement instrument is valid when the observed score matches the true score and the variation due to any systematic or random errors is very low.

Validity of a measuring instrument is of three types – content or face validity, criterion-related or predictive validity and construct validity. The reliability capability measuring instrument will have predictive validity if the evaluation scores for different companies are correlated with the actual reliability of their products (Saraph *et al.*, 1989). This requires correlating the field reliability of products supplied by the company to the maturity score obtained from an evaluation. Unfortunately, these data are extremely difficult to obtain. As an alternative, we rely on content validity and construct validity instead. These are discussed below. As per Nunnally (1978):

> Even though a test that is used specifically for a prediction function should be validated as such, the only recourse is to rely heavily on content validity and construct validity instead. The reason is that in many cases a test must be selected for use before there is an opportunity to perform studies in which it is correlated with a criterion. In many performance situations, the criterion measure might not be available for years, or the ones that are available are obviously biased in one way or the other or are highly unreliable.

3.1 Content validity

A measuring instrument has content validity if the measurement items cover all the aspects of the variable (domain content) being measured. Content validity exists when "a measure is judged by one or more persons as containing a reasonable and representative sample of items from the construct's theoretical domain" (Schriesheim *et al.*, 1993).

In our case, the reliability capability model has some degree of content validity because it was constructed based on the literature and standards on the topic (AIAA, 2004; Bell Communications Research, 1993; DoD, 1980; IEC TC-56, 2001; IEEE SA Standards Board, 1998; SAE, 1998; Williams *et al.*, 2003) and evaluation by academicians and practicing reliability managers from the electronics industry. Although content validity is subjectively judged by researchers and not usually quantitatively measured, a quantitative approach to the assessment of content validity, called the Q-sorts methodology, was also used to establish content validity (Schriesheim *et al.*, 1993).

Q-sorts technique is a method of sorting objects into theoretical categories for statistical purposes (Nunnally, 1978). The first step in this approach is that a panel of judges rates the questionnaire tasks for specific scales (key practices), employing a pre-established theoretical definition for each scale (the purpose of key practices in our study). The method requires judges to classify tasks into categories whose definitions or purposes are provided. Undergraduate or graduate students are appropriate to be used as the panel of judges. According to Schriesheim *et al.* (1993):

> [. . .] the only requirement for a set of judges to be considered adequate for this task is that they possess sufficient intellectual ability to perform the item rating task and that they be relatively free of serious potential biases.

For using this method, a content validity questionnaire was created. The questionnaire random listed the 91 reliability tasks and required the judges to classify them into the eight key reliability practices. The judges were provided with a brief definition of the purpose of each key practice to do the classification. In our case, there were 56 responses to the questionnaire from researchers in electronics, general engineering graduates and nonengineering graduate students.

The 56 responses were classified randomly into two segments (S1 and S2) selecting half from each of the groups above. Data were compiled for the number of times each task was classified under different key practices for each segment, and correlation coefficients were obtained for the two segments. The correlation values between data for the two segments are shown in Table II. The results show good correlation between the key practice classifications for the two segments, at significance levels much lower then 0.01 per cent demonstrating content validity.

3.2 Construct validity

A measuring instrument has construct validity if it measures the trait (theoretical construct) that it was designed to measure (Churchill, 1979). Construct validity of each key practice can be evaluated by using factor analysis. Factor analysis validates a scale (key practice) by demonstrating that its constituents (reliability tasks) load on the same common factor. If all the tasks listed under a key practice load on a single factor, they measure the same trait. Factor analysis and construct validity have long been associated with each other, and construct validity is also sometimes called "factorial validity" (Cronbach and Meehl, 1955; Thompson and Daniel, 1996; Tucker and MacCallum, 2006).

For the analysis of data obtained through the survey, each key practice is treated as a separate measure of an organizational trait. Two factor analyses methods, the principal component analysis (PCA) and principal axis factoring (PAF) methods,

	S2_RRP	S2_TAD	S2_RA	S2_RTST	S2_SCM	S2_FDTA	S2_VAV	S2_RIMP
S1_RRP	*0.944**	−0.066	−0.053	−0.144	−0.151	−0.346	−0.121	−0.139
S1_TAD	−0.094	*0.986**	−0.259	−0.160	−0.161	−0.225	−0.170	−0.161
S1_RA	−0.174	−0.245	*0.937**	−0.007	−0.278	0.098	−0.007	−0.134
S1_RTST	−0.121	−0.185	−0.042	*0.949**	−0.176	−0.179	0.050	−0.160
S1_SCM	−0.111	−0.188	−0.259	−0.192	*0.979**	−0.192	−0.122	−0.185
S1_FDTA	−0.344	−0.221	0.132	−0.169	−0.217	*0.976**	−0.076	−0.017
S1_VAV	−0.154	−0.188	−0.118	0.011	−0.081	0.010	*0.921**	0.023
S1_RIMP	−0.125	−0.135	−0.204	−0.126	−0.228	−0.028	0.051	*0.936**

Notes: *Correlation is significant at the 0.01 level (two-tailed); the italicized values represent the outcome of the analysis and the most relevant numbers

Table II. Q-sorts methodology correlational results were used for this verification, since there is enough evidence to suggest that nearly all factoring methods should provide the same results, if there are really clear groupings of variables in a correlational matrix (Nunnally, 1978). Using the Catell's scree test criterion (Foster, 1998; Kline, 1994), it was found that only one factor should be extracted for each key practice for both types of analyses (PCA and PAF). Task 8-11 was excluded from this analysis.

The outputs from factor analysis are the factor loadings for each measurement task. The factor loadings are the correlation coefficients between variables or measurement tasks and the identified factors. Results were similar for both types of factor analysis techniques. Summary of the results from PAF (Table III) shows that tasks 1-04 and 4-06 should be deleted since they do not have factor loadings of more than the recommended significant value of 0.3 with their respective factors (Bryman and Cramer, 2001; Foster, 1998; Kline, 1994; Nunnally, 1978). After eliminating tasks 1-04 and 4-06, co-efficient α values were re-calculated for RRP and RTST and were found to be 0.784 and 0.857, respectively.

4. Weighting factors for reliability tasks

The validation process resulted in a list of 88 tasks that can be used for reliability capability evaluation. The validation process, however, does not provide any information on the relative importance of these tasks for each key practice. This importance can be expressed in the form of weighting factors that can be assigned to reliability tasks during an evaluation.

Through factor analysis, it was found that each key practice represents a single factor or organizational trait. Since factors represent linear combination of variables that load significantly on it, each key practice can be written as a linear combination of tasks that load significantly on it:

$$A = w_1a_1 + w_2a_2 + \ldots + w_ka_k \tag{4.1}$$

where A is the score on a key practice, a_i is the scores on individual tasks and w_i is the weighting factor assigned to the ith task. The factor loadings for tasks under different key practices obtained from PCA can be used as weighting factors in the above equation. For each key practice, the factor loadings were scaled such that the minimum weighting factor for any task became 1. The factor loadings and weighting factors for each reliability task are included in Appendix, Table AI. Table IV provides the range of weighting factor values and the sum of weighting factors for tasks under each key practice.

Key practice	Range of factor loadings	Tasks with loading < 0.3
RRP	0.238-0.640	1-04 (0.238)
TAD	0.367-0.698	None
RA	0.464-0.652	None
RTST	0.284-0.753	4-06 (0.284)
SCM	0.377-0.715	None
FDTA	0.560-0.763	None
VAV	0.393-0.778	None
RIMP	0.495-0.715	None

Table III.
Summary of results from PAF

Key practice	Number of tasks	Range of weighting factors	Sum of weighting factors
RRP	11	1.00-1.88	*17.00*
TAD	10	1.00-1.70	*14.57*
RA	11	1.00-1.32	*13.17*
RTST	12	1.00-1.66	*16.06*
SCM	15	1.00-1.78	*23.18*
FDTA	11	1.00-1.29	*12.79*
VAV	8	1.00-1.74	*12.68*
RIMP	10	1.00-1.35	*12.07*
	88		*121.52*

Notes: *Correlation is significant at the 0.01 level (two-tailed); the italicized values represent the outcome of the analysis and the most relevant numbers

Table IV. Weighting factors for tasks under different key practices

The sum of the weighting factors for all tasks indicates the maximum score that a company can obtain from an evaluation to be regarded as following best-in-class reliability practices. Based on these weighting factors, a company can be assigned scores from an evaluation. The obtained scores can be used to build a bar chart and a radar chart as graphical illustrations of evaluation results and for comparative analysis among companies.

5. Summary and conclusions

This paper uses the statistical methods suggested in the field of psychometrics for validating the key reliability practices and associated reliability tasks proposed earlier for the reliability capability evaluation model. A survey questionnaire was used to obtain relevance ratings for reliability tasks divided among eight key reliability practices. Item analysis, Cronbach's alpha calculations, Q-sort method and factor analysis were used to demonstrate the internal consistency and validity (content and construct) of the key practices and associated tasks. Factor loadings obtained from factor analysis results were subsequently used to develop weighting factors for reliability tasks useful for a quantitative assessment.

Item analysis resulted in elimination of one task (8-11) since it was found to correlate better with a different key practice than the one to which it was assigned. Cronbach alpha co-efficient values were found to exceed the recommend value of 0.7 for each key practice. The internal consistency coefficient (also called "reliability co-efficient" in psychometric parlance) for the entire reliability capability measuring instrument was found to be 0.972. This value indicates that the key reliability practices and included tasks are significantly necessary and sufficient to evaluate reliability capability.

Content validity of the measuring instrument was demonstrated using the Q-sort method. Factor analysis was used for demonstrating construct validity. Two tasks (1-04 and 4-06) were found to have factor loadings less than the recommended lower limit of 0.3 and were deleted from the model. Weighting factors were then obtained for the remaining 88 tasks using factor loadings from PCA. The list of reliability tasks and the corresponding scaling factors are provided in Appendix, Table AI.

The 88 reliability tasks validated in this paper can be used by decision makers and practitioners to assess the status of the reliability management practices within their organization to direct improvements. The sum of weighting factors for each key practice, and then sum of weighting factors for all key practices, is a maximum score against which electronics companies can be benchmarked during an evaluation of reliability capability. Graphical tools like bar and radar charts can be used for comparative analysis among companies.

The model can be tailored for specific scenarios and weighting factors can be adjusted based on unique requirements using the same approach. Further research can be conducted to evaluate the correlation between maturity level of an organization and reliability-related issues that have business impact. Research can also be conducted to evaluate costs associated with moving through different levels of maturity for an organization and optimum ways of improving maturity. This will enable trade-off decisions from a business standpoint and help organizations in identifying hidden cost benefits associated with delivering reliable products to customers.

References

AIAA (2004), *Objective-oriented Reliability and Maintainability Program Data Product Requirements*, AIAA S-102 Draft Document, AIAA, Reston, VA.

Antonius, R. (2003), *Interpreting Quantitative Data with SPSS*, Sage, London.

Bell Communications Research (1993), *Generic Requirements for Assuring the Reliability of Components Used in Telecommunication Equipment*, Technical Reference TR-NWT-000357, Bell Communications Research (Bellcore, now Telcordia), Morristown, NJ.

Benson, P.G., Saraph, J.V. and Schroeder, R.G. (1991), "The effects of organizational context on quality management: an empirical investigation", *Management Science*, Vol. 37 No. 9, pp. 1107-24.

Berdie, D.R. and Anderson, J.F. (1974), *Questionnaires: Design and Use*, Scarecrow Press, Metuchen, NJ.

Bryman, A. and Cramer, D. (2001), *Quantitative Data Analysis with SPSS Release 10 for Windows: A Guide for Social Scientists*, Routledge, London.

Chiesa, V., Coughlan, P. and Voss, C. (1996), "Development of a technical innovation audit", *Journal of Product Innovation Management*, Vol. 13 No. 2, pp. 105-36.

Churchill, G.A. (1979), "A paradigm for developing better measures of marketing research", *Journal of Marketing Research*, Vol. 16, pp. 64-71.

Cronbach, L.J. (1951), "Coefficient alpha and the internal structure of tests", *Psychometrika*, Vol. 16, pp. 297-334.

Cronbach, L.J. (2004), "My current thoughts on coefficient alpha and successor procedures", *Educational and Psychological Measurements*, Vol. 64 No. 3, pp. 341-418.

Cronbach, L.J. and Meehl, P.E. (1955), "Construct validity in psychological tests", *Psychological Bulletin*, Vol. 52, pp. 281-302.

Crosby, P.B. (1996), *Quality is Still Free: Making Quality Certain in Uncertain Times*, McGraw-Hill, New York, NY.

DoD (1980), *Reliability Program for Systems and Equipment Development and Production*, MIL-STD-785B, FSC RELI, AMSC Number F3104, Department of Defense.

Du Plessis, Y. (2003), "The development of an assessment tool for measuring project management culture in organizations", PhD dissertation, University of Pretoria, Pretoria.

Foster, J.J. (1998), *Data Analysis Using SPSS for Windows: A Beginner's Guide*, Sage, London.

Fraser, P. and Gregory, M. (2002), "A maturity grid approach for the assessment of product development collaborations", *Proceedings of 9th International Product Development Management Conference, Sophia Antipolis, Nice*.

Fraser, P., Moultrie, J. and Holdway, R. (2001), "Exploratory studies of a proposed design maturity model", *Proceedings of 8th International Product Development Management Conference, University of Twentie, Holland*.

Hinkin, T.R. (1995), "A review of scale development practices in the study of organizations", *Journal of Management*, Vol. 21 No. 5, pp. 967-88.

IEC TC-56 (2001), "Process for assessing reliability of equipment", International Electrotechnical Commission (IEC) New Work Item Proposal No. 56/775/NP, International Electrotechnical Commission, Geneva.

IEEE SA Standards Board (1998), *IEEE Standard Reliability Program for the Development and Production of Electronics Systems and Equipment, IEEE Std 1332-1998*, IEEE SA Standards Board, Piscataway, NJ.

Jacoby, J. (1978), "Consumer research: a state of the art review", *Journal of Marketing*, Vol. 42, pp. 87-96.

Kerr, A.W., Hall, H.K. and Kozub, S.A. (2002), *Doing Statistics with SPSS*, Sage, London.

Kline, P. (1994), *An Easy Guide to Factor Analysis*, Routledge, London.

Kuei, C.H., Madu, C.N. and Lin, C. (2001), "The relationship between supply chain quality management practices and organizational performance", *International Journal of Quality & Reliability Management*, Vol. 18 No. 8, pp. 864-72.

McGrath, M.E. (Ed.) (1996), *Setting the PACE in Product Development: A Guide to Product and Cycle Time Excellence*, Butterworth-Heinemann, Boston, MA.

Macbeth, D. and Fergusson, N. (1994), *Partnership Sourcing: An Integrated Supply Chain Management Approach*, Pittman, London.

Nunnally, J. (1978), *Psychometric Theory*, McGraw-Hill, New York, NY.

Paulk, M.C., Weber, C.V., Garcia, S.M., Chrisis, M.B. and Bush, M. (1993), "Key practices of the capability maturity model^{SM}, version 1.1", Technical Report CMU/SEI-93-TR-025, ESC-TR-93-178, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA.

SAE Technical Standards Board (1998), *Design/Process Checklist for Vehicle Electronic Systems*, SAE Document No. J-1938, SAE Technical Standards Board, Warrendale, PA.

Saraph, J.V., Benson, P.G. and Schroeder, R.G. (1989), "An instrument for measuring the critical factors of quality management", *Decision Sciences*, Vol. 20 No. 4, pp. 810-29.

Schriesheim, C.A., Powers, K.J., Scandura, T.A., Gardiner, C.C. and Lankau, M.J. (1993), "Improving construct measurement in management research: comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments", *Journal of Management*, Vol. 19 No. 2, pp. 385-417.

Strutt, J.E. (2001), "Reliability capability maturity briefing document", Report No. R-03/2/1, Reliability Engineering and Risk Management Centre, Cranfield University, Cranfield.

Szakonyi, R. (1994a), "Measuring R&D effectiveness – I", *Research Technology Management*, Vol. 37 No. 2, pp. 27-32.

Szakonyi, R. (1994b), "Measuring R&D effectiveness – II", *Research Technology Management*, Vol. 37 No. 3, pp. 44-55.

Thompson, B. and Daniel, L.G. (1996), "Factor analytic evidence for the construct validity of scores: a historical overview and some guidelines", *Educational and Psychological Measurement*, Vol. 56, pp. 197-208.

Tiku, S. and Pecht, M. (2003a), "Auditing the reliability capability of electronics manufacturers", *Advances in Electronic Packaging*, Vol. 1, pp. 947-53.

Tiku, S. and Pecht, M. (2003b), "Reliability capability assessment methodology", *Proceedings of IMAPS Brazil 2003, the International Technical Symposium on Packaging, Assembling and Testing and Exhibition, Campinas*.

Tiku, S., Azarian, M. and Pecht, M. (2007), "Using a reliability capability maturity model to benchmark electronics companies", *International Journal of Quality & Reliability Management*, Vol. 24 No. 5, pp. 547-63.

Tucker, L.R. and MacCallum, R.C. (2006), "Exploratory factor analysis", available at: www.unc.edu/~rcm/book/factornew.htm (accessed October 7, 2008).

Williams, K., Robertson, N., Haritonov, C.R. and Strutt, J. (2003), "Reliability capability evaluation and improvement strategies for subsea equipment suppliers", *Journal of the Society for Underwater Technology*, Vol. 25 No. 4, pp. 165-73.

Zhang, Z. (2001), "Implementation of total quality management: an empirical study of Chinese manufacturing firms", doctoral dissertation, University of Groningen, Groningen.

Further reading

Fraser, P., Moultrie, J. and Gregory, M. (2002), "The use of maturity models/grids as a tool in assessing product development capability", *Proceedings of the IEEE International Engineering Management Conference, IEMC'02, 1, Cambridge*, pp. 244-9.

Torgerson, W.S. (1958), *Theory and Methods of Scaling*, Wiley, New York, NY.

Appendix

Reliability task or trait		Factor loading XX	Weighting factor XX
1. *Reliability requirements and planning*			
1-01	Presence of a reliability department	0.364	1.00
1-02	Customer inputs in the form of their requirements and expectations	0.407	1.12
1-03	Capturing reliability specifications of competitive products	0.397	1.09
1-05	Product reliability plan that includes reliability goals and reliability activities' schedule	0.639	1.76
1-06	Establishing reliability goals for sub-assemblies and components in a product	0.618	1.70
1-07	Establishing reliability goals as a distribution and not as a point estimate	0.507	1.39
1-08	Establishing reliability goals for products based on specific life cycle conditions	0.605	1.66
1-09	While preparing a reliability plan, planning for required resources like materials, personnel and equipment	0.63	1.73
1-10	Including details on reliability analysis and testing for specific products as part of a reliability plan	0.684	1.88
1-11	Contingency planning and specification of decision criteria for altering reliability plans	0.653	1.79
1-12	Reliability plan includes a process for communicating results from reliability activities	0.683	1.88
1-04	Using reliability specs from old products while establishing requirements for new products	Deleted	
2. *Training and development*			
2-01	Reliability training plan or program	0.669	1.56
2-02	Formally trained reliability engineers	0.687	1.60
2-03	Top management commitment to reliability training	0.709	1.65
2-04	Business managers trained to appreciate importance of reliability to products or business	0.616	1.43
2-05	Reliability managers trained on how specific reliability activities can impact reliability	0.733	1.70
2-06	Reliability engineers trained to identify failure modes and mechanisms in a product design	0.651	1.51
2-07	Reliability engineers trained in statistical methods for reliability prediction and data analysis	0.612	1.42
2-08	Reliability engineers trained in failure analysis, root cause analysis and corrective actions	0.62	1.44
2-09	Reliability training provided to employees not directly associated with reliability, e.g. procurement and purchasing	0.43	1.00
2-10	Tracking new technologies, modeling, or analysis techniques that can impact reliability	0.54	1.26
3. *Reliability analysis*			
3-01	Identification of potential single points of failure and failure modes in a product design	0.66	1.26
3-02	Identification of potential failure mechanisms that can cause failures in a product design	0.69	1.32

(*continued*)

Table AI.
List of reliability tasks and scaling factors

Reliability task or trait		Factor loading XX	Weighting factor XX
3-03	Identification of critical failure modes and mechanisms in a product design	0.666	1.27
3-04	Quantification of risks and weaknesses for critical components in a product design	0.66	1.26
3-05	Checking adherence of a design to design rules	0.594	1.14
3-06	Making reliability point estimates using modeling or reliability prediction handbooks	0.567	1.08
3-07	Making reliability distribution predictions based on times to failure for potential failure mechanisms	0.66	1.26
3-08	Characterization of materials used in a product design	0.602	1.15
3-09	Using reliability predictions for specifying warranty periods and making spares provisioning	0.601	1.15
3-10	Using reliability analysis to design specific reliability tests for a product	0.668	1.28
3-11	Optimizing life cycle costs for a product based on reliability vs cost trade-offs	0.523	1.00
4. *Reliability testing*			
4-01	Tests to identify design margins and destruct limits for a product	0.545	1.17
4-02	Design verification or qualification tests for a product	0.505	1.09
4-03	Using reliability testing to make design changes in a product prior to production	0.563	1.21
4-04	Reliability testing based on generic specifications for all products	0.551	1.18
4-05	Reliability testing based on customer specifications	0.509	1.09
4-07	Reliability tests tailored to specific products	0.537	1.15
4-08	Detailed reliability test plans for products including sample sizes and confidence limits	0.737	1.58
4-09	Accelerated tests based on specific failure mechanisms to determine times to failure	0.761	1.64
4-10	Analysis of the test data to determine statistical failure distributions	0.771	1.66
4-11	Application of failure distributions to make reliability predictions using acceleration factors	0.763	1.64
4-12	Reviewing and updating reliability qualification test requirements for components	0.765	1.65
4-13	Minimizing reliability testing by using burn-in or environmental stress screening	0.465	1.00
4-06	Burn-in or screening of products prior to shipping	Deleted	
5. *Supply chain management*			
5-01	Component engineers for parts selection and supply management	0.574	1.38
5-02	Procuring parts only from authorized distributors and not from part brokers	0.667	1.61
5-03	Using manufacturing quality data for part selection and for in-coming lot rejection	0.662	1.60
5-04	Using reliability test data from suppliers for part selection and for in-coming lot rejection	0.597	1.44
5-05	Considering technology maturity of parts during part selection	0.654	1.58
5-06	Vendor or supplier assessments or audits	0.616	1.48
5-07	Using and maintaining a list of preferred/qualified/approved parts and suppliers	0.661	1.59
5-08	Using and maintaining a supplier rating system	0.707	1.70

(*continued*)

Table AI.

Reliability task or trait		Factor loading XX	Weighting factor XX
5-09	Using techniques like uprating for qualifying parts for use outside their data sheet specs	0.415	1.00
5-10	Supplier contractual agreements containing quality and reliability requirements	0.622	1.50
5-11	Multiple sourcing of parts	0.594	1.43
5-12	Tracking component traceability markings to identify any changes	0.693	1.67
5-13	Tracking part obsolescence to ensure continued supply or to make alternate supply arrangements	0.694	1.67
5-14	Review of supplier product change notices (PCNs) to assess their impact on manufacturability	0.726	1.75
5-15	Review of supplier PCNs to assess their impact on reliability	0.737	1.78
6. *Failure data tracking and analysis*			
6-01	Manufacturing defects and production testing failures tracked and recorded in a database	0.711	1.17
6-02	Reliability testing failures tracked and recorded in a database	0.755	1.24
6-03	Field failures tracked and recorded in a database	0.715	1.17
6-04	Ensuring traceability of products from manufacture to failure	0.755	1.24
6-05	Conducting failure analysis on failed products from all sources from manufacturing to field	0.704	1.16
6-06	Creating Pareto charts based on failure modes and failure sites	0.609	1.00
6-07	Conducting root cause analysis on failed products from all sources	0.657	1.08
6-08	Generating failure analysis reports detailing underlying failure mechanisms for failed products	0.726	1.19
6-09	Creating Pareto charts based on failure mechanisms	0.664	1.09
6-10	Correlating failure mechanisms with specific materials and processes	0.706	1.16
6-11	Creating and updating a database of corrective actions based on identified failure modes and mechanisms	0.784	1.29
7. *Verification and validation*			
7-01	Obtaining certifications like ISO for all management processes including reliability	0.463	1.00
7-02	Updating reliability predictions for products based on field data for present and previous products	0.781	1.69
7-03	Modifying statistical failure distributions used for reliability predictions on the basis of field failure data	0.794	1.71
7-04	Modifying reliability test conditions for current and future products based on failure mechanisms observed in field	0.801	1.73
7-05	Updating the failure modes database to incorporate any new failure modes observed in field	0.802	1.73
7-06	Updating the failure mechanisms database to incorporate any new failure mechanisms observed in field	0.804	1.74
7-07	Verifying and modifying warranty estimates and spares provisioning based on field returns	0.698	1.51
7-08	Internal audits for reliability planning, analysis, and testing activities	0.725	1.57
8. *Reliability improvements*			
8-01	Bill of materials' modification to exclude parts that have had reliability problems in field	0.665	1.20
8-02	Updating product reliability requirements due to business or market considerations	0.555	1.00

(*continued*)

Table AI.

Reliability task or trait		Factor loading XX	Weighting factor XX
8-03	Making design changes, if required, to accommodate changes in life cycle environment	0.679	1.22
8-04	Implementing corrective actions based on field failure modes	0.611	1.10
8-05	Implementing corrective actions based on field failure mechanisms	0.724	1.30
8-06	Requiring engineering change notifications for reliability improvements	0.647	1.17
8-07	Preventing recurrence of failures in future products, which have already been observed in existing products	0.671	1.21
8-08	Using field failure information to improve company design rules and process control requirements	0.748	1.35
8-09	Evaluating and implementing new modeling or analysis techniques to improve product reliability	0.674	1.21
8-10	Evaluating and implementing new technologies to improve product reliability	0.726	1.31

Table AI.

About the authors

Sanjay Tiku has an MS and a PhD degree in Mechanical Engineering from the University of Maryland, College Park. He currently works for Microsoft Corporation in Sammamish, Washington. He is accountable for development of design concepts for complex electromechanical products and creation of reliability specifications for them. He manages a team of engineers and researchers who use advanced computer-aided design analysis tools, experimental techniques, and statistical methods to ensure reliability and drive qualification of hardware products. Previously, he has worked at the Research Center of Tata Motors and also taught mechanical engineering for a while in India. His research interests include quality and reliability of electronic products and electronic parts selection and management. He has written several papers and book chapters in this area. He is a member of the IEEE and an invited member of the academic honor society Phi Kappa Phi. Sanjay Tiku is the corresponding author and can be contacted at: stiku22@gmail.com

Michael Pecht has an MS in Electrical Engineering and an MS and a PhD in Engineering Mechanics from the University of Wisconsin at Madison. He is a Professional Engineer, an IEEE Fellow and an ASME Fellow. He served as a Chief Editor of the IEEE Transactions on Reliability for eight years and on the advisory board of IEEE Spectrum. He is a Chief Editor for Microelectronics Reliability and an Associate Editor for the IEEE Transactions on Components and Packaging Technology. He is the founder of Center for Advanced Life Cycle Engineering at the University of Maryland, College Park, where he is also a Chair Professor in Mechanical Engineering. He has written more than 20 books on electronic products development, use and supply chain management and over 400 technical articles. He has been leading a research team in the area of prognostics for the past ten years and has now formed a new Prognostics and Health Management Consortium at the University of Maryland. He has consulted for over 50 major international electronics companies, providing expertise in strategic planning, design, test, prognostics, IP, and risk assessment of electronic products and systems. He was awarded the highest reliability honor, the IEEE Reliability Society's Lifetime Achievement Award in 2008. He has previously received the European Micro and Nano-Reliability Award for outstanding contributions to reliability research, 3M Research Award for electronics packaging, the IEEE Undergraduate Teaching Award and the IMAPS William D. Ashman Memorial Achievement Award for his contributions in electronics reliability analysis.