

This article was downloaded by: [University of Sydney]

On: 01 May 2015, At: 15:34

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Creativity Research Journal

Publication details, including instructions for authors and subscription information:  
<http://www.tandfonline.com/loi/hcrj20>

### The development, reliability, and validity of the revised creative product semantic scale

Karen O'Quin<sup>a</sup> & Susan P. Besemer<sup>b</sup>

<sup>a</sup> Psychology, Buffalo State College, State University of New York College, 1300 Elmwood, Buffalo, NY, 14222

<sup>b</sup> Reed Library, State University of New York College, Fredonia

Published online: 02 Nov 2009.

To cite this article: Karen O'Quin & Susan P. Besemer (1989) The development, reliability, and validity of the revised creative product semantic scale, Creativity Research Journal, 2:4, 267-278, DOI: [10.1080/10400418909534323](https://doi.org/10.1080/10400418909534323)

To link to this article: <http://dx.doi.org/10.1080/10400418909534323>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

---

# The Development, Reliability, and Validity of the Revised Creative Product Semantic Scale

---

Karen O'Quin

*State University of New York College at Buffalo*

---

Susan P. Besemer

*Reed Library*

*State University of New York College at Fredonia*

---

**ABSTRACT:** *The Creative Product Semantic Scale (CPSS) is based on a theoretical model which conceptualizes three dimensions of product attributes: Novelty, Resolution, and Elaboration and Synthesis. The revision of the CPSS is described in the present article. Study 1 confirmed that the instrument could detect differences in the level of Novelty of three commercially available household products varying in unusualness. Study 2 was designed to validate the three dimensions of product attributes by testing whether or not the rankings of experts would agree with those of lay judges. Results indicated that judgments of Novelty made by naive judges were consistent with those of experts. The validity of Novelty, and to a lesser degree, that of Elaboration and Synthesis, were verified by Study 2. The validity of Resolution is yet to be established (no significant differences among products were observed). In both studies, Novelty and Resolution were independent, but Elaboration and Synthesis subscales migrated between loading with Novelty and Resolution.*

The literature of creativity has traditionally focused on the creative personality, the creative process, and to a lesser degree, the creative product. For more than 30 years, researchers have viewed products as windows through which a glimpse of the creativity of the products' makers might be seen (Ghiselin, 1958). The ultimate focus of such activity has been on the creative personality and the processes used by the creator. Ghiselin's words from 1958 are still true today:

what really is required is an examination of products, in an effort to discover some grounds on which we can assign or deny to them the epithet "creative" in a fully intelligible and defensible way. (p. 142)

Taylor and Sandler (1972) and MacKinnon (1978) reiterated the rationale for using the product as a reflection of the level of creativity of the artist or creator. Although this concept has remained largely uncontested through the years, little empirical research on creative products has followed these authors' strong recommendations. Theoretical papers abound in the field of creativity, but empirical studies of the cre-

---

Karen O'Quin, Psychology, Buffalo State College, 1300 Elmwood,  
Buffalo, NY 14222.

ativity manifested in products have been relatively few (see Amabile, 1982).

## Creative Product Analysis

Creative product analysis attempts to isolate the elements of products to determine what makes them more or less creative. Early research in the area of product analysis considered products in art (Barron, Gaines, Lee, & Marlow, 1973) and engineering and chemistry (Taylor & Sandler, 1972). More recently, Parke and Byrnes (1984) described the *Creative Product Scales* of the Detroit Public Schools, designed to be used and interpreted by professionals in the areas of art, music, creative writing, performing arts, and dance. The goal of developing these scales was to achieve more consistent and explainable evaluations of students' products.

Such research, although valuable within a specific discipline, has limited usefulness to the broader realm of creative product analysis. A generalized theoretical framework for understanding creative products, called the *Creative Product Analysis Matrix*, was proposed by Besemer and Treffinger (1981) and refined by Besemer and O'Quin (1986, 1987). The matrix, or CPAM, was developed to apply to a wide variety of products. It includes three dimensions along which products can be evaluated. Each dimension is further divided into several attributes.

The first dimension in the CPAM is *Novelty*, which has long been identified as a necessary, but typically not a sufficient, criterion for creativity (e.g., Runco, 1988). In the CPAM, the three attributes of this dimension are whether a product is Original, Surprising, and Germinal. The last attribute is related to a product's perceived influence in suggesting other new products.

The second dimension is the *Resolution* of the product, which measures the extent to which a product meets the practical needs of the problem situation. The three attributes included in the CPAM are whether a product is Valuable, Logical, and Useful. The third dimension, *Elaboration and Synthesis*, considers the aspects of style or production values. The CPAM includes five attributes: whether a product is Organic, Elegant, Complex, Understandable, and Well-Crafted.

Research using the CPAM has been aimed at developing a measuring instrument which might be used by diverse groups, such as engineers, new product designers, artists, and students to consider and describe their creative products. The development of the instrument has occurred in three stages. Early work within the CPAM's theoretical framework proceeded using a 110-item checklist (Besemer & Treffinger, 1981). Further studies using the CPAM sought to verify the appropriateness of the theoretical model's three dimensions (Besemer & O'Quin, 1987), using the same 110 items, but answered on a 4-point descriptive scale.

Because of measurement problems revealed by this early work, a shorter instrument--the *Creative Product Semantic Scale* (CPSS; Besemer & O'Quin, 1986)--was developed from the theoretical matrix. It employed the semantic differential format based on the work of Osgood, Suci, and Tannenbaum (1957). The CPSS contained 71 bipolar adjectives divided into 11 subscales, each measuring an attribute of the three dimensions of the CPAM discussed above. Empirical data from a study using everyday products (T-shirts), indicated that the 11 subscales were reliable (Besemer & O'Quin, 1986).

## Pilot Study

The pilot study was conducted with three goals: (a) to begin to establish validity by determining whether or not the CPSS could detect differences in the level of *Novelty* of products judged a priori to differ in *Novelty*, (b) to examine the generalizability of the CPSS by applying it to different products; (c) to further refine the subscales. Products were chosen for the pilot study which were to varying degrees unusual, yet which were alike in one way: applying animal concepts to everyday objects. However, the products were chosen because they differed in the level of *Novelty* of the application, as judged a priori by the authors.

## Method

### *Subjects*

Subjects were 194 undergraduates from the State University College at Buffalo; 119 were female, 69 were male, and 6 did not report their gender. The majority ( $n = 101$ ) were from introductory psychology, with the remainder from Creative Studies and Library Science classes. Their average age was 24 years (range 17 to 66).

### *Materials*

Three products were chosen from the 1984 Neiman-Marcus *Christmas Book*,<sup>1</sup> and were judged by the authors a priori to differ in *Novelty*. Product 1, judged least novel, was a white ceramic pitcher shaped like a plump cat with a red and blue bandanna around its neck; its paw was raised to form a spout, and its tail formed the handle. Product 2

was a wooden armchair with a natural finish, covered with small irregular black dots like a Dalmatian; two small black dog heads decorated the fronts of both arms. Product 3, judged most novel, was a desk shaped like a life-size longhorn steer; the finish was wood inlaid to look like steerhide, and compartments dropped down from the steer's side.

### *Procedure*

Subjects participated in groups of varying sizes. One of two female experimenters presented subjects with informed-consent statements containing a brief description of the rating task. Subjects then received a questionnaire booklet with written instructions which were clarified by the experimenter. Three optical scanning sheets were also included. The booklet contained the 71-item version of Besemer and O'Quin's (1986) CPSS. The experimenter displayed a slide of one of the products (the order of presentation was counterbalanced), and read a one-sentence description of the product. All subjects finished rating one product before the next product was displayed.

## Results

### *Reliability of Subscales*

The 71-item scale was initially divided into 11 subscales on the basis of past research analyses (Besemer & O'Quin, 1986). These subscales were subjected to iterative relia-

---

<sup>1</sup>Neiman-Marcus *Christmas Book* 1984 (Dallas, TX: Neiman-Marcus, p. 60, item b, p. 110 item a, p. 111, item d).

bility analyses. Two criteria were used to delete items from the subscales: a low corrected item-total correlation, and an increase in coefficient alpha if the item were deleted. We also sought to retain the same number of items in each subscale, and took content into account. Table 1 presents subscale reliabilities, most of which were good, and some of which were excellent. The exception was *Geminal*, which showed poor reliability for both Product 2 and Product 3.

### Factor Analysis

An orthogonal analysis was chosen because, theoretically, the three dimensions of the CPSS are independent. Thus, a principal components analysis (varimax rotation) was used to examine relationships among subscales for each product separately. For the cat pitcher, the analysis yielded two factors accounting for 46% and 21% of the total variance, with eigenvalues of 5.01 and 2.29, respectively. For the dog chair, the analysis yielded two similar factors, accounting for 39% and 24% of the total variance, with eigenvalues of 4.31 and 2.66. For the steer desk, two factors were also obtained, accounting for 42% and 29% of the total variance, with eigenvalues of 4.62 and 3.18. The factor loadings for each product are presented in Table 2.

### Comparisons Among Products

A multivariate repeated measures analysis of variance was performed, using the 11 subscale scores as dependent variables across the three products. Results showed that the multivariate main effect for product was significant,  $F_{(22, 172)} = 13.59, p < .001$ . Means, standard deviations, and results of the post

**Table 1**  
*Subscale Items and Reliabilities for Revised CPSS*

	Cat Pitcher	Dog Chair	Steer Desk	Final Version
<b>Novelty</b>				
Original	.84	.76	.82	.82
Over Used-Fresh Predictable-Novel Usual-Unusual *Unique-Ordinary *Original-Conventional				
Surprising	.87	.84	.87	.85
Stale-Startling Customary-Surprising *Astonishing-Commonplace *Shocking-Old Fashioned *Astounding-Common				
Geminal	.79	.56	.58	.73
Warmed Over-Trendsetting *Revolutionary-Average *Radical-Old Hat Uninfluential-Influential *Pioneering-Unprogressive				
<b>Resolution</b>				
Valuable	.82	.78	.84	.78
Worthless-Valuable *Important-Unimportant *Significant-Insignificant Inessential-Essential Unnecessary-Necessary				
Logical	.80	.80	.86	.79
Illogical-Logical *Makes Sense-Senseless Irrelevant-Relevant *Appropriate-Inappropriate *Adequate-Inadequate				
Useful	.82	.75	.82	.83
Ineffective-Effective *Functional-Nonfunctional *Operable-Inoperable Useless-Useful *Workable-Unworkable				

Table 1 (continued)

	Cat Pitcher	Dog Chair	Steer Desk	Final Version
<b>Elaboration and Synthesis</b>				
Organic	.68	.77	.70	.65
Disordered-Ordered				
*Arranged-Disarranged				
*Organized-Disorganized				
Formless-Formed				
Incomplete-Complete				
Elegant	.81	.88	.87	.79
*Graceful-Awkward				
Repelling-Charming				
Coarse-Elegant				
*Attractive-Unattractive				
*Refined-Busy				
Complex	.72	.65	.69	.71
*Intricate-Straightforward				
Simple-Complex				
Plain-Ornate				
*Complicated-Uncomplicated				
Boring-Interesting				
Understandable	.70	.64	.79	.67
*Meaningful-Meaningless				
Mystifying-Understandable				
*Intelligible-Unintelligible				
*Clear-Ambiguous				
Unexplained-Self-Explanatory				
Well-Crafted	.83	.84	.85	.82
*Skillful-Bungling				
*Well Made-Botched				
Crude-Well Crafted				
*Meticulous-Sloppy				
Careless-Careful				

Note: Reliabilities in the first three columns were obtained with an earlier version of the CPSS. Items marked with an asterisk are recoded before averaging to form a subscale score. When the scale is administered, items are randomly ordered.

hoc analyses (Tukey's) are presented in Table 3.

For the *Novelty* subscales, original,  $F_{(2, 386)} = 53.12, p < .001$ , surprising,  $F_{(2, 386)} = 90.97,$

Table 2  
Factor Loadings for Each Product (Pilot Study,  $N = 194$ )

	Pitcher		Chair		Desk	
	F1	F2	F1	F2	F1	F2
<b>Resolution</b>						
Logical	86	10	87	-08	90	-20
Useful	71	20	66	23	79	19
Valuable	60	34	81	-01	82	04
<b>Elaboration &amp; Synthesis</b>						
Understandable	81	-08	86	-13	82	-27
Well-Crafted	75	25	61	51	72	34
Organic	74	11	60	42	70	16
Elegant	74	20	72	14	86	-10
Complex	24	79	13	67	15	82
<b>Novelty</b>						
Surprising	09	92	-17	89	-19	90
Original	11	88	02	85	-06	89
Germinal	17	85	19	75	14	79

$p < .001$ , and germinal,  $F_{(2, 386)} = 74.66, p < .001$ , were all significant. Post hoc tests (Tukey's) showed that both the steer desk and the dog chair were rated significantly more original, surprising and germinal than the cat pitcher; they did not differ from each other.

For the *Resolution* subscales, valuable,  $F_{(2, 386)} = 16.92, p < .001$ , logical,  $F_{(2, 386)} = 16.62, p < .001$ , and useful,  $F_{(2, 386)} = 6.97, p < .002$ , were significant. The steer desk was rated as more valuable than the other two products, which did not differ from each other on this subscale. The cat pitcher and the steer desk did not differ on the logical subscale; both were rated as more logical than the dog chair. The steer desk was seen as more useful than the dog chair; the cat pitcher did not differ from the other two products in usefulness.

For *Elaboration and Synthesis*, significant differences among products occurred on or-

**Table 3**  
**Product Means and Standard Deviations, and Differences from *Poc Hoc* (Tukey's) Comparisons (Pilot Study,  $N = 194$ )**

	Cat Pitcher	Dog Chair	Steer Desk
<b>Novelty</b>			
<b>Original</b>			
M	3.16a	2.23b	2.23b
SD	1.31	0.93	1.12
<b>Surprising</b>			
M	3.85a	2.73b	2.52b
SD	1.29	1.07	1.19
<b>Geminal</b>			
M	4.29a	3.37b	3.16b
SD	1.28	1.04	1.08
<b>Resolution</b>			
<b>Valuable</b>			
M	4.99a	5.04a	4.49b
SD	1.24	1.26	1.38
<b>Logical</b>			
M	4.01a	4.59b	4.18a
SD	1.18	1.27	1.41
<b>Useful</b>			
M	3.25ab	3.40a	3.00b
SD	1.31	1.21	1.23
<b>Elaboration &amp; Synthesis</b>			
<b>Organic</b>			
M	3.02a	3.32b	3.01a
SD	0.91	1.09	1.00
<b>Elegant</b>			
M	3.77a	4.20b	4.35b
SD	1.27	1.67	1.50
<b>Complex</b>			
M	4.02a	3.57b	3.09c
SD	1.22	1.18	1.18
<b>Understandable</b>			
M	3.53a	4.04b	3.88b
SD	1.21	1.20	1.43
<b>Well-Crafted</b>			
M	3.28a	3.31a	3.01b
SD	1.09	1.18	1.18

Note: Scales range from 1 to 7, with low ratings indicating higher Novelty, Resolution, and so on. Means in the same row with the same subscript did not differ ( $p < .01$ ).

ganic ( $F_{(2, 386)} = 9.73, p < .001$ ), elegant ( $F_{(2, 386)} = 12.88, p < .001$ ), complex ( $F_{(2, 386)} = 39.68, p < .001$ ), understandable ( $F_{(2, 386)} = 9.98, p < .001$ ), and well-crafted ( $F_{(2, 386)} = 6.52, p < .003$ ). The cat pitcher was perceived as more elegant and more understandable than either of the other objects, which did not differ from each other (see Table 3). The steer desk was rated as more complex than the dog chair, which in turn was seen as more complex than the cat pitcher. For organic, the cat pitcher and steer desk were rated similarly; both were perceived as more organic than the dog chair. The steer desk was seen as more well-crafted than either of the other two products, which did not differ from each other.

## Discussion

As expected, the CPSS was clearly able to distinguish among products judged a priori to differ in *Novelty*. All three subscales of the *Novelty* dimension significantly differentiated among products chosen as least and most novel in this study. Although inconclusive, this finding supports the validity of these subscales.

A problem with the pilot study was that the products were judged by the authors to differ in *Novelty*, but these judgments may have been biased. Thus, before beginning the second study, judgments were obtained from uninvolved experts regarding their perceptions of the products.

A second decision was to change the type of product. The first reason for this decision was to counter the possibility that some of the products chosen for the pilot study were too unusual or beyond the everyday experience of our subjects. For example, the steer desk, a very unusual and costly

item, drew verbal responses of surprise and amusement from some of the subjects. The second reason for this change was to examine the generalizability of the CPSS. We chose keychains for Study 2 because they were familiar products, with no discernable sex or age differences in adult usage.

A third change was from a within-subjects design to a between-subjects design. One of the problems of the pilot study identified informally was that a certain amount of fatigue was involved in rating three products on many dimensions. Although counter-balancing "evened out" the effects of fatigue, the problem may not have been eliminated. Thus, in Study 2, each subject rated only one product (randomly assigned).

A fourth change was prompted by the likelihood that the intermediate version of the CPSS used in Study 1 contained response set or stylistic bias (e.g., Crano & Brewer, 1986). All the items were answered in the same direction (low ratings indicated creative). The response format of approximately half the items in the CPSS was changed to deal with this problem. In addition, to make the scores easier to interpret, the scoring direction was changed such that a high score was "higher" on each subscale.

The fifth change occurred because the low reliability of the Germinal subscale seemed to be a problem. We decided to expand the Germinal subscale by adding another item, namely, "pioneering-unprogressive." In addition, because the Elegant subscale contained only four items, "refined-busy" was added to it. Finally, we removed some redundancy in the CPSS by contrasting "intricate" with "straightforward," and "uncomplicated" with "complicated" on the Complex subscale.

## Study 2

### Phase 1

#### Method

**Subjects.** Sixteen subjects participated as judges in this phase of Study 2. Ten of the judges (8 males and 2 females) were members of the Marketing Services department of *The Buffalo News*, a large daily newspaper. The majority were graphic designers, and their average age was 33 years, 5 months. The remaining six subjects (five males and one female) were faculty members or graduate students in Creative Studies at State University College at Buffalo. Their average age was 33 years and 10 months.

**Materials and Procedure.** Subjects were provided with several keychains and asked to make global ratings (on 7-point scales) of the *Novelty*, *Resolution*, and *Elaboration and Synthesis* of each keychain. They were given a general definition of each term before they began rating the keychains. For *Novelty*, the definition was:

The extent of newness of the product: in terms of the number and extent of new processes, new techniques, new materials, new concepts included. The newness of the product both in and out of field; the effects of the product on future creative products.

For *Resolution*, the definition was:

Like the resolution of a musical chord, the degree to which the product fits or meets the needs of the problematic situation.

For *Elaboration and Synthesis*, the definition was:

The degree to which the product combines unlike elements into a refined, developed, coherent whole statement or unit.



Subjects were asked not to discuss their ratings with one another.

Four keychains were chosen for further examination because upon inspection of average ratings they represented high, moderate and low levels of *Novelty* and *Elaboration and Synthesis*. The first was a well-styled yellow plastic pen which snapped open and closed in a black metal frame. The second was a brown leatherette case (with an advertisement) containing a magnifying glass which rotated out of the case. The third was a simple two-layer silver metal ring in the shape of a heart. The fourth was a plain sturdy rubber band folded back upon itself.

## Results and Discussion

Means and standard deviations for the four keychains are presented in Table 4. Three separate 2 (Group: Graphic Design Professionals vs. Creative Studies Professionals) X 4 (Keychain) analyses of variance with repeated-measures on the second factor were performed, one for each dimension. For all three analyses, there was no significant effect of Group, nor was the Group X Keychain interaction significant. There was a significant main effect of Keychain for *Novelty* ( $F_{(3, 12)} = 42.60, p < .001$ ) and *Elaboration and Synthesis* ( $F_{(3, 12)} = 9.06, p < .002$ ). For *Resolution*, no significant effects were obtained. Means, standard deviations, and results of the post hoc tests are presented in Table 4.

From these results, it is clear that the groups of experts did not differ in their judgments. Both groups perceived differences in the *Novelty* and *Elaboration and Synthesis*, but not the *Resolution*, of the keychains.

**Table 4**  
Product Means, Standard Deviations (below means), and Differences

	Pen n=16	Magnifier n=16	Heart n=16	Rubber Band n=16
<b>Novelty</b>				
M	6.44a	4.56b	3.63bc	2.25c
SD	0.73	1.63	1.67	1.73
<b>Resolution</b>				
M	5.06a	3.75a	4.31a	3.50a
SD	1.84	1.77	1.58	1.90
<b>Elaboration &amp; Synthesis</b>				
M	6.13a	4.75ab	3.88bc	2.81c
SD	1.26	1.69	1.63	1.76

Note: Scales range from 1 to 7, with low ratings indicating higher Novelty, Resolution, and so on. Means in the same row with the same subscript do not differ ( $p < .05$ ).

## Phase 2

### Method

**Subjects.** There were four groups of subjects ( $N = 108$ ). The first was composed of 14 males and 15 females who volunteered during the Creative Problem Solving Institute in Buffalo, New York. Their average age was 48 years, 11 months (range 15-67 years). The second group was composed of 18 males and 1 female who attended a presentation (in English) on creative product analysis in Noordwijk aan Zee, The Netherlands. Their average age was 43 years and 5 months (range 26-64). The third group was composed of 7 male and 13 female undergraduate students enrolled in an Introductory Psychology course at the State University College at Buffalo. Their average age was 26 years and 1 month (range 17-44 years). The fourth group was composed of

23 males and 17 females enrolled in a business class at the State University College at Fredonia. Their average age was 23 years and 5 months (range 19-49 years).

**Materials and Procedure.** Four key-chains, described in Phase 1, were used: a rubber band, a heart-shaped metal ring, a magnifying glass, and a yellow pen. The 55-item CPSS (see Table 1) was administered. Subjects participated in groups in lecture halls. Subjects were randomly assigned to one of four groups by means of a number on the face of each booklet. Booklets contained informed consent statements which briefly described the study, along with the following instructions:

You will be describing a product or concept on a series of 7-point scales. On each of the scales, please indicate how you perceive the product or concept. There are no right or wrong answers--there are only personal opinions.

You will be using a computer answer sheet to make your ratings. Make your ratings according to how the scale describes the product or concept you are rating.

Here is an example: New 1--2--3--4--5--6--7 Old

If you think the product or concept is somewhat more old than new, you would mark a 5 on your computer answer sheet. If you think the product or concept is very new, you would mark a 1, and so forth. If you feel the product or concept is neither new or old, mark a 4.

Try to rate the product or concept on all scales. Please do not leave any blanks.

The booklet also contained the 55-item CPSS.

The numbered booklets were distributed in randomized blocks of four by a female experimenter. Subjects with the same number were asked to sit together in one corner of the hall. The experimenter, beginning with a randomly-chosen group, showed subjects the keychain they were to rate. Each member of the group was asked to examine

the keychain before beginning to rate it. Subjects were asked not to discuss their ratings with one another.

When all subjects had completed their booklets, they were debriefed. The experimenter explained the study, demonstrated the four keychains, and answered questions.

## Results

**Reliability.** Items in each subscale were recoded as necessary such that a high score indicated greater originality, elegance, and so on (see Table 1). The reliability (Cronbach's alpha) of each 5-item subscale was assessed. The reliabilities for the subscales were generally very good. Subscale scores were created for each subject by averaging the five items in the subscale (after recoding).

**Factor Analysis.** A principal components analysis with varimax rotation was used to examine the relationships among the subscales. Two factors were found, accounting for 39% and 25% of the total variance, with eigenvalues of 4.32 and 2.72, respectively. Factor loadings are presented in Table 5. As in Study 1, *Novelty* and *Resolution* subscales loaded on separate factors. Once more, *Elaboration* and *Synthesis* subscales did not form a separate factor, but loaded largely on the *Resolution* factor.

**Comparisons Among Products.** A one-way (4 Groups/Keychains) between-subjects multivariate analysis of variance was performed using the 11 subscale scores as dependent measures. The multivariate main effect was significant,  $F_{(33, 278)} = 3.71, p < .001$ . Examination of the univariate statistics

**Table 5**  
*Factor Loadings for Study 2*

	Factor 1	Factor 2
<b>Resolution</b>		
Logical	.84	.02
Useful	.82	-.06
Valuable	.68	.18
<b>Elaboration &amp; Synthesis</b>		
Organic	.77	.08
Understandable	.75	-.22
Well-Crafted	.69	.38
Elegant	.64	.43
Complex	.07	.71
<b>Novelty</b>		
Surprising	-.08	.88
Original	.02	.87
Germinal	.22	.83

showed that seven of the subscales were significant. The means and *F* values for these variables are presented in Table 6. Tukey's tests indicated that, for all three *Novelty* subscales, the pen was seen as more novel than the other three keychains, which did not differ from each other. For the four significant *Elaboration and Synthesis* subscales, a more complex pattern of results emerged.

## Discussion

**Reliability.** The subjects in Study 2 were from fairly heterogeneous groups, which contributes to the generalizability of these results. The reliabilities of the subscales were generally very good. They were similar to those of Study 1, and similar to those from earlier research (Besemer & O'Quin,

**Table 6**  
*Means, Standard Deviations, and Post Hoc Comparisons*

	Pen	Magnifier	Heart	Rubber Band	<i>F</i> (3,10)
<i>n</i>	28	28	26	26	
<b>Original</b>					
M	4.95a	3.81b	3.88b	3.55b	6.32**
SD	1.04	1.24	1.50	1.34	
<b>Surprising</b>					
M	4.25a	3.28b	3.18b	3.23b	5.35*
SD	1.05	1.17	1.15	1.27	
<b>Germinal</b>					
M	3.96a	2.92b	3.09b	2.88b	6.60**
SD	1.02	1.11	1.18	0.79	
<b>Organic</b>					
M	5.19ab	4.98bc	5.68a	4.40c	8.74**
SD	0.74	0.88	0.90	1.11	
<b>Elegant</b>					
M	4.46a	3.55b	4.91a	3.36b	15.85**
SD	0.78	0.94	1.04	1.04	
<b>Complex</b>					
M	3.55a	2.75b	2.43bc	1.83c	18.58**
SD	0.91	0.97	0.78	0.78	
<b>Well-Crafted</b>					
M	4.84a	4.06bc	4.76ab	3.38c	11.31**
SD	0.96	1.11	0.81	1.27	

Note: Means in the same row with the same subscripts did not differ (Tukey's test, \**p* < .01, \*\**p* < .001).

1986). The Organic and Understandable subscales yielded reliabilities that were somewhat low. Future research will be conducted to determine whether or not this is a function of the type of product being rated.

**Validity.** As expected, naive subjects perceived the keychains to differ in *Novelty*, as did the expert judges in Phase 1. Both groups perceived the pen to be more novel than the other keychains. One difference

between the pattern of results from experts and naive subjects was that the latter perceived the rubber band to be equal in *Novelty* to the heart and magnifier. The experts rated the rubber band as lower in *Novelty* than the heart and magnifier. This difference may reflect the fact that the experts rated several keychains, but the subjects in Phase 2 rated only one. In general, we believe that the validity of the *Novelty* subscales has been fairly well established.

The validity of the *Resolution* subscales was not adequately examined. Neither the expert judges nor the subjects in Phase 2 perceived significant differences among keychains on *Resolution*. All keychains were seen as equally able to perform the function for which they were created: bearing keys. In order to examine the validity of *Resolution* subscales, creative products which have been judged a priori to vary, with some adequately serving the function for which they were designed and some not, will need to be examined.

The validity of *Elaboration and Synthesis* received some support in Study 2. For both experts and naive subjects, the pen was higher in *Elaboration and Synthesis* than the rubber band. Experts rated the magnifier and the heart keychains as intermediate on the style dimension; with one exception (the heart tended to be viewed as more Elegant than the pen), naive subjects did also. However, the validity of the *Elaboration and Synthesis* subscales needs further research because it is difficult to isolate from *Novelty* and *Resolution*.

## General Discussion

These results are consistent with previous studies with regard to factor loadings of the

subscales (Besemer & O'Quin, 1986, 1987). Results of the factor analyses showed that *Novelty* and *Resolution* were perceived in both Study 1 and Study 2 as independent dimensions. These factor analytic findings support the theoretical underpinnings of the CPAM. As suggested by Besemer and Treffinger (1981), *Novelty* and *Resolution* ratings seem to be independently made.

*Elaboration and Synthesis* did not form a separate factor, but loaded largely with *Resolution* in both Study 1 and Study 2. In fact, the two studies used very different products, and yet the results of the factor analyses were remarkably consistent. The Complex subscale loaded with *Novelty* and the other four subscales loaded on *Resolution*.

Previous examinations of the CPSS also found that *Elaboration and Synthesis* subscales were not independent of *Novelty* or *Resolution* (Besemer & O'Quin, 1986, 1987). Besemer and Treffinger (1981) originally theorized that the three dimensions were independent. In actual testing, however, subjects have repeatedly not differentiated *Elaboration and Synthesis* from the other two dimensions. In addition, the *Elaboration and Synthesis* subscales sometimes migrate from loading mostly on *Novelty* (see Besemer & O'Quin, 1986) to loading mostly on *Resolution*, as in the present report. Future research will contribute to further understanding, and perhaps modification, of the CPAM. This theoretical approach (Besemer & Treffinger, 1981) is only a first step toward understanding the creative product.

## Practical Implications

There are several practical implications of the two studies reported here, especially for those with the goal of training or educating

people responsible for generating new products, or improving the creativity of products. By considering product attributes in a formalized way, areas for improvement and development reveal themselves. Traditional responses to products in development often focus on one dimension, or just a few, unintentionally overlooking the improvement to the product which might be made by looking at a more complete view of the product--through creative product analysis.

Once the techniques of creative product analysis are understood, product designers and engineers might be able to improve candidate products rather than eliminating them in initial screening. Singly, and in teams, product developers can refine existing plans before or after prototypes are tested. Engineers and new product designers might use the instrument in teams to develop a marketable idea. For example, Witt and Beorkrem (1989) measured the value of creative products on which scientists were working by having members of the organization's management evaluate each of the products on overall technical usefulness to end-users. Usefulness is one attribute measured by the CPSS; but other attributes might be found to be important either as criteria or predictors in such a setting.

Artists and artisans might work independently using the scale to refine works in progress. Students may be taught the principles of creative product analysis to improve the overall quality of their original works. They may also use the scale to improve a particular work of art, poem, or term paper. Even researchers might find the dimensions of the CPAM to be helpful. The *Creativity Research Journal*, for example, although not explicitly utilizing the CPAM, cites three major criteria to be used in

evaluating manuscripts submitted for publication: originality, utility, and integration (Runco, 1988). As Runco stated, "originality is vital, but must be balanced with fit and appropriateness" (p. 4).

#### REFERENCES

- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43, 997-1013.
- Barron, F., Gaines, R., Lee, D., & Marlow, C. (1973). Problems and pitfalls in the use of rating schemes to describe visual art. *Perceptual and Motor Skills*, 37, 523-530.
- 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* (pp. 341-357). Buffalo, NY: Bearly Limited.
- Besemer, S. P., & Treffinger, D. J. (1981). Analysis of creative products: Review and synthesis. *Journal of Creative Behavior*, 15, 158-178.
- Crano, W. D., & Brewer, M. B. (1986). *Principles and methods of social research*. Boston, MA: Allyn & Bacon.
- Ghiselin, B. (1958). Ultimate criteria for two levels of creativity. In C. W. Taylor (Ed.), *The 2nd University of Utah Research Conference on the identification of creative scientific talent* (pp. 141-155). Salt Lake City, UT: University of Utah.
- MacKinnon, D. W. (1978). *In search of human effectiveness: Identifying and developing creativity*. Buffalo, NY: Creative Education Foundation.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Parke, B. N., & Byrnes, P. (1984). Toward objectifying the measurement of creativity. *Roeper Review*, 6, 216-218.
- Runco, M. A. (1988). Creativity research: Originality, utility, and integration. *Creativity Research Journal*, 1, 1-8.
- Taylor, I. A., & Sandler, B. E. (1972). Use of a creative product inventory for evaluating products of chemists. *Proceedings of the 80th annual convention of the American Psychological Association*, 7 (part 1), pp. 311-312.
- Witt, L. A., & Beorkrem, M. N. (1989). Climate for creative productivity as a predictor of research usefulness and organizational effectiveness in an R&D organization. *Creativity Research Journal*, 2, 30-40.