Sustainability **2015**, *7*, 15179-15193; doi:10.3390/su71115179

OPEN ACCESS

sustainability

ISSN 2071-1050

www.mdpi.com/journal/sustainability

Article

Repetitive Model Refinement for Questionnaire Design Improvement in the Evaluation of Working Characteristics in Construction Enterprises

Jeng-Wen Lin [1,*], Pu Fun Shen [2] and Bing-Jean Lee [1]

[1] Department of Civil Engineering, Feng Chia University, Taichung 407, Taiwan; E-Mail: bjlee@fcu.edu.tw

[2] Ph.D. Program in Civil and Hydraulic Engineering, Feng Chia University, Taichung 407, Taiwan; E-Mail: p0043264@fcu.edu.tw

* Author to whom correspondence should be addressed; E-Mail: jwlin@fcu.edu.tw; Tel: +886-4-2451-7250 (ext. 3150); Fax: +886-4-2451-6982.

Academic Editors: Adam Jabłoński and Marc A. Rosen

Received: 15 July 2015 / Accepted: 11 November 2015 / Published: 17 November 2015

Abstract: This paper presents an iterative confidence interval based parametric refinement approach for questionnaire design improvement in the evaluation of working characteristics in construction enterprises. This refinement approach utilizes the 95% confidence interval of the estimated parameters of the model to determine their statistical significance in a least-squares regression setting. If this confidence interval of particular parameters covers the zero value, it is statistically valid to remove such parameters from the model and their corresponding questions from the designed questionnaire. The remaining parameters repetitively undergo this sifting process until their statistical significance cannot be improved. This repetitive model refinement approach is implemented in efficient questionnaire design by using both linear series and Taylor series models to remove non-contributing questions while keeping significant questions that are contributive to the issues studied, *i.e.*, employees' work performance being explained by their work values and cadres' organizational commitment being explained by their organizational management. Reducing the number of questions alleviates the respondent burden and reduces costs. The results show that the statistical significance of the sifted contributing questions is decreased with a total mean relative change of 49%, while the Taylor series model increases the *R*-squared value by 17% compared with the linear series model.

Keywords: confidence interval; construction enterprises; questionnaire design; repetitive model refinement; statistical significance; working characteristics evaluation

1. Introduction

The questionnaire approach is widely used for surveying and collecting sample data with regard to an issue, with a list of questions to be answered and the results aggregated for statistical analysis. However, the main factors or questions influencing the findings of the models used need to be validated and simplified for efficient questionnaire design. In order to acquire accurate evaluations of working characteristics in construction enterprises and to alleviate problems of relatively large-dimensional and nonlinear models, this study develops a confidence interval based repetitive parametric model refinement approach for questionnaire design improvement.

1.1. General Information about the Questionnaires

A total of 250 questionnaires were distributed to Taiwanese and Chinese employees of two ranks in the company being studied. After excluding 30 invalid questionnaires (being incomplete or with missing values, or regarded as "outliers" through a set a mathematical analysis) and 39 unreturned ones, a total of 181 questionnaires were valid. The response rate was 72.4%.

1.2. Questionnaire Design Improvement

Questionnaire surveys are a widely used method to collect opinions and views. A customized questionnaire is developed based on the parameters revealed by context immersion in a given field (Kim [1]). However, many factors such as tedious design formats (Saris [2], Saris and Gallhofer [3]), redundant content, and excessive length (Weimiao and Zheng [4]) may lead to an inconsistent comparison matrix for the decision problem. Invalid or bad results from a questionnaire survey may cause decision makers to make faulty inferences (Ergu and Kou [5]). Suzuki *et al*. [6] introduced procedures to design reasonable questionnaires using statistical analysis to obtain high accuracy. Reducing the length of a survey by using a more streamlined set of questions can lead to more reasonable data being acquired and to better explanations of the issues in question. Other examples of this approach include Edwards *et al*. [7], who reduced the effective sample size and introduced bias. Finding ways to increase response rates to postal questionnaires would improve the quality of health research. Landsheer and Boeije [8] used qualitative facet analysis, an application of Guttmann's facet theory, to investigate whether item content sufficiently covered the intended subject area. This form of content analysis constitutes a systematic, effective, and critical tool for improving the content of questionnaires. Jacqui *et al*. [9] improved questionnaire design by enabling iterations of qualitative and quantitative testing, evaluation, and redevelopment. Results from such tests enable evidence-based decisions to be made regarding trade-offs between measurement error, processing error, non-response error, respondent burden, and costs. By enabling targeted improvements at the questionnaire design level according to specific needs, we can create valuable reference resources (Xu *et al*. [10]).

1.3. Model Refinement and Repetitive Computation

To alleviate problems of respondent burden and costs as well as relatively large-dimensional and nonlinear models, the issue of model refinement has increasingly drawn much attention in many fields. Smith [11] addressed the study of algorithms and system designs. Adrian [12] presented a refinement process with respect to data list building using model generators. Kapova and Goldschmidt [13] proposed model-driven application engineering based on the concept of analytical transformations. Liu [14] established two optimization models for a wireless optical communication system based on a four-level pulse amplitude modulation scheme. Ragnhild *et al*. [15] explored the behavior inheritance consistency of both refined and re-factored models with respect to the original model. Steven *et al*. [16] addressed model refinement as an iterative process. Zhuquan *et al*. [17] proposed that measurements permitted the repeated application of a system identification procedure operating on closed-loop data, together with successive refinements of the designed controller.

1.4. Nonlinear Models and Statistical Confidence Intervals

A nonlinear model is often adopted in system applications. Khorshid and Alfares [18] developed a parameter identification technique in creating a mathematical model of vehicle components by solving an inverse problem using a non-linear optimization method. Lin and Chen [19] proposed a statistical confidence interval based nonlinear parameter refinement approach and applied it to the standard power series model (Lin [20], Lin and Betti [21]) for the identification of structural systems. Other statistical confidence interval based studies include Tryon [22], who employed a graphical inference confidence interval approach in analyzing independent and dependent approaches for statistical difference, equivalence, replication, indeterminacy, and trivial difference. Yang *et al*. [23] proposed control limits based on the narrowest confidence interval to analyze problems, if the traditional three-sigma control limits or probability limits were adopted and some points with relatively high probability of occurrence were excluded; yet, some points with relatively small probability of occurrence may still be accepted in asymmetrical or multimodal distributions. Bonett and Price [24] proposed an adjusted Wald interval for paired binomial proportions that was shown to perform as well as the best available methods. In construction management, it has been shown to be feasible to use nonlinear models to deal with construction cost overruns (Ahiaga-Dagbui and Smith [25], Anastasopoulos *et al*. [26]) and schedule forecasting patterns (Kim and Kim [27], Patel and Jha [28]).

1.5. Prime Novelty Statement

In contrast with the conventional tests of reliability and validity, the designed questionnaires in this study were analyzed to identify the main factors and associated questions influencing the model studied using the proposed repetitive model refinement approach so as to streamline the number of questions in surveys of working characteristics in construction enterprises. Problems of respondent burden and costs as well as relatively large-dimensional and nonlinear models were thus alleviated. To reduce the number of questions with a more streamlined set, it was feasible to refine the model by repetitively removing non-contributing questions. Each time non-contributing questions were removed, the questionnaire model would be updated and rerun once again in a multiple regression setting. This model refinement approach for the content validity of the questionnaire was implemented using both linear and Taylor series models by conserving significant questions that were contributive to the issue being studied, *i.e.*, employees' work performance explained by their work values and cadres' organizational commitment explained by their organizational management. The results have been verified by calculating the statistical significance values of the sifted contributing questions and the *R*-squared values of established models.

2. Questionnaires Evaluating Working Characteristics in Construction Enterprises

In this study, the research subjects of the questionnaires were the Taiwanese employees and cadres of Taiwan-based construction enterprises in China. Questionnaire findings of similarities and differences in work values, work satisfaction, organizational management, and organizational commitment were preliminarily reviewed. The effects of work values and organizational management on work satisfaction and organizational commitment, respectively, were analyzed using questionnaires based on the job diagnostic survey by Hackman and Oldham [29]. The "working characteristics questionnaires" included questionnaires for (1) work values; (2) work performance and satisfaction; (3) organizational management; and (4) organizational commitment and identification (Lin and Shen [30], Shen [31]).

3. Repetitive Model Refinement Approach and Analyses

Questionnaire data were used in multiple regression analyses using four models, comprising the linear series, the refined linear series, the Taylor series, and the refined Taylor series model, where for the employees' part the independent variables are X = work values, which are used to explain the dependent variables Y = work performance and satisfaction; and for the cadres' part, X = organizational management, used to explain Y = organizational commitment and identification.

Two linear regression models were generated to identify the causal links between work values and work performance on the one hand, and organizational management and organizational commitment on the other. The original linear series model was refined through an iterative approach. This refined model was developed to streamline the questionnaire by removing non-contributing questions. The Taylor series model expanded the original linear series model up to the third moments. As a consequence, the *R*-squared value in the regression setting was increased. The refined Taylor series model was obtained from the original Taylor series model by the repetitive refinement approach in a regression setting. It was thus feasible to obtain the *R*-squared values of the regression between X and Y defined above and the mean relative change of the statistical significance as two indicators of result verification, so as to prove the accuracy of the refined model and to validate the sifted questions as genuinely significant contributors to the refined model.

The iterative refinement approach provides for the sifting of model components and related questions by repetitively using the 95% confidence interval in a regression setting. The 95% confidence interval is selected by convention and because the higher confidence interval enables more stringent selection of the components and thus a lower possibility of incorporating nonlinear elements, which is generally problematic for systems with a degree of nonlinear behavior; such nonlinearity will be verified in the results, showing the nonlinear Taylor series model significantly increases the

R-squared value when compared with the linear series model. If the estimated confidence interval of a parameter contains the "null" (zero) value, it is statistically valid to remove such a parameter and its corresponding component, while maintaining those parameters whose confidence intervals do not cover the zero value. This component/question sifting process is repeated by rerunning the regression and refining the model until none of the estimated 95% confidence intervals of the remaining parameters cover the zero value (Lin and Chen [19]). In addition, the interval method proposed in this article has proved more reasonable than the mean value method. Using the interval method considers an interval which covers zero or not. However, using the mean value method to remove those close to zero values has a problem; *i.e.*, what values are "close" to zero (e.g., 10^{-10}, 10^{-20}, or 10^{-30}, *etc.*)?

The employees' section of the questionnaire data is used in this study to demonstrate the model refinement approach using 95% confidence intervals in a regression. Using question Ey1 ("I think my work ability is excellent") as an example to show the model refinement approach, we assign Y = Ey1 in the questionnaire for employees' work performance and satisfaction, while X = Ex1–24, being all 24 questions in the questionnaire for employees' work values. In other words, the question Ey1 is explained by the questions Ex1–24. The consequent repetitive sifting process to select the real contributing components/questions out of the 24 questions (Ex1–24) to Ey1 is listed in Tables 1–4 (adapted from Lin and Shen [30], Shen [31]). Each table presents the outcome of a new regression after the component sifting process. Each of the highlighted upper and lower bounds for a given component indicates that the 95% confidence interval covers the zero value in the regression analysis.

Removing those components/questions with 95% confidence intervals covering the zero value in the regression setting of Table 1 and rerunning a new regression of the remaining components leads to Table 2. Continuing this repetitive sifting process by rerunning the regression analysis for the remaining components in Table 2 we obtain Table 3. By the same component sifting process, Table 4 is derived from Table 3. The 95% confidence interval for each remaining component in Table 4 does not cover the zero value, implying that the remaining components are genuine contributing factors in explaining the component Ey1. Hence, it is statistically valid to stop the component sifting process at this point. It is noteworthy that the significance value of each remaining component from Table 2 to Table 4 decreases in average a new regression is conducted in the repetitive refinement approach. The removed components correspond to relatively high significance values while the remaining components correspond to successively declining significance values in each round of regression.

Table 1. Multiple regression of original questionnaire model.

	***R*-square = 0.410**	**[95% Conf. Interval]**		
Ey1	I think my work ability is excellent.	Lower Bound	Upper Bound	Significance
Ex1	New knowledge and technologies can be learned at work.	−0.54	0.732	0.761
Ex2	There are chances for advanced studies at work.	−0.657	0.458	0.719
Ex3	My own dream can be realized at work.	−0.394	0.36	0.929
Ex4	The quality of my life can be improved through my work.	−0.502	−0.244	0.486
Ex5	My life becomes richer due to my work.	−0.476	−0.204	0.421
Ex6	I can have the sense of achievement at work.	0.126	0.612	0.19
Ex7	My boss at work is very understanding.	0.69	0.284	0.402
Ex8	My colleagues always take care of each other.	0.285	0.802	0.34

Table 1. *Cont.*

	***R*-square = 0.410**	**[95% Conf. Interval]**		
Ex9	My colleagues never attack each other for their own benefits.	−0.472	0.502	0.95
Ex10	My colleagues get along with each other well.	−0.45	0.36	0.821
Ex11	I can work in an environment which is not harmful to my body and mind.	0.152	0.499	0.683
Ex12	I can arrange my own schedule properly because of the flexibility of my work.	0.203	1.025	0.183
Ex13	When I am sick, the company takes good care of me.	0.845	2.044	0.404
Ex14	The insurance system of the company is good.	−1.654	2.033	0.836
Ex15	I can get a raise or bonus of a proper amount.	−2.445	−1.391	0.58
Ex16	The welfare system of the company is good.	0.145	2.375	0.605
Ex17	My income is higher than that of others with the same conditions as me.	−3.329	−1.822	0.556
Ex18	I never feel confused or scared while working.	0.371	1.672	0.204
Ex19	There are many chances of promotion.	−1.107	−0.416	0.362
Ex20	I devote myself to my work.	−0.841	0.757	0.916
Ex21	Even if there is no extra pay for working overtime, I would still work overtime to finish my work at night.	−0.529	0.69	0.79
Ex22	I usually go to work earlier to prepare the tasks I have to handle.	−0.474	0.642	0.762
Ex23	I am proud of my work.	0.189	1.407	0.13
Ex24	I want to be perfect when it comes to my work.	−2.01	−0.193	0.019

Table 2. Multiple regression of the refined questionnaire model in the first round.

	***R*-square = 0.399**	**[95% Conf. Interval]**		
Ey1	I think my work ability is excellent.	Lower bound	Upper bound	Significance
Ex4	The quality of my life can be improved through my work	−0.43	−0.174	0.398
Ex5	My life becomes richer due to my work.	−0.384	−0.177	0.461
Ex6	I can have the sense of achievement at work.	0.109	0.431	0.235
Ex7	My boss at work is very understanding.	0.499	0.176	0.339
Ex8	My colleagues always take care of each other.	0.156	0.591	0.247
Ex11	I can work in an environment which is not harmful to my body and mind.	−0.651	0.356	0.558
Ex12	I can arrange my own schedule properly because of the flexibility of my work.	0.131	0.814	0.152
Ex13	When I am sick, the company takes good care of me.	0.566	1.852	0.289
Ex15	I can get a raise or bonus of a proper amount.	−1.991	−1.038	0.529
Ex16	The welfare system of the company is good.	−0.888	2.231	0.39
Ex17	My income is higher than that of others with the same conditions as me.	−3.244	−0.951	0.276
Ex18	I never feel confused or scared while working.	0.117	1.647	0.087
Ex19	There are many chances of promotion.	−1.105	−0.174	0.149
Ex23	I am proud of my work.	0.107	1.113	0.104
Ex24	I want to be perfect when it comes to my work.	−1.674	−0.362	0.003

Table 3. Multiple regression of the refined questionnaire model in the second round.

	***R*-square = 0.395**	**[95% Conf. Interval]**		
Ey1	I think my work ability is excellent.	Lower bound	Upper bound	Significance
Ex4	The quality of my life can be improved through my work.	−0.44	−0.15	0.327
Ex5	My life becomes richer due to my work.	−0.386	−0.153	0.387
Ex6	I can have the sense of achievement at work.	0.084	0.446	0.176
Ex7	My boss at work is very understanding.	0.463	0.197	0.42
Ex8	My colleagues always take care of each other.	0.189	0.529	0.345
Ex12	I can arrange my own schedule properly because of the flexibility of my work.	0.076	0.645	0.119
Ex13	When I am sick, the company takes good care of me.	0.499	1.874	0.249
Ex15	I can get a raise or bonus of a proper amount.	−2.109	−0.823	0.381
Ex17	My income is higher than that of others with the same conditions as me.	−2.258	0.771	0.426
Ex18	I never feel confused or scared while working.	0.03	1.694	0.058
Ex19	There are many chances of promotion.	−1.12	−0.141	0.125
Ex23	I am proud of my work.	0.103	1.05	0.105
Ex24	I want to be perfect when it comes to my work.	−1.633	−0.39	0.002

Table 4. Multiple regression of the refined questionnaire model in the third round.

	***R*-square = 0.392**	**[95% Conf. Interval]**		
Ey1	I think my work ability is excellent.	Lower bound	Upper bound	Significance
Ex4	The quality of my life can be improved through my work.	−0.452	−0.128	0.267
Ex5	My life becomes richer due to my work.	−0.394	−0.139	0.341
Ex6	I can have the sense of achievement at work.	0.088	0.439	0.186
Ex7	My boss at work is very understanding.	0.473	0.18	0.372
Ex8	My colleagues always take care of each other.	0.206	0.5	0.405
Ex12	I can arrange my own schedule properly because of the flexibility of my work.	0.074	0.645	0.117
Ex13	When I am sick, the company takes good care of me.	0.582	1.55	0.065
Ex15	I can get a raise or bonus of a proper amount.	−2.21	−0.409	0.173
Ex18	I never feel confused or scared while working.	0.089	1.453	0.082
Ex19	There are many chances of promotion.	−1.015	−0.134	0.17
Ex23	I am proud of my work.	0.058	1.076	0.077
Ex24	I want to be perfect when it comes to my work.	−1.629	−0.392	0.002

4. Results and Verifications

4.1. Statistical Significance of Question

The relative change of the statistical significance value before and after each round of the repetitive refinement approach in the regression setting is defined as:

$$\frac{x_j^f - x_j^i}{x_j^i} \tag{1}$$

where x_j^f denotes the final statistical significance value for the *j*th component of the model, while x_j^i denotes the initial statistical significance value for the *j*th component of the model. The statistical significance is defined as follows: If the *p*-value is less than or equal to alpha, we say that the data are statistically significant at level alpha. In statistics (where "significant" means "corresponds to a real difference in fact") the term is used to indicate only that the evidence against the null hypothesis reaches the standard set by alpha (Moore and McCabe [32]). Since the lower the significance value of a component the higher will be its contribution to the model, a negative value for the relative change of the statistical significance in Equation (1) signifies that the effect of the corresponding component/question on the model is increased, while the opposite is true for the case of a positive value. Tables 5 and 6 list the relative change of the statistical significance as a percentage (%) for each question of Ey explained by Ex1–24 and for each question of Cy explained by Cx1–8, respectively.

Table 5. Employees' part: relative change of the statistical significance for each question of Ey explained by Ex1–24.

Work Satisfaction / Work Values	Ey1	Ey2	Ey3	Ey4	Ey5	Ey6	Ey7	Ey8	Ey9	Ey10
Ex1		−34%		−50%		−38%	−97%		−19%	−90%
Ex2		42%				−50%			−59%	−17%
Ex3		−13%				−28%		−20%	−37%	
Ex4	−45%				20%	−77%	−74%	−77%	−28%	−32%
Ex5	−19%			−1%		−47%			−55%	0.3%
Ex6	−2%				−45%		−64%	−21%		
Ex7	−7%				−59%	−56%		−42%	−46%	
Ex8	19%			−80%	−26%	−90%	−0.3%	−72%		
Ex9				−31%	−20%		−66%		−44%	−50%
Ex10				−17%			−13%		−8%	
Ex11		−74%		−48%			−67%	−27%	−58%	−100%
Ex12	−36%			−71%			−58%	−43%	−61%	−38%
Ex13	−84%	−70%		−15%		−69%		−7%		−14%
Ex14					−31%	−70%	−32%	−24%	−51%	−23%
Ex15	−70%	−85%			−48%		−8%	−2%		−12%
Ex16		−79%			−59%					
Ex17				−94%		−100%	−21%	−97%		−81%
Ex18		−78%		−27%		−71%		−25%		
Ex19	−53%			−4%		−70%		−42%		
Ex20		−13%		−6%		−34%		−30%		
Ex21				−44%	−37%	−17%				−55%
Ex22		−91%		−28%	−50%		−20%	−77%	−97%	−74%
Ex23	−41%			−15%	−56%		−61%	−46%		−60%
Ex24	−89%	−31%		−40%		−38%	−84%	−58%		−49%
Mean change	−41%	−48%		−37%	−37%	−57%	−48%	−42%	−47%	−46%
Total Mean Change							−45%			

Table 6. Cadres' part: relative change of the statistical significance for each question of Cy explained by Cx1–8.

Organizational commitment / **Organizational management**	**Cy1**	**Cy2**	**Cy3**	**Cy4**	**Cy5**	**Cy6**	**Cy7**	**Cy8**	**Cy9**	**Cy10**
Cx1	−68%		−56%	−40%	−74%	−57%	0%	−5%		−72%
Cx2	−85%	−7%	−64%		−25%	−83%	0%	−33%	−91%	−27%
Cx3		−91%		−83%		−53%	0%	−33%	−92%	−11%
Cx4	−96%		−98%	−74%	−60%		0%	−35%	−93%	−11%
Cx5	−88%	−48%				−53%	0%	12%		−37%
Cx6			−45%		−42%		0%	−19%		−2%
Cx7		−48%		−74%	−69%	−40%	0%	−35%	−93%	
Cx8		1%		−85%	−39%	−36%	0%		−92%	−95%
Mean change	−84%	−39%	−66%	−71%	−52%	−54%	0%	−21%	−92%	−36%
Total mean change	−52%									

In Table 5, a blank indicates that the question used to explain the corresponding question Ey in a model has been removed. All the questions used to explain the question Ey3 have been removed, implying that Ey3 ("My boss thinks I am doing a great job at work") has nothing to do with any of the questions relating Ex1–24. Such a question should be removed to improve questionnaire design for accurate evaluations of working characteristics. It is clear that all the significance values of the remaining questions are decreased except for the four marked values. Such a decrease in the significance value refers to the increase of the effect of the question on a model, verifying that the remaining questions are the real contributing questions/factors for the refined model. The total mean relative change of the statistical significance of the remaining variables is −45%.

Similarly in Table 6, a blank indicates that the question used to explain the corresponding question Cy in a model has been removed. Again, the significance values of the remaining questions are clearly decreased except for the two marked values. Such a decrease in the significance value verifies that the remaining questions are the real contributing questions/factors to the refined model. The total mean relative change of the statistical significance of the remaining variables is −52%. In particular, the question Cy7 "Staying and working for this company doesn't do me any good" needs to be explained by all eight questions Cx1–8 relating to organizational management. In other words, choosing whether to stay and work for the company depends on the entire range of the company's management strategies.

4.2. R-Squared Value of Regression Analysis

In the regression setting, the final R-squared value of each Ey for the employees' part through the repetitive refinement approach implemented in the linear series, refined linear series, Taylor series, and refined Taylor series models is listed in Table 7 (adapted from Lin and Shen [30], Shen [31]). The total mean R-squared value is decreased by 0.02 for the refined linear series model from the linear series model, signifying that the model refinement approach developed here cannot truly affect the R-squared value when searching for the genuinely contributory questions for survey improvement. On the other hand, the Taylor series model increases the mean *R*-squared value by 0.19 from the linear series model, which greatly improves the modeling process in the multiple regression setting.

Table 7. Employees' part: Final *R*-squared values for linear series, refined linear series, Taylor series, and refined Taylor series models.

X **= Work Values** ***Y*** **= Work Performance and Satisfaction**		**Linear Series**	**Refined Linear Series**	**Taylor Series**	**Refined Taylor Series**
Ey1	I think my work ability is excellent.	0.41	0.392	0.593	0.533
Ey2	I can always finish my work rapidly on time.	0.407	0.366	0.624	0.562
Ey3	My boss thinks I am doing a great job at work.	0.285	0.208	0.389	0.26
Ey4	My professional knowledge is enough to do my job.	0.46	0.449	0.684	0.638
Ey5	I am highly cooperative with my team.	0.314	0.302	0.521	0.479
Ey6	I am very satisfied with the welfare provided by the company I work for.	0.555	0.53	0.692	0.632
Ey7	I am very satisfied with what this job has to offer to help improving my future development.	0.521	0.499	0.743	0.694
Ey8	I am very satisfied with my salary.	0.493	0.487	0.699	0.656
Ey9	I am very satisfied with my relationships with my colleagues.	0.495	0.481	0.708	0.661
Ey10	I am very satisfied with the opportunities and the system of promotion.	0.531	0.524	0.713	0.663
	Overall mean per model	0.44	0.42	0.63	0.57

Similarly, the final *R*-squared value of each Cy for the cadres' part obtained by the repetitive refinement approach in the linear series, refined linear series, Taylor series, and refined Taylor series models is listed in Table 8 (adapted from Lin and Shen [30], Shen [31]). The total mean *R*-squared value is again decreased by 0.02 for the refined linear series model. The Taylor series model on average increases the *R*-squared value by 0.17 from the linear series model, greatly improving the modeling process. In Table 8, all the questions implemented in the Taylor series model achieve high *R*-squared values of greater than 0.85, implying a satisfactory result in modeling the causal explanations for questionnaire design.

Table 8. Cadres' part: Final *R*-squared values for linear series, refined linear series, Taylor series, and refined Taylor series models.

X = Organizational Management *Y* = Organizational Commitment and Identification		Linear Series	Refined Linear Series	Taylor Series	Refined Taylor Series
Cy1	I care about the future development of the company.	0.785	0.757	0.942	0.879
Cy2	In order to stay employed by the company, I am willing to accept any assignment.	0.723	0.681	0.911	0.793
Cy3	In order to help the company to be successful, I am willing to pay extra efforts.	0.757	0.753	0.934	0.848
Cy4	It doesn't matter to work for another company as long as job content and conditions are similar.	0.724	0.692	0.894	0.817
Cy5	I think the company I work for is a good company, and it's worthy to work hard for it.	0.769	0.765	0.938	0.842
Cy6	The style of this company is close to my values.	0.797	0.772	0.956	0.844
Cy7	Staying and working for this company doesn't do me any good.	0.97	0.97	0.999	0.999
Cy8	I would leave this company as long as my job status is slightly changed.	0.647	0.613	0.854	0.768
Cy9	I can identify myself with the company's policy for its employees.	0.781	0.771	0.939	0.897
Cy10	I am glad that I decided to take this job instead of others.	0.656	0.653	0.859	0.753
Overall mean per model		0.76	0.74	0.93	0.84

4.3. Reliability and Validity

Verifications and error analyses were also conducted to compare the above results using the repetitive model refinement approach with those using methods of reliability and validity.

This study adopted Cronbach's alpha to represent the reliability in data analysis. Guieford [33] proposed a set of criteria for Cronbach's alpha. The standard value of Cronbach's alpha is 0.5. High alpha values (>0.7) mean high reliability while low ones (<0.35) mean low reliability. Table 9 shows that through the repetitive model refinement approach the number of questions was reduced and all the reliabilities were over 0.7, indicating that the sample was adequately stable and consistent.

Table 9. Reliability analyses.

	Before deleting questions	After deleting questions
Employees' work values	Cronbach's alpha = 0.623	Cronbach's alpha = 0.720
Employees' work performance and satisfaction	Cronbach's alpha = 0.577	Cronbach's alpha = 0.742
Cadres' organizational management	Cronbach's alpha = 0.565	Cronbach's alpha = 0.740
Cadres' organizational commitment and identification	Cronbach's alpha = 0.590	Cronbach's alpha = 0.780

Validity in SPSS on the other hand means "exploratory factor analysis" (according to SPSS online help), whose main features are the following tests:

(1) Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy tests whether the partial correlations among variables are small (KMO > 0.6);
(2) Bartlett's Test of Sphericity tests the null hypothesis that the correlation matrix is an identity matrix, indicating that the factor model is inappropriate (Sig < 0.05);
(3) SPSS analysis defines communality as the proportion of a parameter's variance that is explained by the factor structure.

This repetitive model refinement approach thus reduces the number of questions and can be shown to promote communality significantly; this also indicates that validity was not reduced after questions had been deleted, as illustrated in Table 10.

Table 10. Exploratory factor analysis.

	Before Deleting Questions	**After Deleting Questions**
Employees' work values	KMO = 0.816 Bartlett Test Sig = 0.03 Communality = 0.768	KMO = 0.772 Bartlett Test Sig = 0.01 Communality = 0.811
Employees' work performance and satisfaction	KMO = 0.763 Bartlett Test Sig = 0.01 Communality = 0.798	KMO = 0.733 Bartlett Test Sig = 0.00 Communality = 0.828
Cadres' organizational management	KMO = 0.741 Bartlett Test Sig = 0.00 Communality = 0.739	KMO = 0.709 Bartlett Test Sig = 0.00 Communality = 0.801
Cadres' organizational commitment and identification	KMO = 0.712 Bartlett Test Sig = 0.01 Communality = 0.754	KMO = 0.700 Bartlett Test Sig = 0.01 Communality = 0.799

5. Conclusions

This study is consistent with sustainable development issues, dealing with four areas: employees' work values; employees' work performance and satisfaction; cadres' organizational management; and cadres' organizational commitment and identification. The questionnaire data are available for reference and for enterprises' development. In addition, the questionnaire design improvement can assist researchers to design more precise and effective questionnaires. In this study, an effective repetitive model refinement approach using 95% confidence intervals in a multiple regression setting has been applied to the analysis of questionnaire design improvement for evaluating working characteristics in construction enterprises. Such an approach sifts components/questions by removing non-contributing questions of the model, inducing only a 2% decrease in the model's corresponding *R*-squared value, while keeping the genuinely contributory questions of the model for questionnaire design improvement. This not only reduces the time to complete the questionnaire in surveys, but also reduces the cost of production of the questionnaire. The results prove that the developed Taylor series model significantly increases the *R*-squared value by 17% when compared with the linear series model.

After repeatedly running the screening process of the estimated parameters, almost all the remaining questions of the model for both the employees' and cadres' sections show decreased significance values with a total mean relative change of 49%, verifying that the remaining questions are indeed the real contributing ones to the models studied. In particular, the question "My boss thinks I am doing a great job at work" in evaluating employees' work performance cannot be successfully explained by the contents of the questionnaire relating to employee work values. Such a question should instead be evaluated by a manager within the repetitive model refinement approach. However, the question "Staying and working for this company doesn't do me any good" can be evaluated through the full content of the questionnaire relating to organizational management. In other words, an employee's decision to stay in the company is substantially dependent on the company's management strategies. Further, limitations of the study indicate that the developed questionnaire design improvement should be applied to data with high reliability.

Acknowledgments

The work described in this paper comprises part of the research project sponsored by Feng Chia University (Contract No. 14I42315), whose support is greatly appreciated.

Author Contributions

Jeng-Wen Lin designed the research and wrote the paper; Pu Fun Shen performed research and analyzed the data; and Bing-Jean Lee revised the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Kim, M.J. A framework for context immersion in mobile augmented reality. *Autom. Constr.* **2013**, *33*, 79–85.
2. Saris, W.E. Questionnaire Design. 2014. Available online: http://link.springer.com/referenceworkentry/10.1007%2F978-94-007-0753-5_2392#page-1 (accessed on 12 November 2015).
3. Saris, W.E.; Gallhofer, I.N. Design, Evaluation, and Analysis of Questionnaires for Survey Research. 2014. Available online: https://books.google.com.tw/books?hl=zh-TW&lr=&id=NTKpAgAAQBAJ&oi=fnd&pg=PT12&ots=xHVXLPqMua&sig=Vfoid8qPqQZ2we861FwFBaI3Ccs&redir_esc=y#v=onepage&q&f=false (accessed on 12 November 2015).
4. Weimiao, F.; Zheng, Y. Factors Affecting Response Rates of the Web Survey: A Systematic Review. *Comput. Hum. Behav.* **2009**, *26*, 132–139.
5. Ergu, D.; Kou, G. Questionnaire Design Improvement and Missing Item Scores Estimation for Rapid and Efficient Decision Making. *Ann. Oper. Res.* **2012**, *197*, 5–23.

6. Suzuki, S.; Ando, M.; Hashimoto, H.; Asama, H. Design Procedure and Improvement of a Mathematical Modeling to Estimate Customer Satisfaction. 2014. Available online: http://link.springer.com/chapter/10.1007/978-4-431-54816-4_2#page-1 (accessed on 12 November 2015).
7. Edwards, P.J.; Roberts, I.; Clarke, M.J.; DiGuiseppi, C.; Wentz, R.; Kwan, I.; Cooper, R.; Felix, L.M.; Pratap, S. Methods to Increase Response to Postal and Electronic Questionnaires. *Cochrane Database Syst. Rev.* **2009**, doi:10.1002/14651858.MR000008.pub4.
8. Landsheer, J.A.; Boeije, H.R. In Search of Content Validity: Facet Analysis as a Qualitative Method to Improve Questionnaire Design An Application in Health Research. *Qual. Quant.* **2008**, *44*, 59–69.
9. Jacqui, J.; Pete, B.; Sarah, W.; Jane, C. Improved Questionnaire Design Yields Better Data: Experiences from the UK's Annual Survey of Hours and Earnings. In Proceedings of the International Conference on Establishment Survey (ICES–III), Montréal, QC, Canada, 18–21 June 2007.
10. Xu, S.; Zhou, S.; Cao, Q.; Lei, J.; Li, X.; He, Y. Questionnaire Design and Analysis of Online Teaching and Learning: A Case Study of the Questionnaire of "Education Online" Platform of Beijing University of Technology. 2014. Available online: http://link.springer.com/chapter/10.1007/978-94-007-7618-0_419 (accessed on 12 November 2015).
11. Smith, D.R. Model Refinement: Calculating Refinements in Algorithm and System Design. 2009. Available online: ftp://kestrel.edu/pub/papers/smith/mr.pdf (accessed on 12 November 2015).
12. Adrian, V. Performance Criteria for Software Metrics Model Refinement. *J. Appl. Quant. Methods* **2007**, *2*, 118–128.
13. Kapova, L.; Goldschmidt, T. Automated Feature Model-Based Generation of Refinement Transformations. In Proceedings of the 2009 SEAA '09, 35th Euromicro Conference on Software Engineering and Advanced Applications, Patras, Greece, 27–29 August 2009; pp. 141–148.
14. Liu, X. Solution Refinement of Wireless Optical System Model with Multi-level Modulation. *Electron. Lett.* **2009**, *45*, 475–476.
15. Ragnhild, V.D.S.; Viviane, J.; Tom, M. A Formal Approach to Model Refactoring and Model Refinement. *Softw. Syst. Model.* **2005**, *6*, 139–162.
16. Steven, J.L.; Philip, R.B.; Chiu, W. EMAN: Semiautomated Software for High-Resolution Single-Particle Reconstructions. *J. Struct. Biol.* **1999**, *128*, 82–97.
17. Zhuquan, Z.; Bitmead, R.R.; Gevers, M. H2 Iterative Model Refinement and Control Robustness Enhancement. In Proceedings of the 30th IEEE Conference on Decision and Control, Brighton, UK, 11–13 December 1991; Volume 1, pp. 279–284.
18. Khorshid, E.; Alfares, M. Model Refinement and Experimental Evaluation for Optimal Design of Speed Humps. *Int. J. Veh. Syst. Model. Test.* **2006**, *2*, 80–99.
19. Lin, J.W.; Chen, H.J. Repetitive Identification of Structural Systems using a Nonlinear Model Parameter Refinement Approach. *Shock Vib.* **2009**, *16*, 229–240.
20. Lin, J.W. Adaptive Algorithms for the Identification of Nonlinear Structural Systems. Ph.D. Thesis, Columbia University, New York, NY, USA, 2001.
21. Lin, J.W.; Betti, R. On-line Identification and Damage Detection in Non-linear Structural Systems using a Variable Forgetting Factor Approach. *Earthq. Eng. Struct. Dyn.* **2004**, *33*, 419–444.

22. Tryon, W.W. Evaluating Independent Proportions for Statistical Difference, Equivalence, Indeterminacy, and Trivial Difference Using Inferential Confidence Intervals. *J. Educ. Behav. Stat.* **2009**, *34*, 171–189.
23. Yang, J.; Xie, M.; Goh, T.N. Control Limits Based on the Narrowest Confidence Interval. *Commun. Stat. Theory Methods* **2011**, *40*, 2172–2181.
24. Bonett, D.G.; Price, R.M. Adjusted Wald Confidence Interval for a Difference of Binomial Proportions Based on Paired Data. *J. Educ. Behav. Stat.* **2012**, *37*, 479–488.
25. Ahiaga-Dagbui, D.D.; Smith, S.D. Dealing with Construction Cost Overruns Using Data Mining. *Constr. Manag. Econ.* **2014**, *32*, 682–694.
26. Anastasopoulos, P.C.; Haddock, J.E.; Peeta, S. Cost Overrun in Public-Private Partnerships: Toward Sustainable Highway Maintenance and Rehabilitation. 2014. Available online: http://ascelibrary.org/doi/abs/10.1061/%28ASCE%29CO.1943-7862.0000854 (accessed on 12 November 2015).
27. Kim, B.C.; Kim H.J. Sensitivity of Earned Value Schedule Forecasting to S-Curve Patterns. *J. Constr. Eng. Manag.* **2014**, *140*, 04014023. Available online: http://ascelibrary.org/doi/10.1061/%28ASCE%29CO.1943-7862.0000856 (accessed on 12 November 2015).
28. Patel, D.A.; Jha, K.N. Neural Network Model for the Prediction of Safe Work Behavior in Construction Projects. 2015. Available online: http://ascelibrary.org/doi/10.1061/%28ASCE%29CO.1943-7862.0000922 (accessed on 12 November 2015).
29. Hackman, J.R.; Oldham, G.R. Development of the Job Diagnostic Survey. *J. Appl. Psychol.* **1975**, *60*, 159–170.
30. Lin, J.W.; Shen, P.F. Factor-analysis Based Questionnaire Categorization Method for Reliability Improvement of Evaluation of Working Conditions in Construction Enterprises. *Struct. Eng. Mech.* **2014**, *51*, 973–988.
31. Shen, P.F. Impact of Employees' Work Values at Two Sides of Taiwan Straits on Corporate Performance. Master's Thesis, Feng Chia University, Taichung, Taiwan, 2011.
32. Moore, D.S.; McCabe, G.P. *Introduction to the Practice of Statistics*, 3rd ed.; W.H. Freeman and Company: New York, NY, USA, 2000.
33. Guieford, J.P. *Fundamental Statistics in Psychology and Education*, 4th ed.; McGraw Hill: New York, NY, USA, 1965.