Creativity Research Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hcrj20

Confirming the Three-Factor Creative Product Analysis Matrix Model in an American Sample

Susan P. Besemer & Karen O'Quin

Published online: 08 Jun 2010.

Creativity Research Journal
1999, Vol. 12, No. 4, 287–296

Confirming the Three-Factor Creative Product Analysis Matrix Model in an American Sample

Susan P. Besemer
State University of New York College at Fredonia
Karen O'Quin
State University of New York College at Buffalo

ABSTRACT: The purpose of this study was to confirm, in an American sample, the validity of a three-factor structural equation model that had been previously tested and refined with a sample of young adult students in Norway (Besemer, 1998). The three-factor model, a Creative Product Analysis Matrix (Besemer & Treffinger, 1981), was tested by analyzing the responses to the Creative Product Semantic Scale by a sample of American college students from two State University of New York colleges. Confirmatory factor analyses provided strong support for construct validity of the questionnaire and the three-dimensional creativity model. Participant judges were able to detect differences perceived in Novelty, Resolution, and Elaboration and Synthesis of the 4 stimulus items.

Persons involved in the development and evaluation of products need ways of assessing the creativity of those products. In the business world, for example, it is critical to new product development to have adequate, reliable, and valid means for evaluating the quality, including the creativity, of the products brought to market (Stone-Romero & Stone, 1997). Teachers also need to assist their students to improve the creativity of their products: works of art, term papers, and class presentations (Treffinger & Poggio, 1972). Those who review public displays of art and performances of dance, theater, and music need to be able to make reasoned judgments of the quality and creativity of the works they review (Perkins, 1979).

Much rests on the accuracy and validity of such creativity judgments in all of the areas just listed, yet valid and reliable means of evaluation are used infrequently, and judgments are often made quickly and intuitively (Cooper & Kleinschmidt, 1986). Although it is possible that intuition may have internalized important criteria for judgment on the basis of long and valuable experience for some judges, current needs for accountability and objectivity require more explicitness in the statement of review criteria and standards of judgment.

Creative Product Analysis Matrix

The Creative Product Analysis Matrix (CPAM; Besemer & Treffinger, 1981) theory was developed to help cultivate more careful observation of created products and to focus judges' attention on relevant attributes of products. The CPAM is a three-dimensional model of creativity in products, which hypothesizes Novelty, Resolution, and Elaboration and Synthesis as the three factors. *Novelty* considers newness in materials, processes, concepts, and methods of making the product. *Resolution* considers aspects of how well the product works or functions. *Elaboration and Synthesis* describes stylistic components of the product. Making up the three factors are nine facets. These are, for Novelty, originality and surprise; for Resolution, logical, useful, valuable, and understandable; and for Elaboration and Synthesis, organic, well-crafted, and elegant.

Manuscript received January 5, 1999; accepted May 1, 1999.
Correspondence and requests for reprints should be sent to Susan P. Besemer, 10056 Patterson Lane, Fredonia, NY 14063. E-mail: IdeaFusion@AOL.com

The Novelty factor is based on the widely accepted concept that novelty is a critical component of creativity in products. Although in practice, the term *novelty* is sometimes used synonymously with *creativity,* suggesting a unidimensional model, most researchers and many practitioners stress the importance of another major dimension related to a product's usefulness or functionality.

The Resolution factor is based on this concept. Many researchers develop their models of creativity on this two-factor model (e.g., Csikszentmihalyi, 1996). Considering a product's resolution prevents one from acknowledging an admittedly bizarre or unrealistic product as creative, as the product must be workable, in addition to being original. Unless a product solves a problem for its creator, it is not a creative product.

The CPAM also includes a third factor, Elaboration and Synthesis. This factor considers the perceived attributes of style in the product's production. Although it is less often cited in the literature, considerations of style are often unexpressed criteria of creativity. In some fields, such as the visual and performing arts, these criteria are more explicitly expressed. One might expect to find stylistic considerations in evaluating art, but judgments of style are also expressed in other fields such as mathematics, when a solution may be termed *elegant* for the simplicity, insight shown, and conciseness of its presentation.

Based on this theory, the Creative Product Semantic Scale (CPSS) was developed. It has been in the process of validation for more than a decade (Besemer, 1998; Besemer & O'Quin, 1986, 1987, 1993; O'Quin & Besemer, 1989). Development of a scale to measure creativity in products began in 1980 when the theoretical model was created, based on a review of the literature in which criteria for creativity in products were cited (Besemer & Treffinger, 1981). Item generation took the form of lists of criteria cited in the many articles reviewed. From 1981 until 1986 several versions of self-report instruments (varying from 70 to 110 items) were developed to test the CPAM model. These instruments were tested using different types of products and various groups of lay judges.

CPSS

The CPSS is an evaluation instrument designed to assess the creativity that is perceived to be manifested in products. The CPSS is based on the model of creativity described earlier, the CPAM. The CPSS posits that untrained judges, using a validated and reliable instrument, can make informed judgments of creativity in products. The ability to use untrained judges in studies of creativity increases the generalizability of results to the natural environment and demystifies the notion of creativity in products. It also allows users to improve the creativity of products under consideration by strengthening weaker attributes. The questionnaire is designed to elicit the participants' perceptions of the products being evaluated through the instrument, along lines of the model. The CPSS has been in formal development since 1986 when preliminary testing of a series of bipolar adjective scales began (Besemer & O'Quin, 1986). The instrument underwent statistical and psychometric analyses, including reliability testing of items and scales, and factor analysis and the deletion of items that did not improve reliability (Besemer, 1998). The subscales have undergone many changes and have been reduced in number not only to improve their reliability but also to make them easier to administer (Besemer, 1998).

In 1994 a Norwegian language version of the CPSS was tested in Norway through the collection of data there at two folk high schools (Besemer, 1998). Participants completed evaluations of three highly novel chairs using the CPSS. The version of the CPSS discussed here reflects the changes that were made to the instrument at that time. The CPSS instrument was shortened from 55 item pairs to be scored to 43, and a subscale (Understandable) that had been hypothesized to be an indicator of Elaboration and Synthesis was found to be a better indicator of Resolution.

Purpose and Hypotheses

It was possible that the results found in the Norwegian study might not generalize to another culture, or to less novel products. The purpose of this study was first to confirm the three-factor CPAM model found in the Norwegian study in an American sample. Second, we wanted to further assess the generalizability of the CPAM model by including a more traditional product.

As in the Norwegian study, chairs were chosen as the products to be evaluated by lay judges. For replication purposes, three of the products chosen for this study were novel chairs that had been used in Norway.

A fourth traditional American chair was included to extend the testing of the model.

It was predicted that Novelty ratings would be lower for the traditional chair and that Resolution ratings would be higher for the traditional chair. We expected that confirmatory factor analysis (CFA) would show that the CPAM model, developed using Norwegian students, would be replicated in the American sample using both the same products and a more traditional product.

Method

Participants

Participants ($N = 185$) were students in introductory psychology courses in the fall of 1997 at two colleges of the State University of New York: Buffalo and Fredonia. Mean age of participants was 20.79 years ($SD = 3.76$). Seventy-four percent of the participants were female. The students participated voluntarily in the study but received extra credit in their courses as a result of their participation. Table 1 presents the demographics of the sample.

Materials and Procedure

Chairs were used in this study for several reasons. Chairs are products with which everyone has experience. It is also possible by looking at a slide of a chair to imagine how comfortable and sturdy it would be, because chairs are ubiquitous in our daily life. Weber (1996) wrote persuasively about the value of studying commonplace inventions to reveal conceptual principles that underlie the language of invention. In addition to this theoretical reason for studying chairs, they are good objects to use in studies such as this because people feel comfortable sharing their opinions about these common objects, so much a part of Western society. Three of the chairs used in this study were highly novel and only somewhat similar to other chairs that most people have used. This helped to focus participants' attention on the task and consider their judgment of each chair with attention to detail. In this study, the three novel chairs that were evaluated in Norway were used, and a fourth chair (a traditional colonial-style wooden chair) was added as a contrast. All four chairs were evaluated by all participants.

Table 1. *Demographic Description of the Study Sample*

Group	Gender	*n*	%	*M* Age	*SD*
Buffalo		125.00	67.57	21.38	4.26
	Male	34.00	18.38	21.34	2.97
	Female	90.00	48.65	21.41	4.69
	Missing[a]	1.00	0.54		
Fredonia		60.00	32.43	19.55	1.93
	Male	14.00	7.57	19.54	2.15
	Female	46.00	24.86	19.56	1.88
Total		185.00	100.00	20.79	3.76

Note: $N = 185$.
[a]One participant omitted age and gender.

In the Norwegian study, the stimulus slides of chairs were selected by a group of four expert judges (two creativity instructors, a working architect, and a professional painter) as the most novel, yet least useful of six products selected from a book that describes a gallery exhibit of art chairs. (See Besemer, 1998, for photographs of the three novel chairs.) The three chairs were termed Ritz Boxes, which looked as though it had been made of assorted cans of soup, cereal boxes, and other food items from the grocery store; Soft Auto, which resembled an upholstered sedan of the 1950s; and Garden Chaise, a sculptural bench made of smooth, hard, triangular elements. The addition of the traditional chair called Patriot was made to extend the testing of the model in the American study and to further contrast the scores of the highly novel chairs. The reason that highly novel, less useful chairs were selected initially was to help focus attention on the Resolution factor of the scale in the validation effort.

The CPSS instrument is administered in a group setting and it takes approximately 15 min to evaluate one product. The CPSS is scored on 7-point Likert-type scales, ranging from 1 to 7 between bipolar adjectives such as *old–new*. Each of the nine subscales is created of four or five items.

Subscale scores are constructed by taking the mean of the items that make up the subscale. For example, the subscale Elegant has five items (pairs of adjectives): graceful–awkward, refined–busy, coarse–elegant, repelling–charming, and attractive–unattractive. A participant's score for Elegant is computed by taking the mean of the scores for these items. Some items are presented in reverse order, requiring recoding so that higher scores consistently represent higher ratings.

The three scales can be used together or individually to fit the researcher's needs.

The CPSS instrument was administered to the students outside of class in 11 small-group sessions. After completing informed consent statements and receiving instructions on the instrument, they completed the scale items while viewing a slide of the stimulus product, one of the four chairs. The order of presentation of slides was counterbalanced to avoid bias from fatigue or comparison. When the group was finished evaluating one slide, the next one was presented. The total time for evaluating four products was about 60 min.

Results

The data set was examined for missing data and normality. The data appeared to be normally distributed. Recoding of reversed items was performed, such that higher scores meant higher ratings. Missing data were imputed by mean substitution, and items were assembled into subscales, according to the model. To remove the effects of any non-normality in the study sample, the data were transformed by log function, which reduced the range of the means but ensured more meaningful comparisons.

Statistical Analyses

Reliability analyses were performed on the scales and subscales to check for internal consistency of the judgments made by the participants. A repeated measures multivariate analysis of variance (MANOVA) was performed to compare the rankings of the products on each subscale to determine if participants perceived differences among the four chair stimuli, according to the attributes identified in the model. Finally, CFA was performed to determine if the model proposed in the Norwegian study adequately represented the data collected in the American study sample. In CFA, robust statistics were requested, which further reduced the influence of any non-normality in the sample.

Reliability

Internal consistency reliability estimates (i.e., Cronbach's alphas) for the measures of Novelty, Resolution, and Elaboration and Synthesis were computed for each of the four stimulus slides. Table 2 presents the scale reliabilities, most of which were good and some of which were excellent. They ranged from a low of .69 on Novelty for the Patriot chair to a high of .86 for Garden Chaise. It is interesting to note that the mean alpha for the scales applied to the Patriot (traditional) chair were the lowest.

Comparisons Among Products

A repeated measures MANOVA was performed using the nine subscale scores as dependent variables across the four products. The multivariate main effect for product was significant, $F(27, 157) = 78.77, p < .001$. Detailed results of the MANOVA are presented in Table 3.

All of the multivariate *F*s are significant, as may be seen in Table 3. Examination of the univariate *F*s revealed that there, too, all of the perceived variance among stimulus items on each variable was significant at $p < .001$. As might be expected, Novelty rankings were highest of all for Ritz Boxes, Soft Auto, and Garden Chaise (the three novel chairs of the Norwegian study). Novelty was perceived as the weakest dimension for the traditional chair, Patriot, which rated 2.47 on originality and 2.473 on surprise. The highest ratings for Patriot were 2.80 for its usefulness and 2.80 for its organic qualities.

As predicted, Patriot received the highest ratings of any of the chairs on the subscales of the Resolution dimension. It received the highest rating for logicalness,

Table 2. *Reliability of the Scales By Dimension and Product*

Dimensions	Ritz Boxes α	Soft Auto α	Garden Chaise α	Patriot α
Novelty	0.76	0.72	0.84	0.69
Resolution	0.84	0.85	0.85	0.83
Elaboration & Synthesis	0.81	0.80	0.86	0.81
Mean α	0.80	0.79	0.85	0.78

Table 3. *Comparisons Among Products: Multivariate Analysis of Variance*

Dimension	Scale	Ritz Boxes[a]	Soft Auto[a]	Garden Chaise[a]	Patriot[a]	*F*
Novelty						
	Surprising					532.71*
	M	2.74	2.73	2.70	2.47	
	SD	0.07	0.06	0.08	0.06	
	Originality					634.50*
	M	2.77	2.76	2.74	2.47	
	SD	0.06	0.06	0.07	0.07	
Resolution						
	Logicalness					251.50*
	M	2.55	2.66	2.56	2.78	
	SD	0.09	0.09	0.10	0.05	
	Usefulness					318.01*
	M	2.51	2.70	2.58	2.80	
	SD	0.09	0.09	0.11	0.05	
	Value					234.72*
	M	2.52	2.62	2.54	2.75	
	SD	0.09	0.09	0.09	0.06	
	Understandability					207.75*
	M	2.60	2.69	2.58	2.77	
	SD	0.09	0.08	0.10	0.04	
Elaboration and Synthesis						
	Organic qualities					84.43*
	M	2.69	2.73	2.66	2.80	
	SD	0.11	0.07	0.09	0.04	
	Well-craftedness					140.81*
	M	2.61	2.72	2.65	2.78	
	SD	0.11	0.08	0.09	0.05	
	Elegance					173.48*
	M	2.55	2.63	2.60	2.75	
	SD	0.09	0.09	0.08	0.06	

Note: Originally, scales ranged from 1 to 7, from *low* to *high* on each variable. Scales were transformed using the natural logrithmic function.
[a]$n = 185$.
*$p < .001$.

usefulness, value, and understandability. For these subscales, Ritz Boxes scored lower than Patriot. The scores for Garden Chaise were slightly higher on this dimension, and Soft Auto's scores were generally higher than Ritz Boxes and Garden Chaise. It is interesting to consider the understandability subscale to realize that the participants regarded the Patriot as the most understandable product and the Garden Chaise as the least understandable.

Participants found the Patriot chair to have many organic qualities (e.g., orderliness and completeness). They rated this subscale highest for Patriot (2.80), and it was the highest rating of any subscale for any product, even higher than originality for Ritz Boxes. Patriot had high ratings for Elaboration and Synthesis, followed by those for Soft Auto, Ritz Boxes, and Garden Chaise.

CFA

CFAs were performed on the sample data set ($N = 185$) for each of the four chairs, using a generalized three-factor model derived and tested in the Norwegian study. A path diagram for this model may be seen in Figure 1.

The theory implied by the model hypothesized in Figure 1 is that variances in judgment about the creativity of each of the four products analyzed can be adequately explained by three correlated factors (Novelty, Resolution, and Elaboration and Synthesis), that each factor has a nonzero loading on the factor it was designed to measure, and that error terms are uncorrelated. The three-factor confirmatory model discussed here included fixing the first of the regression paths for each factor to a value of 1.0, allowing the others to be freely estimated. The 21 parameters that were estimated were for the three factor covariances among the three factors, the nine regression coefficients between the dependent variables and the factors, and the nine measurement error variances.

The purpose for the CFA was to test the model using new data collected in the study using American student participants. CFA was used to estimate parameters and assess how well correlations that were reproduced, given the model specified, fit the set of correlations of the new data set. CFA thus allows for a quantitative test of a theoretical model. The analysis was performed using the software program EQS for Windows, Release 5.1 (Bentler & Wu, 1995).

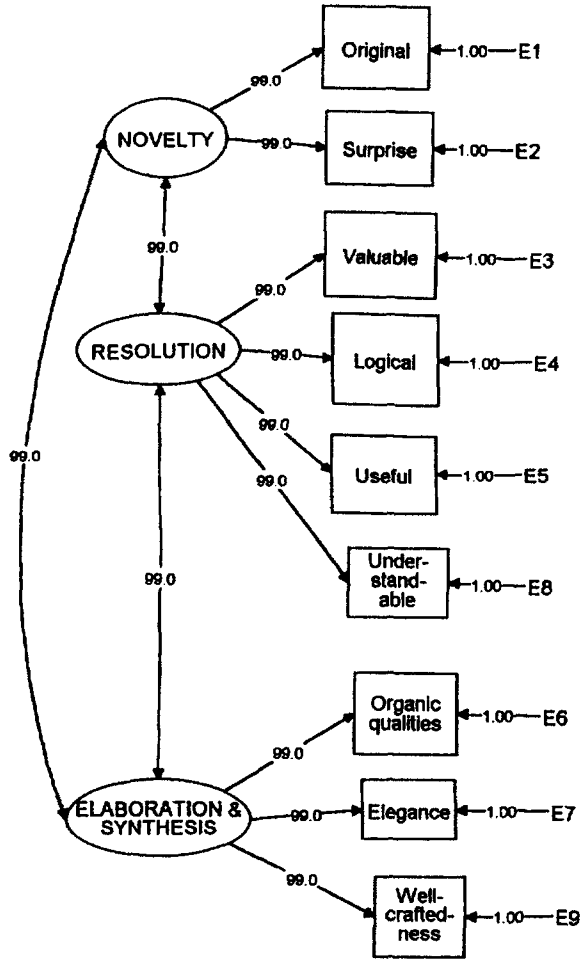

Figure 1. *Generalized three-factor model.*

In CFA, multiple criteria are used to evaluate the fit of the model to the data in the study. No single criterion is completely sufficient on its own, so using multiple criteria helps to evaluate the model thoroughly. The paths linking the variables to the factors are reviewed. These should all be significant. Another criterion is the value of the average off-diagonal absolute standardized residuals. One looks for a symmetric distribution of the residuals around zero, with most falling within the 95% confidence interval around zero. Another criterion is the chi-square statistic, which should be low, with a nonsignificant probability. The ratio of chi-square to the degrees of freedom is examined. Various authors have recommended values from as low as 2 or 3 to 5 (Bollen & Long, 1993, p. 3; Schumacker & Lomax, 1996, p. 121). Finally, one reviews the various fit indexes and modification indexes, looking for values to exceed .9 on the fit indexes. The modification indexes provide statistically based guidance in adding or dropping parameters from the model to improve the fit. These indexes must be viewed with caution, as they are specific to the data sample at hand and have no bearing to the theoretical constructs being tested. No matter how much adding some parameters or deleting others will improve the fit, one should not alter the model without verifying the reasonableness of the change as it relates to the theory of the model.

For each of the chairs tested the sample size was n = 185 and the chi-square statistic for each chair tested p < .01. The goodness-of-fit indexes generated by EQS suggested that the general model of product creativity specified in the path diagram was a reasonable representation of the data collected.

Looking at the individual assessments of each chair's model and data can elucidate the value of the model. These results may be consulted in Table 4.

Starting with Ritz Boxes, the loadings of all variables on their respective factors were reasonable and significant. The standardized residuals were distributed evenly and close to zero. The average off-diagonal absolute standardized residuals was .06. The chi-square for the null model for Ritz Boxes was 778.89 (df = 36). When considering the hypothesized model, a much better fit was seen, with a Satorra–Bentler (Sartorra & Bentler, 1988a, 1988b) scaled chi-square (S–Bχ^2) of 63.23 (df = 24). The Robust Comparative Fit Index (CFI; Byrne, 1994) for Ritz Boxes was .94.

Moving to Soft Auto, the loadings of all variables on their respective factors were reasonable and significant. Although the model for Soft Auto in the Norwegian study suggested small correlated errors between organic and understandable, these did not appear in this sample. The standardized residuals were distributed fairly evenly and close to zero. The average off-diagonal absolute standardized residuals was .07. The chi-square for the null model for Soft Auto was 969.19 (df = 36). When considering the hypothesized model, a somewhat better fit was seen, with an S–Bχ^2 of 102.88 (df = 25). The Robust CFI for Soft Auto was .94. It is clear that improvement could be made in fitting the model to the data of this sample, but because such modifications would be based on the characteristics of this sample only, and because the model is adequate without the modifications, the model was allowed to remain as initially specified.

Garden Chaise, an uncomfortable-looking bench, was represented adequately with no modifications. For Garden Chaise, as for the other chairs, the loadings of all variables on their respective factors were reasonable and significant. No correlated errors emerged, and the model was evaluated in its initial hypothesized form. The standardized residuals were distributed evenly and close to zero. The average off-diagonal absolute standardized residuals was .07. The chi-square for the null model for Garden Chaise was 968.04 (df = 36). The hypothesized model fit the data somewhat better, with an S–Bχ^2 of 98.07 (df = 25). The Robust CFI for Garden Chaise was .91. This model is another that could benefit from fitting based on the unique aspects of this sample. Again, because the model was adequate without revision, and because the conceptual theory did not warrant them, these modifications were not made.

For Patriot, the loadings of all variables on their respective factors were again reasonable and significant. As in the Norwegian model, the hypothesized model included correlated errors between valuable and logical and between organic and logical. Although one never hopes for correlated error values, these were small and reasonable to the content of the concept. In fact, in earlier versions of the CPAM, understandable was hypothesized to load with organic on Elaboration

Table 4. *Evaluation of the Models*

Chair	Null Model χ^2	df	Hypothesized Model S–B χ^2	df	S–B χ^2/df	RCFI
Ritz Boxes	778.89	36.00	63.23	24.00	2.63	0.94
Soft Auto	969.19	36.00	72.21	25.00	2.89	0.94
Garden Chaise	968.04	36.00	98.07	25.00	3.92	0.91
Patriot	745.60	36.00	53.28	23.00	2.32	0.93

Note: S–B = Satorra–Bentler Scaled (Sartorra & Bentler, 1988a, 1988b) Statistic (chi-square); RCFI = Robust Comparative Fit Index.

and Synthesis. The standardized residuals were distributed evenly and close to zero. The average off-diagonal absolute standardized residuals was .05. The chi-square for the null model for Ritz Boxes was 745.60 (df = 36). The hypothesized model fit the data much better, with an S–Bχ^2 of 53.28 (df = 23). The Robust CFI for Patriot was .93.

Several types of fit statistics were calculated by EQS. The chi-square to the degrees of freedom ratio (χ^2/df) as well as the Robust CFI, recommended by Byrne (1994), was used. To minimize the effects of any non-normality of the sample, robust statistics (including robust standard errors) were requested, as well as the S–Bχ^2. Byrne (1994) stated that this statistic can be very helpful in evaluating covariance structure models under "various distributions and sample sizes" (p. 86).

Another measure of the adequacy of the ability of the model to accurately account for the variance in the data collected are the various fit indexes generated by the analysis. Although many fit indexes are provided, the most reliable index for data in smaller samples that may be distributed somewhat non-normally is the Robust CFI (Byrne, 1994). This index can range from 0 to 1, with values in excess of .9 representing an adequate fit of the data with the model.

For this study, Robust CFIs were computed for each of the four chairs under consideration. For the chair called Ritz Boxes, the Robust CFI was .936. For Soft Auto, the Robust CFI was .942. For Garden Chaise, the Robust CFI was .913. For the Patriot chair, the Robust CFI was .930.

When the χ^2/df ratio was calculated for each of the modified models for each product, the resulting statistic was acceptable, although the statistic reached significant levels. Acceptable χ^2/df statistics range between 2 and 5 (Hayduk, 1987; Schumacker & Lomax, 1996), with lower values being preferable. The S–Bχ^2/df for the chair called Ritz Boxes was equal to 63.23/24, or 2.63. The S–Bχ^2/df for Soft Auto was equal to 72.21/25, or 2.89. The S–Bχ^2/df for the chair called Garden Chaise was equal to 98.07/25, or 3.92. The S–Bχ^2/df for the chair called Patriot was equal to 53.28/23, or 2.32.

CFA provides modification indexes that can be used to more closely fit the model to the data by adding and dropping parameters from the model. Two of the chairs required modifications to the generalized model to improve the fit, as had been the case with the Norwegian data set. It appears that the unique aspects of any product create a certain degree of lack of fit to the generalized model. Yet, with relatively minor modifications, the model adequately explained the data. Although some fitting was done, as just described, it was decided to decline any additional modifications to the model.

Recall that the purpose of this study was to test the generalizability of the CPAM using new data gathered about the stimulus items that had been tested in Norway. It is interesting to compare the Robust CFIs computed in the Norwegian study to those found here. In the Norwegian study (Besemer, 1998), the Robust CFI for Ritz Boxes was .96, that for Soft Auto was .997, and that for Garden Chaise was .94. The Patriot chair was not used in the Norwegian study.

Discussion

It appears that the models created during the Norwegian study adequately fit the new sample. The lack of fit demonstrated by the decline in the fit statistics shows the value of limiting the extent of fitting the model to a particular set of data. Although it is possible to achieve a perfect fit by adding and releasing constraints, the resultant model has reduced generalizability. It is far preferable to develop a flexible, multipurpose model that can be used with a variety of data samples.

This study yielded considerable support to confirm the validity of a three-factor model of creativity in products. Although CFA cannot prove that any model is the one correct solution for explaining variance in data, it can identify acceptable solutions and test competing models. In this study, the three-factor model was adequately confirmed for each of the four chairs (including an additional product that had not been tested in the earlier Norwegian study). These results contribute to the construct validity of the three-factor model and its usefulness in two different cultures. Both in the Norwegian study and in the current American study, participants were able to use the CPSS to differentiate perceptions among the different chairs. Further, the addition of a traditional chair allowed the model to be tested again on a different product. The addition of this fourth chair provides an example of the predictive capability of CFA. Given a tested and adequate model, based on conceptually strong theory and the experience of several exploratory factor analyses, it should be possible to predict that perceptions of the new product would fall within the dimensions that had been established. In the case of Patriot, this was clearly the case, even considering the differing perceptions of this product from the others evaluated in the study. Although the judgments of this traditional chair were significantly different from those of the novel chairs, the factor structure established in earlier studies and in this one with the novel chairs could adequately represent the new data.

The ability of the lay judges to demonstrate significant differences among their judgments of the chairs shows the usefulness of the instrument at detecting perceived differences in product creativity, even when naive participants serve as judges of product creativity.

Earlier versions of the instrument have already been used by researchers to judge graphic designs created under two different settings (Howe, 1992) and in making judgments about design elements of quilts (Franklin, 1997), for example. The CPSS can, therefore, be used to judge products in a research setting. It can also be used with other measures, as part of a multitest screening process, in the evaluation of new product ideas in the development process. It may be used to improve creative works in progress by focusing attention on the three dimensions (Novelty, Resolution, and Elaboration and Synthesis). When the nine subscale scores are determined, attention may be given to strengthening lower scoring attributes.

Implications for Future Research

Another important line of research can proceed from this basis. Much study has been undertaken on the creative personality, with certain measures of cognitive style (e.g., Kirton, 1985; Martinsen, 1993) receiving considerable attention. It could be interesting to explore the relation, if any, between cognitive style and product creativity. For example, do innovators prefer products that score higher in Novelty? Do assimilators prefer products that are higher in Resolution? These questions warrant more consideration, as does the continued study of creative product analysis in various other cultures. Judgments of products in non-Western cultures would be especially interesting and would provide the opportunity for further confirmation of the factor structure of the CPAM through the CPSS.

References

Bentler, P. M., & Wu, E. J. C. (1995). EQS for Windows (Version 5.1) [Computer software]. Encino, CA: Multivariate Software.

Besemer, S. P. (1998). Creative Product Analysis Matrix: Testing the model structure and a comparison among products—Three novel chairs. *Creativity Research Journal, 11,* 333–346.

Besemer, S. P., & O'Quin, K. (1986). Analyzing creative products: Refinement and test of a judging instrument. *Journal of Creative Behavior, 20,* 115–126.

Besemer, S. P., & O'Quin, K. (1987). Creative product analysis: Testing a model by developing a judging instrument. In S. G. Isaksen (Ed.), *Frontiers of creativity research: Beyond the basics* (pp. 341–357). Buffalo, NY: Bearly Ltd.

Besemer, S. P., & O'Quin, K. (1993). Assessing creative products: Progress and potentials. In S. G. Isaksen, M. C. Murdock, R. L. Firestien, & D. J. Treffinger (Eds.), *Nurturing and developing creativity: The emergence of a discipline* (pp. 331–349). Norwood, NJ: Ablex.

Besemer, S. P., & Treffinger, D. J. (1981). Analysis of creative products: Review and synthesis. *Journal of Creative Behavior, 15,* 158–178.

Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models.* Newbury Park, CA: Sage.

Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming.* Thousand Oaks, CA: Sage.

Cooper, R. G., & Kleinschmidt, E. J. (1986). An investigation into the new product process: Steps, deficiencies, and impact. *Journal of Product Innovation Management, 3,* 71–85.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention.* New York: HarperCollins.

Franklin, J. (1997). *Relationship between cognitive style and creative products.* Unpublished master's thesis, State University of New York College at Buffalo.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essential and advances.* Baltimore: Johns Hopkins University Press.

Howe, R. P. (1992). Uncovering the creative dimensions of computer-graphic design products. *Creativity Research Journal, 5,* 233–243.

Kirton, M. (1985). Adaptors, innovators, and paradigm consistency. *Psychological Reports, 57,* 487–490.

Martinsen, Ø. (1993). Insight problems revisited: The influence of cognitive styles and experience on creative problem solving. *Creativity Research Journal, 6,* 435–447.

O'Quin, K., & Besemer, S. P. (1989). The development, reliability, and validity of the revised Creative Product Semantic Scale. *Creativity Research Journal, 2,* 267–278.

Perkins, D. N. (1979). Are matters of value matters of fact? In C. J. Nodine & D. J. Fisher (Eds.), *Perception and pictorial representation* (pp. 301–315). New York: Praeger.

Satorra, A., & Bentler, P. M. (1988a). Scaling corrections for chi square statistics in covariance structure analysis. In *American Statistical Association. 1988 proceedings of business and economic sections* (pp. 308–313). Alexandria, VA: American Statistical Association.

Satorra, A., & Bentler, P. M. (1988b). *Scaling corrections for statistics in covariance structure analysis* (University of California at Los Angeles statistics series 2). Los Angeles: University of California at Los Angeles, Department of Psychology.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stone-Romero, E. F., & Stone, D. L. (1997). Development of a multidimensional measure of perceived product quality. *Journal of Quality Management, 2,* 87–112.

Treffinger, D. J., & Poggio, J. P. (1972). Needed research on the measurement of creativity. *Journal of Creative Behavior, 6,* 253–267.

Weber, R. J. (1996). Toward a language of invention and synthetic thinking. *Creativity Research Journal, 9,* 353–367.