

# REFLECTIONS ON A TEST FOR CONSTRUCT VALIDITY IN SOME Q TYPOLOGIES

G. Norman Van Tubergen  
*University of Kentucky*

Robert A. Olins  
*Communications Research Incorporated*

Some of the most consequential issues which face researchers, but which nonetheless are among the least examined, have to do with the reliability and/or validity of a particular investigation or its methodologies. Recently, Fairweather (1981) has done a great service to researchers who use Q methodology by reporting a series of tests bearing on reliability and validity questions arising from two studies of attitudes toward land use. While acknowledging that "no empirical test can make a conclusive judgment" on these issues, he provides strong evidence in support of the reliability of typological constructs derived from Q, both across time and across subjects--even in the face of revised instrumentation. He also approaches the matters of external validity and construct validity.

In connection with construct validity (which he defines as "finding out whether the beliefs, values, and attitudes are in fact the real basis for the discovery of" different Q typologies), Fairweather sug-

*Operant Subjectivity*, 1982(Oct), 6(1), 22-30.

gests that it is the most difficult to assess, and recognizes that the case he makes for construct validity is somewhat weak. We would add that while construct validity is the most difficult form of validity to demonstrate, it is at the same time the most critical in terms of demanding demonstration.

This knotty problem prompted us to review data from a series of studies conducted for one of our clients to see if we could provide any additional support to the issue of construct validity in Q typologies. Of course, we recognized, as did Fairweather, that whatever evidence we might find would not be conclusive, but we felt strongly that a shared accumulation of such evidence would bring social scientists much closer to conclusiveness--if in no other way than to stimulate communication on the subject.

#### BACKGROUND AND METHOD

In mid-1979, a client of Communications Research Incorporated commissioned the company to conduct a segmentation study of the market for a particular type of agricultural chemical. Using a large-sample adaptation of Q methodology which we have mentioned in a previous report (Van Tubergen & Olins, 1979), we identified three major types of farmer-consumers. In that study, 305 respondents Q sorted 71 statements which expressed attitudes and beliefs about farming and cultural practices and the role which the chemical product category under study was seen to play in these practices. Each of the three hypothetical types of farmers was summarized for the client in a one- or two-page description, setting forth our perception of the major attitudinal "facts" and our interpretations of the linkages among those "facts."

A year and a half later, the same client commissioned us to conduct two brief telephone surveys among new samples drawn from the same farmer population with which the segmentation study had dealt. The purpose of these surveys was to seek answers to specific questions bearing on the client's immediate marketing problems which required gathering information not touched upon in the earlier investigation.

It was recommended to the client that an attempt be made to categorize the respondents interviewed in the telephone surveys according to the previously developed Q typologies, with the hopes that such categorization would enhance interpretation of the survey results. It was, of course, impractical to readminister the Q-sort instrument by telephone. Instead, discriminant analysis of the original segmentation data was utilized to select four statements from the complete set of 71 which, it was discovered, were capable of correctly classifying about 80% of the respondents in the original study. (In the discriminant analysis, four statements appeared to be the minimum that could be used and still achieve relatively high classification accuracy; to increase the accuracy to the order of 90% would have required more than twice as many statements, which would certainly have increased the length of the telephone interviews.) The four statements selected in this manner were added to the questionnaires for each of the two surveys, with respondents asked to indicate agreement or disagreement for each statement on a five-point scale. During analysis, information from the discriminant analysis was used to assign respondents to the Q typology into which they appeared to fit best, based on their ratings of the four statements.

It is acknowledged that the procedure described lacks a good deal of rigor and severely pressures the reliability of the typologies. Some misclassifications would be expected even if the original respondents were recontacted. In this instance, we have new respondents, a year and a half later, responding to a totally different instrument. With regard to the analysis presented in this paper, there is no reason to believe that misclassification would systematically favor any of the hypotheses to be tested; on the contrary, weakened reliability of the typologies could reasonably be expected to reduce the chances for demonstrating construct validity.

As Fairweather points out, "independent evidence for construct validity is necessary" and this can best be provided "by deriving hypotheses from the theory underlying the construct." For the present

analysis, we went back to the two telephone questionnaires and reviewed each question (other than the rating of the four attitude statements) which was put to the respondents. We also reviewed the summary typological descriptions from the original segmentation study (the constructs). For each question in the surveys, we asked ourselves what prediction(s), if any, could be made regarding the responses to the question based on our knowledge of the typological constructs. Since we felt the survey data might not be sufficiently robust to support elaborate statistical and analytic procedures, we confined ourselves to hypotheses which could be tested in the most straightforward manner--either for a significant difference of proportions, or for a significant difference of means.

This examination yielded 13 research hypotheses which could be tested with data occurring in one or both of the survey questionnaires. This set of hypotheses covers such diverse areas as demographic characteristics, cultural practices used, brand preferences in the product category, product usage practices, and economic behavior. Although the ethics of confidentiality for our client's proprietary data prohibit us from presenting the specific detail of these hypotheses, we have tried to present them as explicitly as possible so that the reader may sense the nature and range of predictions we attempted to make for this analysis. The 13 hypotheses derived from the typological constructs are as follows:

1. Respondents classified as type I will be located in the client's marketing regions 2 and 3 more frequently than will other respondents.
2. Respondents classified as type III will be located in the client's marketing region 4 more frequently than will other respondents.
3. Respondents classified as type I will operate larger farms than will other respondents.
4. Respondents classified as type II will have used more brands in the product category than will have been used by other respondents.
5. Respondents classified as type I will more fre-

- quently report using products in combination (rather than singly) than will other respondents.
6. Respondents classified as type I will more frequently report using a combination of brand A with brand B than will other respondents.
  7. Respondents classified as type II will more frequently report using a combination of brand C with brand D than will other respondents.
  8. Respondents classified as type III will less frequently report using products in combination (rather than singly) than will other respondents.
  9. Respondents classified as type I will report a perception of product problem X as being less consequential than will respondents classified as type III.
  10. Respondents classified as type II will more frequently report using cultural practice M than will other respondents.
  11. Respondents classified as type III will less frequently report using cultural practice N than will other respondents.
  12. Respondents classified as type III will less frequently report using cultural practice P than will other respondents.
  13. Respondents classified as type III will report spending less money on products in the product category than will other respondents.

It will be observed that five of the hypotheses make predictions regarding respondents classified as type I, three refer to respondents classified as type II, and six make predictions about answers given by respondents classified as type III. Note that one hypothesis, #9, calls for an explicit comparison of the answers from type I farmers with those from type III farmers.

## RESULTS AND DISCUSSION

Before looking at the tests of these hypotheses, let us consider briefly the implications of the possible

outcomes. If our research hypotheses are not supported, there would be several possible explanations, among which the most critical would probably be:

- (a) the typologies lack construct validity,
- (b) the research hypotheses do not provide an adequate test of typological construct validity,
- (c) the method used to assign survey respondents to typological constructs is inadequate, and/or
- (d) the typological constructs lack reliability.

These last two points emphasize that our analysis is based on the assumption that our Q typologies would be reliable across time and across subjects and that we can correctly classify new subjects in the way described. Although the work of Fairweather (as well as the unreported experiences of ourselves and other Q researchers) supports the reliability assumption, no direct test of reliability is possible since the procedure used compelled each respondent to be assigned to one of the previously observed typologies. The assumption does receive some encouraging support indirectly, however, as a consequence of the typological distribution in the samples studied. Because Q studies usually involve small samples, the distribution of typologies cannot be generalized to a larger population; nevertheless, the samples in these studies were large enough (original segmentation  $n = 305$ , telephone survey I  $n = 399$ , telephone survey II  $n = 149$ ) to permit reasonable comparison of typological distribution. If the distribution of typologies in the surveys conforms with the distribution in the original segmentation study, then it is reasonable to say that some support exists for the assumption of typological reliability across time and across subjects.

As the figures in Table 1 indicate, the proportion of respondents in each survey who were classified as type II was the same (within sampling error). Further, in all three studies, the largest portion of respondents was classified as type I. While there

Table 1  
CROSS-STUDY COMPARISONS OF TYPOLOGY DISTRIBUTIONS

Type	Segmentation	Survey I	Survey II
I	40%	63%	63%
II	25	24	19
III	35	13	17
<i>n</i>	305	399	149

appears to be an increase in type I and a decrease in type III membership, this difference can be explained by the fact that type III is affiliated with a specific geographic area (see the test for hypothesis #2 below), from which a smaller proportion of the survey samples was drawn than was the case in the original segmentation study. With this explainable difference, the distributions of typologies from the surveys are quite similar to those associated with the development of the constructs.

From this reasoning, if our research hypotheses are not supported, it would be more likely to be due to one of the first two reasons cited above. If, on the other hand, the research hypotheses are supported (and if such an occurrence is assumed not to be spurious), then they provide evidence that:

- (a) the typological constructs in this series of studies have construct validity,
- (b) the method used to assign survey respondents to typological constructs was probably a good estimator, and
- (c) the typological constructs found in the original segmentation study display reliability across time and across subjects, although the support here would be somewhat inferential.

Returning to the research hypotheses, the data gathered in the two telephone surveys made possible 18 tests of the 13 hypotheses. The results of these

Table 2  
RESULTS OF TESTS OF RESEARCH HYPOTHESES

Hypothesis	Difference Tested	Survey I	Survey II
1	proportion	ns	x
2	proportion	.01	x
3	mean	.0025	ns
4	mean	ns	.05
5	proportion	.025	x
6	proportion	.001	.0005
7	proportion	.025	.05
8	proportion	.0005	x
9	mean	x	.025
10	proportion	.01	x
11	proportion	.025	x
12	proportion	x	.0005
13	mean	.001	.01

Note: Numerical entries are probability levels for tested differences. ns = no significant difference, x = no test of hypotheses in specified survey.

tests are summarized in Table 2, which shows that 12 of the 13 hypotheses were supported by data from at least one of the two telephone surveys; three hypotheses were supported in both surveys. A total of 15 of the 18 tests performed showed significant differences at or beyond the .05 level. Since all of the hypotheses were directional, there is further encouragement in the fact that of the three tests which were unresponsive none ran significantly in the opposite direction. Finally, since each of the three typological constructs was involved independently in various hypotheses, it is gratifying to see that none of the three constructs was left without at least some support in this series of tests.



## CONCLUSION

By the reasoning presented earlier, we believe that this analysis presents strong evidence of construct validity for the original segmentations typologies. We grant that the ad hoc nature of this analysis means that many hypotheses which one might derive from the constructs involved were left untested, but we consider the strong results from those which are tested (in what could be considered a random choice fashion) to be remarkable, especially in the face of serious obstacles to a favorable outcome posed by the use of different respondents at a much later time and classified to typologies by a procedure which is clearly less than perfect. Again each of these obstacles would seem much more likely to introduce randomness which would obscure any observed differences than they would to introduce systematic bias favoring the discovery of differences. Thus, we feel that while the analysis we present is in no way definitive, and applies only to this one domain of study, it provides strong demonstration of the construct validity and predictive power of typological descriptions derived by Q methodology.

- G. Norman Van Tubergen, Department of Communication, University of Kentucky, Lexington, KY 40506 (and Senior Vice-President, Communications Research Incorporated)*
- Robert A. Olins, President and Chairman, Communications Research Incorporated, 303 East Wacker Drive, Suite 1907, Chicago, IL 60601*

## REFERENCES

- Fairweather, J.R. Reliability and validity of Q-method results: Some empirical evidence. *Operant Subjectivity*, 1981, 5, 2-16.
- Van Tubergen, G.N. & R.A. Olins. Mail vs. personal interview administration for Q sorts: A comparative study. *Operant Subjectivity*, 1979, 2, 51-59.