British Journal of Psychology (1993), **84**, 289–295 *Printed in Great Britain*

Abstraction of covariation in category learning: A critical note on K. Richardson's studies

Pierre Perruchet* and **Chantal Pacteau**
Laboratoire de Psychologie Différentielle, Université René Descartes, 28 rue Serpente 75006 Paris, France

Jorge Gallego
Université Pierre et Marie Curie

Several studies directed by K. Richardson (Richardson, 1986, 1987; Richardson & Carthy, 1989, 1990) claimed that people are able to abstract imperfect covariations among feature variables defining artificial categories, and to use that information in various subsequent tasks. We challenge this claim on a twofold basis. First, a large part of the performance variance Richardson attributed to covariation knowledge may be accounted for by a very simple exemplar model, which assumes no abstractive processes. Second, the residual influence of relational information on performance may be attributed to the knowledge subjects have acquired in real-world situations before their training with the study exemplars.

A recent series of experiments carried out in K. Richardson's laboratory aimed at demonstrating the ability of people to abstract imperfect or 'fuzzy' covariations among feature variables defining categories, and to use that information in various subsequent tasks (Richardson, 1986, 1987; Richardson & Carthy, 1989, 1990).

Let us consider Expt 1 in Richardson & Carthy (1990) which will serve as a privileged example throughout this comment. Stimuli were pictures of robots. They differed along three variables (I, J and K): the length of the arms, the size of the body and the size of the head. Each variable had three possible values (1: small; 2: medium; 3: large). All the possible combinations of variables defined a set of 27 (3^3) items whose configuration is reported in Table 1, column 1. In the first phase of the experiment, subjects were asked to study and remember nine out of the 27 possible pictures. These items, along with their occurrence frequencies (see Table 1, column 2), were selected so that the three variables exhibited a moderately high overall level of covariation (in fact, only the I and J, and J and K associations were significant).

In a subsequent test phase, subjects were shown the full set of 27 exemplars. They were asked to indicate whether each of the pictures was a member of the experienced set, and to rate their level of confidence about their decision. The mean results (after coding the responses on a six-point scale) are reported in Table 1, column 3.

These empirical results were compared with the values predicted by a covariation

* Requests for reprints.

coding model, based on the application of log-linear statistics. The method of computation is described at length in all the Richardson and associates studies, and it will not be reproduced here in order to save space. It is sufficient for our concern to point out that a log-linear model allows the simulation of the performance of a subject who would have abstracted the covariations embedded in the study material, and would make recognition judgements on the sole basis of this relational information. Pearson's r between recognition judgements and predicted values (shown in Table 1, column 4) was a substantial .701.

Several earlier studies investigated recognition or typicality judgements in similar situations (Richardson, 1986; Richardson, 1987, Expt 1). They also gave evidence that values predicted by the covariation coding model fit subjects' observed performance well. In other studies (Richardson, 1987, Expt 2; Richardson & Carthy, 1989) subjects were trained as above, but they had to complete the missing part of an incomplete stimulus in the test phase. For instance, after viewing exemplars exhibiting a covariation between I, J and K variables, subjects were asked to infer the value of K given the values on I and J. Subjects' completions were also predicted fairly well by the covariation coding model.

We do not intend to challenge the reliability of the experimental results just outlined. We take hereafter for granted that the covariation coding model is a good predictor of performance, a finding which provides evidence, in a clever and original way, that subjects' performance is sensitive to the moderately strong covariations structuring the environment. What we do challenge is the inference Richardson draws about the nature of the psychological processes underlying the performance pattern. According to Richardson, the results just outlined demonstrate that subjects abstract knowledge about covariation among variables while studying exemplars, and use this abstract knowledge in subsequent recognition, typicality, and completion judgements.

Covariation coding versus exemplar-similarity models

The validity of Richardson's inference rests on the postulate that a concurrent model cannot account for observed performance better than, or at least as well as, the covariation coding model. In this regard, Richardson deals with two other models of categorization, framed around the notions of prototype representation and independent feature frequency respectively. He shows through quantitative simulations that these concurrent models fail to explain a substantial part of the performance variance. However, Richardson pays only marginal attention to another class of models, which does not call for any abstraction process. In these so-called exemplar or 'exemplar-similarity' models, the studied exemplars are stored in memory with their specific properties, and subsequent categorization judgements of new items are made on the basis of their degree of similarity with the stored exemplars (Brooks, 1978; Medin & Shaffer, 1978). Richardson rejects the exemplar models on the basis of anecdotal case reports illustrating a discrepancy between the predictions derived from these models and the empirical data (e.g. Richardson & Carthy, 1990, p. 425). The present section aims at demonstrating that these models, even when instantiated by a few simple rules, are in fact as powerful as the covariation coding model in accounting for observed typicality or recognition judgements.[1]

A very simple exemplar-similarity model

In the model used here, the expected typicality or recognition judgement for a given item is the sum of two components, which are respectively proportional to: (1) the frequency of occurrence of the target item. Proportionality coefficient was set to one, so that this component value was in fact equal to the number of presentations of the item; and (2) the frequency of occurrences of items similar to the target item, the similarity between two items being defined by the fact that they differ only by one value on one variable. Proportionality coefficient was arbitrarily set to 1/2. Thus, for instance, an item similar to an item displayed 10 times was credited an additional expected value of 5 units.

For the sake of illustration, consider again the Richardson & Carthy (1990) Expt 1. Table 1 shows that item *I*1 *J*1 *K*1 (the robot with all small parts) was credited an expected value of 3.5. Three units were attributed because the item was displayed three times in the study phase. The additional .5 unit stems from the fact that the target item is similar to item *I*1 *J*1 *K*2, which was displayed once in the study phase. The target item is also similar to items *I*2 *J*1 *K*1 and *I*1 *J*2 *K*1, but these items were not previously displayed, and hence provide no additional value. The predictions generated for the remaining items according to the same principles are shown in Table 1, column 5.

A variety of other computational modes would have been possible. One option which deserves special mention pertains to the operationalization of the concept of similarity. The model we choose here instantiates featural, and not relational similarity (e.g. Goldstone, Medin & Gentner, 1991). For instance, *I*2 *J*2 *K*2 is not taken to be similar to *I*1 *J*1 *K*1, and hence no transfer is assumed to occur between both items. This choice was intended in order to avoid spurious reintroduction of covariation processing into the notion of similarity. For the other aspects, simplicity was the main criterion of choice.

Comparing predictive power of covariation coding and exemplar models

Correlations between recognition ratings and the predictions generated by covariation coding and exemplar-similarity models were computed over the 27 test items. Pearson's r between predicted values displayed in Table 1 were respectively .701 and .762.

Similar analyses were carried out for all the sets of data available in Richardson's studies. These include (1) results for two different categories ('Gleeks' and 'Loopers') tested in the same subject sample in Richardson (1986), (2) results on

[1] Quantitative comparisons will be presented only for the studies in which subjects were asked for typicality or recognition judgements in the test phase. Conceptual generalization to the studies requiring inference from incomplete information is straightforward. However, the results of most of these studies are not easily prone to *post hoc* analyses, because part of the required results are unavailable; indeed, only test items relevant to Richardson's concerns were presented to subjects during testing.

Table 1. Observed performance (recognition scores on a six-point scale) and values predicted by the two concurrent models for each of the 27 test items. Values in columns 1–4 are taken from Richardson & Carthy (1990, Tables 1*c*, 2*b* and 3). The data represented in italics serve as illustration in the text

Items			Predicted values	
Configuration *IJK*	Number of presentation	Observed values	Covariation coding	Exemplar similarity
111	*3*	*4.81*	*10.49*	*3.5*
121	—	*3.06*	*3.42*	*1.5*
131	—	1.56	1.11	0.5
112	1	*4.19*	*7.34*	*3.5*
122	—	2.83	3.42	0.5
132	—	1.98	1.59	1
113	2	3.61	5.14	2.5
123	—	2.72	3.42	1
133	—	2.03	2.27	1.5
211	—	*3.44*	*4.88*	*1.5*
221	—	3.42	3.42	1.5
231	1	3.06	2.39	2
212	—	3.17	3.42	0.5
222	—	3.64	3.42	2.5
232	2	3.75	3.42	4
213	—	3.09	2.39	1
223	—	3.61	3.42	2
233	3	3.94	4.88	4
311	—	2.00	2.27	1
321	2	3.69	3.42	3.5
331	—	2.53	5.14	1.5
312	—	2.13	1.59	1.5
322	3	3.83	3.42	4.5
332	—	3.44	7.34	2.5
313	—	1.97	1.11	0.5
323	1	3.64	3.42	2.5
333	—	3.70	10.49	2

single category learning for two independent groups of subjects (A and B) differing by their amount of training in Richardson (1987, Expt 1) and (3), four successive tests (T1, T2, T3, T4) on the same subjects for a single category in Richardson & Carthy (1990, Expt 1; T1 serves herein as illustrative example). Predictions from the exemplar-similarity model were generated in each case as indicated above.

Results are shown in Table 2, columns 1 and 3. On the whole, subjects' performance fitted better with the exemplar model than with covariation coding model in six out of the eight comparisons, and the reverse result was obtained in the two others.

Table 2. Summary of the results from hierarchical multiple regression analyses. The data represented in italics serve as illustration in the text

	Predictors of performance				
	Covariation coding		Exemplar similarity		
Sources	*r*	*F* to enter ES	*r*	*F* to enter CC	Multiple R
Richardson (1986)					
'Gleeks'	.483	5.096	.533	3.168	.606
'Loopers'	.594	1.641	.439	8.276	.628
Richardson (1987)					
A	.693	1.108	.610	6.253	.709
B	.568	3.645	.619	1.139	.642
Richardson & Carthy (1990, Expt 1)					
T1	*.701*	*24.003*	*.762*	*15.557*	*.863*
T2	.651	22.749	.763	9.456	.837
T3	.615	14.315	.704	7.065	.781
T4	.568	11.572	.672	4.813	.737

Alternative or complementary models?

Do covariation coding and exemplar-similarity models account for the same or for independent components of variance on typicality or recognition judgements? Consider again the data in Richardson & Carthy (1990, Expt 1, first test). The correlations between the sets of values predicted by the two models was .441. This moderate value makes it possible that each model accounts for at least partially independent components of variance in performance.

This hypothesis was tested with a hierarchical multiple regression technique. Entering predictions of covariation coding and exemplar models into the regression equation produced a multiple R of .863. The squared R, which described the proportion of the performance variance shared with the optimally weighted predictions from the two models, is .745. Covariation coding model alone accounts for a proportion of variance of .491. Entering exemplar model predictions into the regression equation increases the part of explained variance significantly ($F = 24.00$). Similar results were found when the variables were entered in the reverse order: the addition of the covariation coding model preditions significantly increases ($F = 15.56$) the amount of variance explained by the exemplar model alone (.58).

Table 2 shows the results of similar computations for each of the eight sets of data available in the Richardson papers. Multiple R are shown in the right-hand column; they range from .606 to .863, hence demonstrating that combination of the two models achieved good to excellent predictions of subjects' judgements. More importantly, F tests reported in columns 2 and 4 show that in most cases one model adds significant predictive power to the other. This occurs in five out of the eight sets of values when the covariation coding model is entered first in the regression equation, and in six out of the eight sets of values when the entrance order is reversed.

To sum up, the present results show that a very simple exemplar-similarity model is as good a predictor of performance as the covariation coding model, if not a better one. Such a conclusion considerably softens Richardson's claim pertaining to the overall supremacy of the covariation coding model over concurrent ones. However, our results also show that the covariation coding model may still account for a reliable part of variance in performance when predictions of the exemplar model are partialled out, hence suggesting that the apparent sensitivity to covariation cannot be considered only as the by-product of the memory for specific exemplars.

Covariation abstraction or background knowledge?

We now argue that the evidence provided for the use of relational information during the test phase does not testify that covariations are abstracted during the study session.

In most of Richardson's studies, the to-be-abstracted relations bear on the size of the different components of the study items. More precisely, the best exemplars of the concepts are those in which the different components – arms, head and body for the robots – have the same relative size (all small, all medium, or all large). The point is that this kind of dimensional covariation is ubiquitous in everyday subjective experience. Consequently, subjects could be sensitive to dimensional covariation independent of any specific training in the experimental setting.[2]

A close scrutiny of the data in Richardson & Carthy (1990) provides empirical support for this hypothesis. In this study, covariations experienced in the study phase were restricted to only two (*IJ* and *JK*) out of the three pairs of variables in Expt 1, and to only one pair of variables (*IJ*) in Expt 2. Although this aspect was ignored by the authors, the other pairs of variables (namely *IK* in Expt 1, and *IK* and *JK* in Expt 2) may serve as controls. If sensitivity of subjects to relational information in the test phase stems from the abstraction of such information from study exemplars, then subjects' performance must be random when analysis is circumscribed to the variables for which a no-contingency relation was experienced in the study phase. This prediction is clearly confirmed by the data. For instance, Table 7 in Richardson & Carthy (1990) shows that, in Expt 2, the best three recognized items among the nine (3^2) which instantiate all the possible combinations of the *IK* variables are: *I*1 *K*1, *I*2 *K*2, and *I*3 *K*3. Thus subjects trained with items in which *I* and *K* features exhibited no particular relationship to preferentially recognized items in which *I* and *K* features are well-proportionate, and this trend appears to be general.

The major conclusion to be drawn from these data is that subjects' judgements

[2] The single exceptions to the size covariation principle are in studies where stimuli are houses. One of the variables in Richardson (1987, Expt 2) is the number of windows (1, 2 or 3), and in Richardson & Carthy (1989), the shape of the houses' roof (single-, double-, or triple-pitched). These exceptions do not detract from our argument, insofar as the study items still exhibited covariations which parallel natural expectations (e.g. large houses were more likely to have three windows or a triple-pitched roof than small houses).

may be sensitive to a positive covariation between the relative size of stimulus components, although the stimulus displayed in the study phase instantiate a no-contingency relation for these variables!

In order to demonstrate that subjects really abstract the fuzzy covariations embedded in the study material, one definitely needs to use variables that are unrelated in the subjects' background knowledge. Pending results collected in experiments fulfilling this condition, data provided by Richardson's studies offer no compelling evidence that subjects are able to abstract covariations among feature variables.

Acknowledgements

The study was supported by the Centre National de la Recherche Scientifique and Université René Descartes.

References

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds), *Cognition and Categorization*, pp. 169–215. Hillsdale, NJ: Erlbaum.

Goldstone, R. L., Medin, D. L. & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology*, **23**, 222–262.

Medin, D. L. & Shaffer, M. M. (1978). A context theory of classification learning. *Psychological Review*, **85**, 207–238.

Richardson, K. (1986). Evidence for relational-coding in concept formation, and a note on the effects of feedback. *International Journal of Psychology*, **21**, 641–657.

Richardson, K. (1987). The coding of relations versus the coding of independent cues in concept formation. *British Journal of Psychology*, **78**, 519–544.

Richardson, K. & Carthy, T. (1989). Concept models and concept functions: Inference from partial information. *Acta Psychologica*, **72**, 81–102.

Richardson, K. & Carthy, T. (1990). The abstraction of covariation in conceptual representation. *British Journal of Psychology*, **81**, 415–438.

Received 2 January 1992; revised version received 23 June 1992