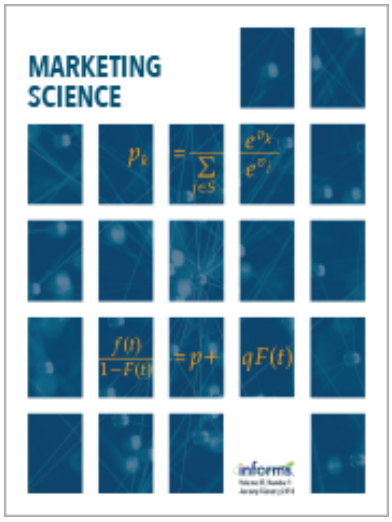

Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Comment—Recall, Recognition and the Dimensionality of Memory for Print Advertisements: An Interpretative Reappraisal

Daniel J. Howard, Alan G. Sawyer,

To cite this article:
Daniel J. Howard, Alan G. Sawyer, (1988) Comment—Recall, Recognition and the Dimensionality of Memory for Print Advertisements: An Interpretative Reappraisal. Marketing Science 7(1):94-98. https://doi.org/10.1287/mksc.7.1.94

COMMENT

MARKETING SCIENCE
Vol. 7, No. 1, Winter 1988
Printed in U.S.A.

RECALL, RECOGNITION AND THE DIMENSIONALITY OF MEMORY FOR PRINT ADVERTISEMENTS: AN INTERPRETATIVE REAPPRAISAL

DANIEL J. HOWARD AND ALAN G. SAWYER
Southern Methodist University
The University of Florida-Gainesville

Bagozzi and Silk (1983), B&S hereafter, report the results of LISREL analyses of the covariance structure of aggregate recognition and recall data collected 30 years ago by the *Advertising Research Foundation* (1956) in a study of Printed Advertising Rating Methods (PARM). The authors conclude that "memory is multidimensional." Several conceptual and methodological issues, however, cloud the interpretation of their results.

Conceptual Issues

The issue of whether a "lack of fit" between recognition and recall reflects the "dimensionality of memory" depends on the nature of the measurements employed and exactly what is meant by the "dimensionality of memory." It should first be pointed out that recognition in the PARM study was measured by the proportion of magazine readers *claiming* to have "noted" a given ad at some earlier time; it was not possible to verify the accuracy of reader claims, as is typical of the psychological studies cited by B&S. This contributes to the lack of theoretical clarity in the meaning of this measure and the interpretability of its covariation with recall scores.

The major conceptual problem is that B&S have not presented a clear view of what their "dimensionality of memory" construct means. Concerning the recognition and recall controversy, the authors begin by noting that "the critical questions are why and how do these two measures covary" (pp. 99–100). Unfortunately, the authors seem only justified concluding in the end that recognition and recall may sometimes behave differently (i.e., evidence on "how" they covary), which is something that has been known for many years (Brown 1976). In our view, it is not very helpful at this stage of theoretical development to suggest that the lack of covariation between recognition and recall may reflect different memory processes (i.e., "memory is multidimensional") if one is unable to say *which* memory processes (i.e., evidence about "why" they do or do not covary).

Although B&S never explicitly define the construct of "dimensionality of memory," their extensive references suggest that "multidimensionality" refers to either different *acquisition* processes, such as levels of processing theory, (Craik and Lockhart 1972; Craik and Tulving 1975), dual-coding theory (Paivio 1971), and Tulving's (1972) distinction between episodic and semantic memory; different *retrieval* mechanisms, such as Anderson and Bower's (1972) dual-process theory; or some unspecified combination of these processes. Thus, grounds are not provided for an unambiguous interpretation of the dimensionality of memory construct examined. Unless B&S can better specify the theoretical meaning and source of multiple dimensions of memory represented by the extent of covariation of recall and recognition (i.e., the "why" question), very little has been learned.

0732-2399/88/0701/0094$01.25

First, let us examine the position that the relationship between the two measures can be used to make inferences concerning differential properties of information *acquisition.* Any such inference would require data that conform (in function) to the encoding/retrieval paradigm used in investigations of memory processes (see Tulving 1983). This paradigm allows for the manipulation or measurement of differential conditions of encoding across differential conditions of retrieval. In an encoding experiment, retrieval conditions are held constant while encoding conditions are varied; in a retrieval experiment, encoding conditions are held constant while retrieval conditions are varied. Thus, implementation of the full paradigm allows the "teasing apart" of influences on data from the different sources. With an encoding/retrieval interaction, inferences can be made about either the nature of encoding processes, retrieval processes, or both. The PARM data, as utilized by B&S, conform most closely to the minimal necessities of a retrieval experiment. The problem when using the relationship between two different conditions of retrieval (e.g., recognition vs. recall) to make inferences concerning the presumed "multidimensionality" of an equivalent condition of acquired information is that differential patterns of variation may reflect known or suspected differences in the means of retrieval and not differences in terms of acquisition.[1] Recognition and recall differ at least in terms of the retrieval cues provided and the memory performance required. Such differences can by themselves result in a breakdown in the relationship between the two measures. There was no way of determining whether any of the acquisition influences cited by B&S were operating since encoding conditions were not meaningfully varied nor measured; a lack of fit between recognition and recall cannot be interpreted in terms of differences in memorial acquisition processes when any such differences cannot be distinguished from processes unique to retrieval.[2]

The second issue is whether the covariation between recognition and recall can be used to infer whether the measures involve similar (unidimensional) or distinct (multidimensional) forms of *retrieval.* When recognition and recall scores are examined across a common condition of input and differential patterns of performance are observed, one interpretation is that distinct retrieval processes are involved. Dual process theory asserts that the retrieval mechanisms are distinct.[3] However, Tulving's (1983) update of his earlier (Tulving 1976) synergistic ecphory model proposes a one-step process for recognition and recall that can accommodate much the same patterns of data once used as indirect evidence for dual process theories.[4] As Tulving notes,

[1] Although Bagozzi and Silk stratified the data by the time interval between last reading of the magazine and the interview data, that variable was not discussed in terms of implications of any of the "multidimensional" memory theories they cite.

[2] The reader is referred to Paivio (1976) for a discussion of the implications of the covariation of recognition and recall performance as a function of dual vs. non-dual process coding (as well as levels of processing and semantic/nonsemantic memory). Although a casual reading of Paivio's analyses might suggest a close resemblance to that of B&S, Paivio, unlike B&S, included either manipulations or measures of encoding (i.e., acquisition) conditions; this difference is a critical one if one believes, as B&S do, that acquisition differences may be important. In particular, it should be noted that in analyses involving direct comparisons of recognition and recall when attempting to make inferences about acquisition, encoding conditions were either directly manipulated (e.g., words encoded verbally vs. words encoded as images), or directly measured (e.g., rated imagery-evoking value of items). However, like B&S, Paivio's analyses can be criticized for reliance on aggregated data.

[3] The version of dual process theory described by Anderson and Bower (1972) assumes that recognition and recall are basically different processes not only because recognition is viewed as a subprocess in recall (i.e., items are first recalled from memory, then the "correct" ones are recognized), but also because the inter-item (node) pathway associations responsible for retrieval (necessary for recall) is a process considered independent of the list-tag item node associations responsible for recognition decisions.

[4] Tulving's (1983) updated synergistic ecphory model is a relatively new model of retrieval and is not developed to the point where all of its premises are testable although some of them, such as the different

> Recognition can now be shown to be higher than recall as well as lower; recognition and recall scores may be correlated positively, negatively or not at all; in one situation, recall of one thing may involve recognition of another, whereas in another situation, recognition of one thing may require recall of another; variations in certain variables may produce changes in recall but not in recognition, or changes in recognition but not in recall. Even this partial summary of general experimental findings clearly suggests that a simple statement of the relationship between recognition and recall is no more possible (p. 301).

At issue here is a concern with the interpretability of any "dimensionality of memory" construct that might be derived from simply examining the covariance structure of recognition and recall scores when evidence is not provided to support or discount either preferred (i.e., dual process) or competing (i.e., any other) explanations that might be involved.

Methodological Issues

As discussed above, the meaning of the recognition measure in the PARM study is unclear. In addition to making comparisons with past psychological experiments tenuous, the lack of control for guessing could have created specific measurement problems with the recognition measure such as "yea-saying bias" (Wells 1961). "Yea-saying" bias would likely inflate the correlation between the recognition and reader interest measures (but not the recall-interest correlation), a result consistent with B&S's results.[5]

There is also empirical and anecdotal evidence of problems with the PARM recognition data collection procedure. In B&S's Table 1 (p. 98), it can be seen that recognition scores increased from the first to the second time interval and decreased only slightly for the third, with the third period's score still higher than the first's. The A.R.F. report was suspicious of the plausibility of increasing recognition over time and suggested that instead there was most likely something in the interview situation itself which produced the erratic behavior of recognition scores (*Advertising Research Foundation,* Vol. IV, No. 6, p. 3, 1956). Similar problems were not reported for the aided recall measure.

TABLE 1

Partitioning of Variance Due to Trait and Error for Bagozzi and Silk's (1983) *Models* 1, 3 *and* 4.

	Proportion of Variance due to:					
	Model 1		Model 3		Model 4	
Measure	Trait	Error	Trait	Error	Trait	Error
PNR1	0.466	0.534	0.651	0.349	0.466	0.534
PNR2	0.390	0.610	0.474	0.526	0.397	0.603
PNR3	0.325	0.675	0.495	0.505	0.324	0.676
NOT1	0.801	0.199	0.590	0.179	0.802	0.198
NOT2	0.806	0.194	0.602	0.197	0.802	0.198
NOT3	0.666	0.334	0.462	0.173	0.668	0.332

informational requirements for recognition versus recall, are testable. See Tulving (1976, 1983) for a discussion and review of past research providing indirect support for dual-process mechanisms.

[5] This speculation assumes that the measure of reader interest is at least partially valid. We have refrained from discussing conceptual or theoretical issues concerning B&S's findings with the "reader interest" measure since they acknowledge that "its meaning is ambiguous" (p. 124). To reinforce that view, however, it should be noted that Darrell Lucas, Technical Director of the *Advertising Research Foundation* at the time of the PARM study and an original member of the PARM Committee, cautioned against the use of the measure: "I don't know what it is. I don't know anyone who claims to know what it is. I would have left it out" of the PARM study (Lucas 1956, p. 2).

Consistent with the explanation that there may have been an interview bias with the recognition but not the recall measure is B&S's finding that ". . . our empirical test of the model allowing correlated errors for measures of recall and measures of recognition reveals that only the latter achieve statistical significance" (p. 114). B&S note that ". . . ultimately one would like to identify the specific causes of correlated errors because this is necessary for the development of a true theory of memory processes" (p. 117). However, a plausible cause of the correlated errors may simply be the interview bias. If controlling for systematic error variation due to an erratic interviewing situation for the recognition measure can plausibly and parsimoniously account for model fit, it is then unnecessary to conclude that the recall and recognition measures also reflect "other mental states" (p. 120).

Another methodological problem involves a possibly misleading conclusion about the effect of reader interest on recognition and recall. In their final model, Bagozzi and Silk achieve marginally acceptable model fit ($p \cong 0.09$) when controlling for reader interest. However, comparing that preferred model to only the third model when making conclusions concerning the measurement of recognition and recall may be misleading. They state, ". . . when interest was held constant, trait variation exhibited by recognition measures was impressive, ranging from about 67 to over 80 percent. The control of interest seems to be a viable means to enhance the measurement of recognition but not necessarily recall" (p. 121). The problem is that B&S do not compare the trait variation exhibited in their final model with the trait variation exhibited in their initial model. The trait variation for recognition and recall measures is presented for models 1, 3, and 4 in Table 1.[6]

The results in Table 1 show that, although model 1 did not yield an adequate statistical fit to the data, the trait variation exhibited by recognition and recall measures before reader interest is controlled is essentially equivalent to the trait variation exhibited in the final model 4, after reader interest is controlled. In fact, the only thing which is demonstrated to influence the amount of trait variation attributed to recognition and recall measures is whether or not the correlated recognition error scores are explicitly modeled. In other words, the reason trait variation in recognition scores increased in the final model may be because the methods factor was removed from the design and not because the reader interest variable was added.

Finally, although B&S are not the first to use aggregate data to investigate memorial processes, we question the assumption that meaningful inferences concerning *individual* level memory processes are possible with *aggregate* level proportions from samples of different respondents. The PARM data (correlations of measures of recall, recognition, and reader interest across 95 advertisements) represent properties of advertisements and not individuals. The most meaningful study of individual level cognitive processes would seem to require individual level analyses. The dangers of aggregation bias are especially worrisome in the study of memory where aggregation can result in misleading conclusions concerning the type of individual processes involved (see Baloff and Becker 1967).

Conclusion

In his award-winning philosophy of science article, Bagozzi (1984) describes the Bagozzi and Silk research as an "illustration of Holistic Construal in (a) substantive context" (p. 20). He also stresses

> the importance of making any piece of research an integral part of a body of knowledge . . . drawing upon prior learning in the literature, making special efforts to develop sound measure-

[6] B&S calculated the trait variation for recognition and recall scores for models 3 and 4. The trait variation for model 1 was calculated from Bagozzi and Silk's Table 3.

> ments, . . . Marketing is entering a new era that will demand of its researchers excellence in both theory and method. The irony is that as we become more explicit in our theorizing and modeling, so too will we become more vulnerable to criticism . . . Rather than playing it safe and hiding behind well-developed theory that is never really tested, or a sophisticated methodology that is applied without sound conceptual groundwork, we should demand of ourselves the best that can be offered in both substance and method (p. 27).

We agree with Bagozzi's challenge to marketing researchers and are critical about the ability of the B&S 1983 research to meet Bagozzi's very tough standards.[7] We caution against any inferences involving the cited psychological literature. This is not meant as a criticism of the authors for a re-analysis of the PARM data, but only the attempted interpretation in terms of concepts, theories and issues that the PARM study was neither developed nor intended to address. Elegant LISREL analysis of the covariation of advertising recall, recognition, and reader interest scores cannot offset the conceptual and methodological problems discussed above.[8]

Acknowledgements. The authors wish to thank several colleagues at the University of Florida, The Ohio State University, the University of Wisconsin and anonymous referees for their helpful comments.

[7] We also agree that it is easier to *criticize* than it is to *do* research that follows the tenants of critics unfettered by real data or other problems. We salute B&S for their efforts to make some sense of the PARM data even though we also believe that their efforts have fallen considerably short of the standards prescribed by Bagozzi's Holistic Construal.

[8] This paper was received April 1984 and has been with the authors 20 months for 2 revisions.

References

Advertising Research Foundation (1956), *A Study of Printed Advertising Rating Methods.* Vol. I–V. New York: Advertising Research Foundation.

Anderson, J. R. and G. H. Bower (1972), "Recognition and Retrieval Processes in Free Recall," *Psychological Review,* 79, 97–123.

Bagozzi, R. P. (1984), "A Prospectus for Theory Construction in Marketing," *Journal of Marketing,* 48, 11–29.

——— and A. J. Silk (1983), "Recall, Recognition and the Measurement of Memory for Print Advertisements," *Marketing Science,* 2, 95–134.

Baloff, N. and S. Becker (1967), "On the Futility of Aggregating Individual Learning Curves," *Psychological Reports,* 20, 183–191.

Brown, J. (1976), *Recall and Recognition.* New York: John Wiley.

Craik, F. I. M. and R. S. Lockhart (1972), "Levels of Processing: *A Framework for Memory Research," Journal of Experimental Psychology: General,* 104, 268–294.

——— and E. Tulving (1975), "Depth of Processing and the Retention of Words in Episodic Memory," *Journal of Experimental Psychology: General,* 104, 268–294.

Lockhart, R. S., F. I. M. Craik and L. Jacoby (1976), "Depth of Processing, Recognition and Recall," in J. Brown (Ed.), *Recall and Recognition.* New York: John Wiley.

Lucas, D. B. (1956), "What Space Sellers and Buyers Can Expect from the PARM Study," paper delivered to the New York Chapter of the American Marketing Association, April 12, 1956.

Paivio, A. (1971), *Imagery and Verbal Processes.* New York: Holt, Rinehart and Winston.

——— (1976), "Imagery in Recall and Recognition," in J. Brown (Ed.), *Recall and Recognition.* New York: John Wiley.

Tulving, E. (1983), *Elements of Episodic Memory.* New York: Oxford University Press.

——— (1976), "Ecphoric Processes in Recall and Recognition," in J. Brown (Ed.), *Recall and Recognition.* New York: John Wiley.

——— (1972), "Episodic and Semantic Memory," in E. Tulving and W. Donaldson (Eds.), *Organization and Memory.* New York: Academic Press.

Wells, W. D. (1961), "The Influence of Yea-saying Response Style," *Journal of Advertising Research,* 4, 1–12.