

Measuring the Involvement Construct*

JUDITH LYNNE ZAICHKOWSKY**

A bipolar adjective scale, the Personal Involvement Inventory (PII), was developed to capture the concept of involvement for products. The scale successfully met standards for internal reliability, reliability over time, content validity, criterion-related validity, and construct validity. Tests of construct validity demonstrated that the scores were positively related to perceived differences among brands, brand preferences, interest in gathering information about the product category, and comparison of product attributes among brands.

Researchers of consumer behavior have historically developed a number of complex theories in the attempt to explain and predict the behavior of the consumer (e.g., Bettman 1979; Engel, Kollat, and Blackwell 1978; Howard and Sheth 1969). These theories propose that consumers actively search for and use information to make informed choices. This implies that the consumer is an intelligent, rational, thinking, and problem-solving organism, who stores and evaluates sensory inputs to make a reasoned decision (Markin and Narayana 1975).

However, a great deal of consumer behavior does not involve extensive search for information or a comprehensive evaluation of the choice alternatives, even for the purchase of major items (Olshavsky and Granbois 1979). The average consumer makes dozens of mundane decisions each day, few of which may be of importance. For such decisions, it may be inappropriate to assume an active information processor (Kassarjian 1978, 1981). This idea has led theorists to view consumer behavior in terms of a two-fold dichotomy: low involvement consumer behavior and high involvement consumer behavior (Engel and Blackwell 1982).

THE PROBLEM

Although researchers agree that the study of low versus high involvement states is interesting and important, there is currently little agreement about how to best define, and hence measure, the construct of involvement (Cohen 1983). The reasons for the diverse definitions and measures of involvement are perhaps due

to the different applications of the term "involvement." The literature suggests that a person can be involved with advertisements (Krugman 1962, 1965, 1967, 1977), with products (Howard and Sheth 1969; Hupfer and Gardner 1971), or with purchase decisions (Clarke and Belk 1978). Involvement with these different objects leads to different responses. For example, involvement with ads leads one to give more counterarguments to the ad (Wright 1974). Involvement with products has been hypothesized to lead to greater perception of attribute differences, perception of greater product importance, and greater commitment to brand choice (Howard and Sheth 1969). Involvement with purchases leads one to search for more information and spend more time searching for the right selection (Clarke and Belk 1978). Therefore, each area might have its own idiosyncratic result of the state of being involved with the object.

Researchers generally use the resulting behaviors as indicators of the level of involvement. Previous research has examined involvement with advertisements via a five-point scale that measures the degree of attention to the ad (Wright 1973, 1974). Involvement with products has been measured by several methods: rank-ordering products (Sheth and Venkatesen 1968), rating a series of products on an eight-point concentric scale as to their importance in the subject's life (Hupfer and Gardner 1971), asking how important it is to get a particular brand (Cohen and Goldberg 1970), or finding the total times that subjects report "don't know" for a series of brands (Ray 1973). On a broader level, involvement has been measured by administering Likert statements that were thought to tap the underlying concept—e.g., the product means a lot to me, it matters to me, or the product is important to me (Lastovicka and Gardner 1978a; Traylor 1981).

These diverse measures pose several problems for researchers. If conflicting results are obtained, we do not know if the discrepancy is due to different measures or to different behaviors. Second, many scales are single-item measures and may not capture the total involvement concept. Finally, single-item measures have low reliability, and current multiple-item measures have not

*This article was a finalist in the 1984 Robert Ferber Award for Consumer Research competition for the best interdisciplinary article based on a recent doctoral dissertation. The award is cosponsored by the Association for Consumer Research and the *Journal of Consumer Research*.

**Judith Lynne Zaichkowsky is Assistant Professor of Marketing, Faculty of Business Administration, Simon Fraser University, Burnaby, B.C., Canada V5A 1S6. The author wishes to thank Hal Kassarjian, who chaired the dissertation on which this article is based at the University of California, Los Angeles, and the Canada Council for financial support. A special thanks goes to The American University and Nanette Brown for help in word processing the manuscript.

been tested for internal reliability, stability, or validity. Hence a standardized, general, valid, and multiple-item measure of involvement should be useful.

BACKGROUND AND CRITERIA FOR MEASURING INVOLVEMENT

A measure of involvement—independent of the behavior that results from involvement—would allow the researcher to use the same measure across various research studies. This measure should also be sensitive to the proposed areas that affect a person's involvement level. These areas might be classified into three categories (Bloch and Richins 1983; Houston and Rothschild 1978):

1. Personal—inherent interests, values, or needs that motivate one toward the object
2. Physical—characteristics of the object that cause differentiation and increase interest
3. Situational—something that temporarily increases relevance or interest toward the object

In Houston and Rothschild's (1978) framework, different situations and different people are two factors that lead to various levels of involvement. Houston and Rothschild integrate physical characteristics of the product as part of the situational factor. Coinciding with Bloch and Richins (1983), the present article separates the physical from the situational and allows the same physical object to be subjected to different levels of involvement given different situations.

The evidence for the three factors—physical, personal, or situational—that influence the consumer's level of involvement or response to products, advertising, and purchase decisions is found in the literature. For example, Wright (1974) found that variation in the type of media—print versus audio—influenced the response given to the same message (physical). Lastovicka and Gardner (1978a) demonstrated that the same product has different involvement levels across people (personal), and Clarke and Belk (1978) demonstrated that different purchase situations for the same products cause differences in search and evaluation or raise the level of involvement (situational). Based on this prior reasoning, a measure of involvement might be developed that would pick up differences across people, objects, and situations.

Different types of scales were pretested before selecting a measurement approach that seemed to be generalizable across all product categories. First, a series of vignettes was developed to represent involvement. The vignettes were similar to scenarios found in Lastovicka and Gardner (1978b). Problems arose with developing enough generalizable scenarios for a reliable scale. Likert scale items proposed a problem because items that seemed to be appropriate for frequently purchased

goods did not seem to apply to durable goods and vice versa.

The most effective and generalizable type of scale appeared to be a semantic differential type (Osgood, Suci, and Tannenbaum 1957). The Semantic Differential consists of a series of bipolar items, each measured on a seven-point rating scale. It is easy to administer and score, takes only a few minutes to complete, and is applicable to a wide array of objects. The descriptors or phrases easily relate across product categories and can be appropriate to other domains, such as purchase decisions or advertisements. (However, the main focus of this article and scale development is involvement with products.) The steps taken to develop the measure were:

1. Define the construct to be measured.
2. Generate items that pertain to the construct.
3. Judge the content validity of generated items (item reduction).
4. Determine the internal reliability of items judged to have content validity (item reduction).
5. Determine the stability of internally reliable items over time (item reduction).
6. Measure the content validity of the 20 selected items as a whole.
7. Measure the criterion-related validity, which is the ability of the scale to discriminate among different products for the same people and different situations for the same product and same people.
8. Test the construct validity or theoretical value of the scale by gathering data and testing whether the scale discriminates on self-reported behavior.

DEFINING THE CONSTRUCT

This article will adopt the general view of involvement that focuses on personal relevance (Greenwald and Leavitt 1984; Krugman 1967; Mitchell 1979; Rothschild 1984). In the advertising domain, involvement is manipulated by making the ad "relevant:" the receiver is personally affected, and hence motivated, to respond to the ad (e.g., Petty and Cacciopo 1981). In product class research, the concern is with the relevance of the product to the needs and values of the consumer. In purchase decision research, the concern is that the decision is relevant, and hence that the consumer will be motivated to make a careful purchase decision (e.g., Clarke and Belk 1978). Although each is a different domain of research, in general, high involvement means personal relevance (Greenwald and Leavitt 1984).

In this study, the definition of involvement used for the purposes of scale development was:

A person's perceived relevance of the object based on inherent needs, values, and interests.

This definition recognized past definitions of involvement (e.g., Engel and Blackwell 1982; Krugman 1967;

Mitchell 1979). Judging from previous writings, this definition may be applied to advertisements, products, or purchase decisions. Early work by Krugman (1962, 1967) in advertising focused on personal connections. Wright (1974) defined involvement with advertising as the receiver's perception of the relevancy of the ad content to some pending problem. In the area of product class involvement, Howard and Sheth (1969) used the terms importance of purchase and involvement interchangeably; they defined involvement in terms of a person's needs or values. Hupfer and Gardner (1971) defined involvement as a general level of interest in or concern about an issue without reference to a specific position. Finally, Houston and Rothschild (1978) referred to response involvement and defined it as a function of enduring involvement or a need derived from a value in the individual's hierarchy of needs.

ITEM GENERATION AND CONTENT VALIDITY

A semantic differential scale was to be developed based on the earlier definition of involvement. Thus, a list of 168 word pairs was generated to represent this concept of involvement. Examples of those pairs are important-unimportant, interested-uninterested, and exciting-unexciting. The first step was to judge the proposed items for content validity—how well the chosen items represent the defined concept. Content validity of the 168 word pairs was tested in two phases: (1) initial deletion of poor word pairs, and (2) finer judging of the more appropriate word pairs.

Three expert judges (senior Ph.D. candidates in consumer behavior) were given this study's definition of involvement and instructed to rate the 168 word pairs three times: first, replacing the word "object" in the definition with "product;" second, replacing the word "object" with "advertisement;" and third, replacing the word "object" with "purchase decision." Each word pair was rated on the following scale: (1) clearly representative of involvement, (2) somewhat representative of involvement, and (3) not representative of involvement. Word pairs that were not rated as representative of involvement for any advertisement, purchase decision, or product were dropped. Examples of deleted word pairs are adequate-inadequate, controversial-noncontroversial, and naive-sophisticated.

Word pairs that were dropped included traditional measures of attitudes used in the psychology and marketing literature. Word pairs such as good-bad, pleasant-unpleasant, nice-awful, and like-dislike (e.g., Loken 1984; Mitchell and Olson 1981) were judged to be unrepresentative of involvement. The judges decided that other word pairs, such as valuable-worthless and appealing-unappealing would remain, as they seemed to measure involvement. Items at the low end of the bipolar scale that represent the low end of involvement

were generally not negative—as they would be if measuring attitudes—but rather were "who cares" descriptors, e.g., unimportant, unexciting, doesn't matter, or of no concern.

Five new judges then rated the remaining 43 word pairs using the same procedure. Only 23 items were consistently rated as representing the involvement construct (80 percent agreement over products, purchase decisions, and advertisements for each word pair). This meant that at least 12 of the possible 15 judgments for each word pair (five judges over three objects) had to be rated as representative of the involvement construct. Agreement across judges and within each area for the 23 word pairs was as follows: advertisements, 84 percent; products, 87 percent; and purchase decisions, 77 percent.

Twenty-three was assumed to be too low a number of items with which to start data collection (French and Michael 1966; Nunnally 1978). Thus, seven additional items were added to the item pool to raise the initial number to 30 (five of these seven were eventually dropped). For example, trivial-grand (45 percent agreement) was changed to trivial-fundamental, and inspiring-discouraging (55 percent agreement) was changed to inspiring-uninspiring and returned to the list. Therefore, a thirty-item scale emerged from the content validity phase that trained and knowledgeable judges agreed measured involvement over three domains: products, advertisements, and purchase decisions. However, this study focused on, and further validation procedures were carried out on, involvement with products.

INTERNAL SCALE RELIABILITY

The next task was to administer the 30 items as a scale over different product categories to measure the internal consistency or inter-item correlation. Two product classes—watches and athletic shoes—were selected because they were thought to be used by the subjects. One hundred and fifty-two undergraduate psychology students completed the scale during class time. Approximately half of the subjects filled out the scale pertaining to athletic shoes and the other half filled out the scale pertaining to watches. The results show that for both product categories, 26 bipolar items had an item-to-total score correlation of 0.50 or more, and a Cronbach alpha level of 0.95.

Six adjective pairs with relatively low item-to-total correlations were dropped; interestingly, most of these adjective pairs had been returned earlier to the item pool. Factor analyses, using varimax rotation with squared multiple correlations in the diagonals for factor extraction, were carried out over both products to check if the items selected for deletion loaded onto one particular dimension or were amorphous across factors. For both watches and athletic shoes, one factor ex-

plained the major variation in the data, accounting for 70.3 percent and 69.3 percent of the common variance, respectively (eigenvalues 13.3 and 13.2). Watches had two more factors, accounting for 11.6 percent and 5.6 percent of the common variance (eigenvalues 2.2 and 1.1), and athletic shoes had three more factors, accounting for 11.7 percent, 5.9 percent, and 5.7 percent of the common variance (eigenvalues 2.2, 1.2, and 1.1).

The results of the factor analyses showed that the items selected for deletion did not load together on any unique factor across either product category. Since the first factor accounts for approximately 70 percent of the variance, and none of the remaining items had a loading of zero or less on that first dimension, the scale development continued on the assumption of a simple linear combination of the individual items (Comrey 1973). The assumption is that no individual item is sufficient, and that it is the scale taken as a whole that tends to measure the involvement construct (Nunnally 1978).

TEST-RETEST RELIABILITY

Test-retest reliability of the remaining 24 items was examined over two new subject samples and four new product categories. Sixty-eight psychology students initially rated calculators and mouthwash. Forty-five MBA students rated breakfast cereals and red wine. The order of the products was counterbalanced—half of the subjects in each group rated one product category first, and the other half rated the other product category first. The scales were administered during class time and took about five minutes to complete.

Three weeks later the scales were administered over the same product categories to the same subjects. Thirteen psychology students and 19 MBA subjects were lost to attrition; thus, 55 psychology students and 26 MBA students were used to measure test-retest reliability. The average Pearson correlation between Time 1 and Time 2 on the 24 items was 0.90. Individual item-to-item correlations ranged from 0.31 to 0.93. Four additional items with average test-retest correlations below 0.60 were deleted. The resulting twenty-item involvement score test-retest correlations for each product were as follows: calculators, $r = 0.88$; mouthwash, $r = 0.89$; breakfast cereals, $r = 0.88$; and red wine, $r = 0.93$. These product categories were also tested for internal scale reliability. The Cronbach alpha ranged from 0.95 to 0.97 over the four products.

Therefore, a twenty-item scale emerged from the internal reliability and stability phases of scale development for products. Twenty items allowed an adequate sampling of the possible items that represent involvement with products and yet was long enough to ensure a high level of reliability.¹ On a practical level, the scale

¹Although the current analyses do not suggest what the reliability is for subsets of the scale items, the case may be that a smaller number

fits neatly on one page and only takes a few moments to complete. The scale was then counterbalanced so that ten random items were reverse scored. Since each bipolar item was rated on a seven-point scale, the total possible score ranged from a low of 20 to a high of 140. The scale was named the Personal Involvement Inventory (PII) and is listed in Appendix A.

SECOND CONTENT VALIDITY

A second measure of content validity was obtained from the open-ended responses of 45 MBA students over three product categories: 35mm cameras, red wine, and breakfast cereals. After completing the scales for each product, subjects answered the following open-ended question:

Now we would like you to state, in your own words, *why* you rated each product category as you did.

Subjects were then divided into three groups—high, medium, or low—for each product class according to their scale scores.² Examples of the open-ended responses appear in the Exhibit.

Two expert judges (senior Ph.D. candidates in consumer behavior) blind to the scale scores evaluated the total set of open-ended responses. For each product category, the judges sorted the comments into three groups indicative of low involvement, medium involvement, and high involvement with the product category, based on how well the responses represented involvement, as defined earlier.

Interjudge reliability on the classification of the responses was 80 percent agreement for 35mm cameras, 84 percent agreement for red wines, and 80 percent agreement for breakfast cereals. Classifications on which the two expert judges did not agree were then given to

of items would be almost as reliable as the 20 items. The problem of reducing the scale to fewer items lies in deciding which items to select as subsets, since individual items differed in their reliability across product categories. A subset of items that may approach the reliability of the 20 items for one product may not approach the same reliability for another product. This variation is evident in that the test-retest total score correlation ranged from 0.88 to 0.93 over products, and test-retest for the 20 individual items ranged from 0.44 to 0.93 over various products. The twenty-item measure should outperform any subset of the scale; besides, decreasing the number of items would not really make the scale any easier to administer, but may serve to decrease the domain of items judged as being representative of involvement and also lower the reliability of the scale. Researchers who may use this scale are warned not to haphazardly reduce the number of items.

²The classification of subjects into low, medium, and high scores was based on an overall distribution developed over 13 product categories (Table 3) and several hundred subjects. All scores were tabulated on the PII scale range presented in the Figure. Subjects whose PII scores fell into the bottom 25 percent of the overall distribution were classified as having low involvement with the product. Subjects whose PII scores fell into the middle 50 percent of the distribution were classified as having medium involvement, and subjects whose PII scores were in the top 25 percent of the distribution were classified as having high involvement with the product. For development of this classification scheme see Appendix B.

EXHIBIT
OPEN-ENDED RESPONSES ON CONTENT VALIDITY

35mm Cameras

1. High involvement for cameras (score greater than 110).
 - a. Subject 1. Cameras are important, but not essential. They provide a creative and historical outlet for me.
 - b. Subject 12. Cameras interest me and are an important hobby to me.
2. Low involvement for cameras (scores less than 70).
 - a. Subject 17. Because I never use 35mm cameras and am not extremely interested in them.
 - b. Subject 37. It's a nice product to have but not a high priority. I have several but as I recall, none of the purchases was an "involved" purchase.

Red Wine

1. High involvement for red wine (score greater than 110).
 - a. Subject 22. Red wine adds a lot to the appropriate meals.
 - b. Subject 6. I have always wanted to know more about wines and enjoy it when people I know teach me about them.
2. Low involvement for red wine (score less than 70).
 - a. Subject 20. I'm not interested in wines nor do I particularly appreciate the mystique that surrounds wines, in general.
 - b. Subject 36. OK for socials and getting drunk.

Breakfast cereals

1. High involvement for breakfast cereals (score greater than 110).
 - a. Subject 27. I eat cereal, healthy efficient "wake up America." Cereal is good for you.
 - b. Subject 8. Because they are diet foods.
2. Low involvement for breakfast cereals (score less than 70).
 - a. Subject 3. I think breakfast cereals are a sham. I only eat grape-nuts. It infuriates me to see breakfast cereals advertised to be eaten with toast, juice, etc. What's the use, jaw exercise? I refuse to buy cereal for my child.
 - b. Subject 31. I eat cereal for convenience; it is easy and fast. I have no interest in them nor am I fascinated with them.

a third judge to classify. The categories of responses, as grouped by the scale scores, were compared to the categories of responses as grouped by the expert judges. The results are presented in Table 1. These data indicate a significant relationship between the scale scores and the open-ended responses from the subjects, thus adding an additional modicum of support to the validity of the scale.

CRITERION-RELATED VALIDITY

Criterion-related validity is demonstrated by comparing the scores from the developed instrument with one or more external variables that provide a direct measure of the characteristic in question (French and Michael 1966). The external variable selected as a criterion was the simple ordering or classification of products into low or high involvement categories.

Twenty-one products classified in other studies as

TABLE 1
RELATIONSHIP BETWEEN THE SCALE SCORES AND THE OPEN-ENDED RESPONSES

Scale scores	Judges' ratings				Collapsed for Chi-square		
	Low	Medium	High	(Total)	Low	Medium	High
35 mm Cameras^a							
Low	7	1	0	(8)	11	13	7
Medium	4	12	7	(23)	0	4	10
High	0	4	10	(14)			
(Total)	(11)	(17)	(17)	(45)			
Red wine^b							
Low	12	1	0	(13)	12	1	0
Medium	8	9	8	(25)	8	10	14
High	0	1	6	(7)			
(Total)	(20)	(11)	(14)	(45)			
Breakfast cereals^c							
Low	19	3	0	(22)	19	3	0
Medium	9	9	1	(19)	9	11	3
High	0	2	2	(4)			
(Total)	(28)	(14)	(3)	(45)			

^a $\chi^2 = 10.4, df = 2, p < 0.01.$

^b $\chi^2 = 17.0, df = 2, p < 0.001.$

^c $\chi^2 = 11.2, df = 2, p < 0.01.$

NOTE: As more than 20 percent of the expected cell frequencies dropped below 5, either the low or high row was collapsed into the medium row to compute the statistic.

having either high or low involvement (Bowen and Chaffee 1974; Hupfer and Gardner 1971; Lastovicka and Gardner 1978a; Traylor 1981) were presented to a group of 68 undergraduate psychology students. As in Hupfer and Gardner (1971), subjects rated each product on an eight-point scale: extremely unimportant in my life (1) to extremely important in my life (8).

From these 21 products, four were selected for measurement: bubble bath (mean \bar{X} = 2.35); facial tissue (\bar{X} = 5.25); jeans (\bar{X} = 6.6); and automobiles (\bar{X} = 7.9). Bubble bath was previously selected as a low involvement product, and jeans a high involvement product by Clarke and Belk (1978). Facial tissue was previously identified as a low involvement product, and automobiles as a high involvement product by Lastovicka and Gardner (1978a).

The twenty-item involvement scale (PII) was administered for each of the four product categories to a fresh sample of 47 undergraduate psychology students during class time. The PII mean scores and standard deviations for each product were as follows: bubble bath \bar{X} = 69, s = 38 (Males \bar{X} = 55, Females \bar{X} = 74); facial tissue \bar{X} = 87, s = 26; jeans \bar{X} = 99, s = 21; and automobiles \bar{X} = 122, s = 19.

A repeated measures analyses of variance showed an overall significant difference among the product means

($F(3, 138) = 39.9, p < 0.001$). Furthermore, each mean was found to be significantly different from each of the others ($p < 0.01$). These results are in agreement with previous studies that have stated that facial tissue and bubble bath have lower involvement levels than jeans and automobiles.

CONSTRUCT VALIDITY

Studies of construct validity check the theory underlying the test (French and Michael 1966). Three steps are involved in construct validity. First, from the involvement literature, propositions are made about the behavior of people with high and low scores. Second, data is gathered to test if the scale discriminates on behavior, and third, an inference is made as to whether the theory is adequate to explain the data collected.

Theoretical Propositions of Involvement

Various propositions about differences in low and high involvement behavior were selected after reviewing theoretical papers by several authors (e.g., Belk 1981; Bowen and Chaffee 1974; Lastovicka 1979; Lastovicka and Gardner 1978b; Mitchell 1979; Robertson 1976; Tyebjee 1979). Generally, there seems to be some agreement on what constitutes the differences between having high or low involvement in a product class. Under the low involvement condition, researchers propose:

1. A relative lack of active information seeking about brands
2. Little comparison among product attributes
3. Perception of similarity among different brands
4. No special preference for a particular brand

Based on these various theoretical propositions, the following specific statements were developed and then administered to subjects with the PII over various product categories:

1. I would be interested in reading information about how the product is made.
2. I would be interested in reading the *Consumer Reports* article about this product category.
3. I have compared product characteristics among brands of this product.
4. I think there are a great deal of differences among brands of this product.
5. I have a most preferred brand of this product.

These statements were rated on a seven-point scale from strongly disagree (1) to strongly agree (7).

Method. The PII and the specific questions related to involvement were administered over three products to 28 clerical and 29 administrative staff members (7

males and 50 females) at a major university. Subjects were personally contacted at work by the researcher and asked if they would participate. Those who agreed were given a questionnaire and asked to complete it at their office desk. The researcher then returned in about an hour to collect the questionnaires from the subjects. Seven additional subjects agreed to fill out the questionnaire but never completed it. The median age range of the subjects was 35–44 years and the median education level was some college. The products selected for evaluation were instant coffee, laundry detergent, and color television.³ These were chosen to represent a range of products thought to be used by the subjects.

Subjects were classified into three groups to compare the responses between subjects who had low involvement with the product category and those who had high involvement with the product category. This classification scheme is the same as that found in the second content validity section and further explained in Appendix B. The particular question of interest was: "Did subjects having low PII scores for the product category respond to the five statements differently than did those having high PII scores for the product category?" Planned comparisons, by simple t-tests, were carried out between the low and high PII scores for each statement and product category. Before comparing the differences between the low and high group in their responses to the statements within each product category, a one-way MANOVA was computed over the five statements for each product category to determine if the overall pattern of responses across the five propositions was significant.

In addition, the Pearson correlations between the scale score (range = 120) and the responses to the statements (range = 7) were computed. The cell means, correlations, and results of t-tests between the low and high cells are presented and summarized in Table 2.

Results. The Cronbach alpha, the mean, and the standard deviation of the PII scores for the three product categories were as follows:

	Cronbach alpha	Mean (<i>s</i>)
Instant coffee	.97	66 (40)
Color television	.99	97 (30)
Laundry detergent	.97	103 (23)

These scores provide two unexpected results: first, a relatively high PII score for laundry detergent (103), and second, a relatively low PII score for color television (97). These results should be interpreted in the context of the sample population used in this study: this relatively homogeneous group of middle-aged females may have viewed laundry detergent as more involving than

³PII scales for the product category of red wine and two purchase situations for red wine were also included for other purposes.

TABLE 2
RELATIONSHIP BETWEEN CONSTRUCT VALIDITY STATEMENTS AND LOW, MEDIUM, OR HIGH PII SCORES:
MEANS, STANDARD DEVIATIONS, AND CORRELATIONS

Construct validity statement ^a	Instant coffee			<i>r</i> ^b	Laundry detergent			<i>r</i>	Color television			<i>r</i>
	Low (32)	Medium (12)	High (12)		Low (4)	Medium (28)	High (25)		Low (9)	Medium (26)	High (12)	
1. I would be interested in reading information about how the product is made.	3.28 (2.0)	4.42 (2.3)	4.25 (2.3)	.30 ^c	1.25 [*] (.5)	4.04 (1.7)	4.48 (2.4)	.37 ^c	3.56 (2.1)	4.00 (1.9)	4.23 (2.1)	.14
2. I would be interested in reading the <i>Consumer Reports</i> article about this product.	3.00 [*] (1.8)	4.75 (2.3)	4.92 (2.3)	.47 ^c	2.75 [†] (2.9)	4.46 (2.0)	5.00 (2.1)	.33 ^c	4.56 (1.9)	4.65 (1.9)	5.36 (2.1)	.27 ^c
3. I have compared product characteristics among brands.	2.59 [*] (1.8)	3.42 (2.1)	5.25 (2.0)	.52 ^c	1.75 (1.5)	4.36 (1.8)	4.80 (2.4)	.42 ^c	3.11 [*] (1.9)	3.85 (1.9)	4.59 (2.3)	.23 ^d
4. I think there are a great deal of differences among brands.	3.94 [*] (1.6)	4.67 (1.1)	6.33 (.8)	.63 ^c	2.25 [*] (1.0)	4.00 (1.7)	5.20 (2.1)	.42 ^c	4.11 [*] (1.2)	4.85 (1.5)	5.73 (1.7)	.33 ^c
5. I have a most-preferred brand of this product.	2.88 [*] (1.9)	4.83 (1.8)	6.17 (1.7)	.68 ^c	2.50 [*] (3.0)	4.68 (1.6)	5.44 (1.9)	.42 ^c	2.56 [*] (1.4)	4.77 (1.7)	5.55 (1.9)	.50 ^c

^a The construct validity statements are measured on a seven-point scale: (1) strongly disagree to (7) strongly agree.

^b *r* = Pearson Correlation between PII score and response to construct validity question

^c *p* < 0.01.

^d *p* < 0.05.

^{*} Low scores significantly different than high scores *p* < 0.01.

[†] Low scores significantly different than high scores *p* < 0.05.

NOTE: Numbers in parentheses in Table heading are numbers of subjects in each group. Numbers in parentheses in Table body are standard deviations.

color television because it may be their responsibility to do the family laundry. If this is true, they would value the product's benefits and they would be likely to be interested in the quality of the product because they need the product to perform their household duties. Televisions, however, may not fall under their responsibility for maintenance or interest them as much. Electronics, solid state, or color tuning may not be relevant to them. And if a television does not affect them personally, housewives might have relatively low involvement with this product.

The results of the MANOVA for all five statements were significant for the three products (instant coffee $F(10, 98) = 6.56, p < 0.001$; laundry detergent $F(10, 100) = 2.34, p < 0.05$; and color television $F(10, 100) = 2.00, p < 0.05$). This indicated that there were significant differences due to the PII scores on the responses to all five behavioral statements pertaining to involvement. These overall significant results allow the interpretation of each proposition separately.

Search for product information. High involvement consumers should be more interested in acquiring information about the product than low involvement consumers. For example, Engel and Blackwell (1982) defined involvement as the activation of extended problem solving behavior, and Bettman (1979) cited

level of involvement as a mediating variable in information search. Thus, high scale scorers should indicate more interest in product information than low scorers. Subjects were given two statements pertaining to information search over the three product categories. The first statement was "I would be interested in reading information about how the product is made (instant coffee and laundry detergent) or works (color television)." The second statement was "I would be interested in reading the *Consumer Reports* article about _____."

The results of the analyses of the information search questions are generally in agreement with the theory of involvement. High scorers tended to be more interested in information pertaining to the product than low scorers. All correlation coefficients were significantly different from zero, with the exception of the product category of televisions. Perhaps the change in the wording of the question led to the weak results for that product category. Perhaps interest in how televisions work can indicate interest in technology or interest in quality of performance other than interest in the product per se.

Alternative evaluation. One of the characteristics of high involvement is the evaluation of competing alternatives. Since the highly involved consumer searches for relevant information, the available alternatives are thought to be consciously compared before a selection

is made. To tap this dimension, subjects were asked the extent to which they agreed with the statement "I have compared product characteristics among brands of _____." For all products, the high scorers had significantly greater agreement with the statement than low scorers.

Perception of brand differences. The next proposition tested was that high involvement scorers would perceive greater differences among brands in the product class than low involvement scorers. This proposition stems from writings of Robertson (1976), who suggests that high involvement implies that beliefs about product attributes are strongly held, whereas low involvement individuals do not hold strong beliefs about product attributes. Thus, the strength of the belief system to the attributes emphasizes the perception of differences among brands on the attributes where beliefs are strongly held. Subjects were asked to respond to the statement "I think there are a great deal of differences among brands of _____." High scorers always perceived greater differences ($p < 0.01$) among brands than low scorers in the product class.

Brand preferences. People highly involved in a product class were hypothesized to have a most preferred brand in the product category. The preference of a particular brand stems from the perception of differences among brands. Since high involvement implies perceiving greater differences about product attributes, then the consumer should have a greater preference based on that product differentiation. Again, over all three products, high scorers showed a significantly ($p < 0.01$) greater agreement with the statement "I have a most preferred brand of _____" than low scorers.

In conclusion, the various measures of construct validity used the correlation of two paper and pencil tests on the same subjects as evidence that the proposed scale does tap the construct of involvement, as applied to product categories. Although no one result is an excellent test of the scale, each finding adds to the weight of evidence that the scale is an acceptable measure of involvement, as applied to product categories.

FACTOR ANALYSES OF THE PII

An investigation of the dimensionality of the twenty-item scale was carried out for each product category used in the scale development. The items were factor analyzed using varimax rotation with squared multiple correlations in the diagonal for factor extraction. The general pattern of results showed one main factor and (usually) one minor or residual factor for every product category. The major factor accounted for a range of common variance from 65 percent for jeans to 100 percent for instant coffee. Over all products, all items loaded positively on the first factor, which indicates that the assumption of a simple linear combination of the scale items was not violated.

SENSITIVITY TO SITUATIONAL DIFFERENCES

The second content validity, the criterion validity, and the construct validity sections have demonstrated that the level of involvement with product categories varies greatly over individuals. For any product category, there seems to be individuals who have low involvement with the product and individuals who have high involvement with the product. Additionally, the average level of involvement varies across the different products. For example, students rated bubble bath 69 on the PII and rated automobiles 122 on the PII. This demonstrates that different products are perceived differently by the same people. The scale is also proposed to be sensitive to different situations, a third factor that causes involvement, given the same people and the same products.

Previous studies by Clarke and Belk (1978) and Belk (1981) demonstrated that some purchase situations can be more involving than others. They found that the purchase of some previously uninvolved products for gifts can raise the level of involvement in the purchase decision. To investigate the possibility of rating purchase situations on the scale, the PII was administered over two purchase situations for wine to 41 members of the clerical and administrative staff used in the previous construct validity study.⁴ Each subject rated two purchase situations: (1) the purchase of a bottle of wine for everyday consumption, and (2) the purchase of a bottle of wine for a special dinner party. The scale items were internally reliable for these purchase decisions. Cronbach alphas were 0.98 and 0.97 respectively, and the item-to-total correlations were generally above 0.50.

For this data collection, these situations were counterbalanced across subjects. The mean scale score for the everyday consumption was 78 ($s = 34$), and for the special dinner party was 106 ($s = 24$). A related measures t-test was significant at $t(40) = 5.42$, $p < 0.001$; therefore, the two purchase situations were rated differently on the PII. The analysis for differences between the two purchase situations was also carried out as a between-subjects design. Twenty-two subjects first rated the everyday consumption ($\bar{X} = 76$) and 19 subjects first rated the special dinner party ($\bar{X} = 98$). The between-subjects t-test for the first rating was also significant at $t(39) = 2.34$, $p < 0.05$. These results show that the PII is sensitive to different situations, if people and product remain constant. The PII may hold promise as a measure of involvement with purchase decisions, perhaps even applicable as a manipulation check for experiments that deal with manipulation of the situation as manipulation of involvement level (e.g., Park and Young 1983; Petty and Cacioppo 1979, 1981; Petty, Cacioppo, and Goldman 1981).

⁴The other 16 subjects did not receive these scales as part of their questionnaire.

SUMMARY

The purpose of the study was to develop a scale to measure the construct of involvement. Hence, a semantic differential scale was developed to capture the concept of involvement for products. This Personal Involvement Inventory was developed over four data sets of 268 undergraduate psychology students; two data sets with 49 MBA students; and two data sets with 57 clerical and administrative staff members. The scale was demonstrated to have content validity by expert judges at two phases of the scale development: first, for the selection of items, and second, through classification of open-ended responses from subjects. The reliability or stability of the scale over time was checked over two subject populations for an average test-retest correlation of 0.90. The criterion-related validity of the scale was checked by demonstrating agreement with the order of various products as found in previous studies. The construct validity—the test of the scale to theoretical propositions—was then carried out. The scale was administered to clerical and administrative staff and covered three different product categories and several statements of behavior proposed to be representative of involvement. Over all three product categories there was a positive relationship between the scale scores and the subjects' responses to the statements of theoretical propositions pertaining to involvement.

Limitations

Missing from this scale development are tests of convergent and discriminant validity. The tests of convergent validity with another measure of involvement were not carried out because at the time of this scale development no other general involvement measure in the literature had been tested for reliability and validity.⁵ Tests of discriminant validity to the concept of expertise or knowledge structure were carried out and are reported in Zaichkowsky (1985). PII scores were found to be unrelated to expertise but related to product use. Further tests of discriminant and convergent validity need to be carried out with respect to other constructs. In particular, the relationship of PII scores to attitudes should be further examined, since several items on the scale appear to be similar to a measure of attitudes.

Although the usefulness of the PII was demonstrated for products, the initial aim was to select items so that the same scale might also be applied to advertisements or purchase decisions. Some data were collected over different purchase decisions and showed that the PII was internally reliable for different purchase situations for the same product. Additionally, the purchase situations differed in their involvement scores, as would be expected; the special dinner party scored higher on the

PII than the everyday purchase situation. Further, some preliminary research indicates that the PII is an internally reliable measure when applied to advertisements. However, more research needs to be carried out to verify the stability and construct validity of the PII to advertisements and purchase decisions.

The PII should have several benefits to the study of consumer behavior. It offers the potential of a valid instrument to replace the ad hoc and untested approaches that have previously been used in the field. Since involvement is proposed to be a variable in the decision process, the PII offers researchers a quickly administered tool, generalizable across product categories, that can be used as a covariate to other research questions. The ultimate test of the scale is whether or not the instrument can be used in empirical studies to test various aspects of involvement. I am conducting such research, and as others use the instrument and generalizable norms develop, its true validity will be determined.

APPENDIX A

Personal Involvement Inventory

The following Personal Involvement Inventory is designed to measure a person's involvement with products. To change the instructions to measure involvement with advertisements or purchase decisions, the words in the parentheses should be changed accordingly. To measure involvement with advertisements, the words "various products they regularly purchase or have purchased in the past" would be changed to "the advertisements you have just seen (read)." To measure involvement with purchase decisions, the words "various purchase decisions people make" would be substituted.

The name of the object to be judged should be inserted at the top of the scale page. Examples for three applications of the different contexts of the object are: (1) if the product was the object, then "red wine" would be judged; (2) if an ad was the object, then "the ad for Gallo wine" would be judged; and (3) if the purchase decision was the object, then the "purchase of a bottle of wine for a special dinner party" would be judged. The reader is reminded that the construct validity of the scale has only been supported for products.

Instruction Page

Instructions

The purpose of this study is to measure a person's involvement or interest in (various products they regularly purchase or have purchased in the past). To take this measure, we need you to judge various (products) against a series of descriptive scales according to how YOU perceive the product you will be shown. Here is how you are to use these scales:

If you feel that the (product) that appears at the top of the page is *very closely related* to one end of the scale, you should place your check mark as follows:

⁵There is currently some research translating this scale to French for possible validation to an independently developed Likert scale.

TABLE 3
PRODUCTS USED FOR SCALE DISTRIBUTION

Product	Mean PII score	(s)	Sample (N)
Instant coffee	66	(40)	57 ^a
Bubble bath	69	(38)	45 ^b
Breakfast cereal	69	(29)	43 ^c
Mouthwash	74	(28)	68 ^b
Red wine	82	(31)	45 ^c
Red wine	84	(31)	57 ^a
Facial tissues	87	(26)	47 ^b
Headache remedy	91	(25)	68 ^b
35mm camera	96	(26)	45 ^c
Color TV	97	(30)	57 ^a
Jeans	99	(21)	47 ^b
Laundry detergent	103	(23)	57 ^a
Calculator	112	(16)	68 ^b
Automobile	122	(19)	47 ^b

^a Clerical and administrative staff.

^b Undergraduate psychology students.

^c MBA students.

NOTE: Theoretical mean = 80. Actual mean based on above products = 90 (s = 32).

defined as those scoring in the middle 50 percent of the distribution; they had scores ranging from 70 to 110. High scorers were defined as those scoring in the top quartile of the distribution; they had scores ranging from 111 to 140. The distribution was also used for the development of a two-group classification when necessary—i.e., those relatively high scores and those relatively low scores. The scale mean (89.55) was used as the break point for this grouping. By using this overall distribution as a guide to classification, some comparison is possible between low and high scoring subjects on the PII.

[Received March 1984. Revised June 1985.]

REFERENCES

- Belk, Russell W. (1981), "Effects of Gift-Giving Involvement on Gift Selection Strategies," in *Advances in Consumer Research*, Vol. 9, ed. Andrew Mitchell, Ann Arbor, MI: Association for Consumer Research, 408-411.
- Bettman, James R. (1979), *An Information Processing Theory Of Consumer Choice*, Reading, MA: Addison-Wesley.
- Bloch, Peter H. and Marsha L. Richins (1983), "A Theoretical Model for the Study of Product Importance Perceptions," *Journal of Marketing*, 47 (Summer), 69-81.
- Bowen, Lawrence and Steven H. Chaffee (1974), "Product Involvement and Pertinent Advertising Appeals," *Journalism Quarterly*, 51 (Winter), 613-621.
- Clarke, Keith and Russell W. Belk (1978), "The Effects of Product Involvement and Task Definition On Anticipated Consumer Effort," in *Advances In Consumer Research*, Vol. 5, ed. H. Keith Hunt, Ann Arbor, MI: Association for Consumer Research, 313-318.
- Cohen, Joel B. (1983), "Involvement and You: 1000 Great Ideas," in *Advances in Consumer Research*, Vol. 10, eds. Richard Bagozzi and Alice Tybout, Ann Arbor, MI: Association for Consumer Research, 325-328.
- _____ and Marvin E. Goldberg (1970), "The Dissonance Model in Post-Decision Product Evaluation," *Journal of Marketing Research*, 7 (August), 315-321.
- Comfrey, Andrew L. (1973), *A First Course in Factor Analyses*, New York: Academic Press.
- Engel, James F., David Kollat, and Roger D. Blackwell (1978), *Consumer Behavior*, New York: Dryden Press.
- _____ and Roger D. Blackwell (1982), *Consumer Behavior*, New York: Dryden Press.
- French, John W. and William B. Michael (1966), *Standards for Educational and Psychological Tests and Manuals*, Washington, D.C.: American Psychological Association.
- Greenwald, Anthony G. and Clark Leavitt (1984), "Audience Involvement in Advertising: Four Levels," *Journal of Consumer Research*, 11 (June), 581-592.
- Houston, Michael J. and Michael L. Rothschild (1978), "Conceptual and Methodological Perspectives in Involvement," in *Research Frontiers in Marketing: Dialogues and Directions*, ed. S. Jain, Chicago: American Marketing Association, 184-187.
- Howard, John A. and Jagdish N. Sheth (1969), *The Theory of Buyer Behavior*, New York: John Wiley.
- Hupfer, Nancy and David Gardner (1971), "Differential Involvement with Products and Issues: An Exploratory Study," in *Proceedings: Association for Consumer Research*, ed. David M. Gardner, College Park, MD: Association for Consumer Research, 262-269.
- Kassarjian, Harold H. (1978), "Presidential Address," in *Advances in Consumer Research*, Vol. 5, ed. H. Keith Hunt, Ann Arbor, MI: Association for Consumer Research, xiii-xiv.
- _____ (1981), "Low Involvement: A Second Look," in *Advances In Consumer Research*, Vol. 8, ed. Kent B. Monroe, Ann Arbor, MI: Association for Consumer Research, 31-34.
- Krugman, Herbert E. (1962), "An Application of Learning Theory to TV Copy Testing," *Public Opinion Quarterly*, 26 (Winter), 626-634.
- _____ (1965), "The Impact of Television Advertising: Learning Without Involvement," *Public Opinion Quarterly*, 29 (Fall), 349-356.
- _____ (1967), "The Measurement of Advertising Involvement," *Public Opinion Quarterly*, 30 (Winter), 583-596.
- _____ (1977), "Memory Without Recall, Exposure Without Perception," *Journal of Advertising Research*, 17 (August), 7-12.
- Lastovicka, John L. (1979), "Questioning the Concept of Involvement Defined Product Classes," in *Advances in Consumer Research*, Vol. 6, ed. William L. Wilkie, Ann Arbor, MI: Association for Consumer Research, 174-179.
- _____ and David M. Gardner (1978a), "Components of Involvement," in *Attitude Research Plays For High Stakes*, eds. John C. Maloney and Bernard Silverman, Chicago: American Marketing Association, 53-73.
- _____ and David M. Gardner (1978b), "Low Involvement Versus High Involvement Cognitive Structures," in *Advances in Consumer Research*, Vol. 5, ed. H. Keith Hunt,

- Ann Arbor, MI: Association For Consumer Research, 87-92.
- Loken, Barbara (1984), "Attitude Processing Strategies," *Journal of Experimental Social Psychology*, 20 (July), 272-296.
- Markin, Rom J. and Chem Narayana (1975), "Behavior Control: Are Consumers Beyond Freedom and Dignity?" in *Advances in Consumer Research*, Vol. 3, ed. Beverlee B. Anderson, Ann Arbor, MI: Association for Consumer Research, 222-228.
- Mitchell, Andrew A. (1979), "Involvement: A Potentially Important Mediator of Consumer Behavior," in *Advances in Consumer Research*, Vol. 6, ed. William L. Wilkie, Ann Arbor, MI: Association for Consumer Research, 191-196.
- and Jerry C. Olson (1981), "Are Product Attribute Beliefs the Only Mediator of Advertising Effects on Brand Attitude," *Journal of Marketing Research*, 18 (August), 318-332.
- Nunnally, Jum C. (1978), *Psychometric Theory*, New York: McGraw-Hill.
- Olshavsky, Richard W. and Donald H. Granbois (1979), "Consumer Decision Making—Fact or Fiction?" *Journal of Consumer Research*, 6 (September), 93-100.
- Osgood, Charles, George J. Suci, and Percy H. Tannenbaum (1957), *The Measurement of Meaning*, Urbana, IL: University of Illinois Press.
- Park, C. Whan and S. Mark Young (1983), "Types and Levels of Involvement and Brand Attitude Formation," in *Advances in Consumer Research*, Vol. 10, eds. Richard Bagozzi and Alice Tybout, Ann Arbor, MI: Association for Consumer Research, 320-324.
- Petty, Richard E. and John T. Cacioppo (1979), "Issue Involvement Can Increase Or Decrease Message Relevant Cognitive Responses," *Journal of Personality and Social Psychology*, 37 (October), 1915-1926.
- and John T. Cacioppo (1981), "Issue Involvement As A Moderator of the Effects On Attitude of Advertising Content and Context," in *Advances in Consumer Research*, Vol. 8, ed. Kent B. Monroe, Ann Arbor, MI: Association for Consumer Research, 20-24.
- , John T. Cacioppo, and Rachel Goldman (1981), "Personal Involvement as a Determinant of Argument-Based Persuasion," *Journal of Personality and Social Psychology*, 41 (November), 847-853.
- Ray, Michael L. (1973), "Marketing Communications and The Hierarchy-of-Effects," in *New Models for Mass Communication Research*, ed. P. Clarke, Beverly Hills, CA: Sage Publications, 147-176.
- Robertson, Thomas S. (1976), "Low Commitment Consumer Behavior," *Journal of Advertising Research*, 16 (April), 19-24.
- Rothschild, Michael L. (1984), "Perspectives in Involvement: Current Problems and Future Directions," in *Advances in Consumer Research*, Vol. 11, ed. Tom Kinnear, Ann Arbor, MI: Association for Consumer Research, 216-217.
- Sheth, Jagdish N. and M. Venkatesan (1968), "Risk Reduction Process in Repetitive Consumer Behavior," *Journal of Marketing Research*, 5 (August), 307-310.
- Traylor, Mark B. (1981), "Product Involvement and Brand Commitment," *Journal of Advertising*, 21 (December), 51-56.
- Tyebjee, Tyzoon T. (1979), "Response Time, Conflict, and Involvement in Brand Choice," *Journal of Consumer Research*, 6 (December), 259-304.
- Wright, Peter (1973), "Cognitive Processes Mediating Acceptance of Advertising," *Journal of Marketing Research*, 10 (February), 53-62.
- "Analyzing Media Effects on Advertising Response," *Public Opinion Quarterly*, 38 (Summer), 192-205.
- Zaichkowsky, Judith Lynne (1985), "Familiarity: Product Use, Involvement or Expertise?" in *Advances in Consumer Research*, Vol. 12, eds. Elizabeth C. Hirshman and Morris B. Holbrook, Provo, UT: Association for Consumer Research, 296-299.