



Selectivity in Problem-Solving

Alison C. Campbell

The American Journal of Psychology, Vol. 81, No. 4. (Dec., 1968), pp. 543-550.

Stable URL:

<http://links.jstor.org/sici?sici=0002-9556%28196812%2981%3A4%3C543%3ASIP%3E2.0.CO%3B2-1>

The American Journal of Psychology is currently published by University of Illinois Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/illinois.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

SELECTIVITY IN PROBLEM-SOLVING

By ALISON C. CAMPBELL,

Department of Social and Preventive Medicine, University of Manchester

In studies of human problem-solving, apart from a few notable exceptions,¹ the Ss are typically presented with just sufficient data for problem-solution. No more than these are given, nor any fewer. A prime obstacle to solving problems in 'real life' situations, on the other hand, is selecting the data relevant to solution from among the welter of data available. If psychological studies of human problem-solving are to bear on situations outside the laboratory, they must accordingly present S with data over and above those merely sufficient for solution. In line with this logic, the present study deals with problem-situations into which supraminimal data have been introduced. These supraminimal data also open up varying numbers of false leads. By studying such problems and noting how they are coped with, some understanding may be achieved of what goes on in the real-life situation when problems are solved in the face of masses of data that are largely superfluous, irrelevant, or misleading.

It is usual to derive one's hypotheses for experimental testing from the body of literature appropriate to the field. There is, however, no body of literature directly relevant to the present research, due to the tendency in studies of human problem-solving for only such material to be presented to S as is sufficient for problem-solution. In those few studies where supraminimal data *have* been used, the type of problem-situation was vastly different from that used here. The hypotheses formulated in connection with the present study were accordingly derived from the previous code-item studies of the author.² Hypothesis 1 predicted that, holding probability

* Received for publication March 20, 1968. The research reported was done while the author was a doctoral student in the Department of Psychology at the University of Edinburgh. The author thanks Professor J. Drever and Dr. Margaret Donaldson for their advice and guidance throughout.

¹ E.g. K. Dunker, On problem solving, *Psychol. Monogr.*, 58, 1945 (No. 270); A. S. Luchins and E. H. Luchins, New experimental attempts at preventing mechanisation in problem solving, *J. gen. Psychol.*, 42, 1950, 279-297; N. R. F. Maier, Reasoning in humans, I, On direction, *J. comp. Psychol.*, 10, 1930, 115-143; Maier, Reasoning in humans, II, The solution of a problem and its appearance in consciousness, *J. comp. Psychol.*, 12, 1931, 181-194; Maier, An aspect of human reasoning, *Brit. J. Psychol.*, 24, 1933, 144-155; M. Wertheimer, *Productive Thinking*, 1945.

² A. C. Campbell, Redundant and irrelevant data in problem solving, unpublished Ph.D. thesis, University of Edinburgh, 1964.

of starting out on a wrong path constant, items with supraminimal data relevant to solution would be found more difficult (take longer to solve) than items without such data. Hypothesis 2 predicted that, holding probability of starting out on a wrong path constant, items with supraminimal data irrelevant to solution would be found more difficult than items without such data. Hypothesis 3 predicted that items would be found the more difficult the greater the probability of starting out on a wrong path.

METHOD

Problems used. The items that are used in tests of intellectual ability constitute problem-situations in miniature. A type of code item was devised into which supraminimal data, either relevant or not relevant to item-solution, could readily be introduced. The item consisted of five five-letter words listed down the page with coded versions of the words, in a different order, alongside. The order in which the coded words were listed relative to the uncoded words was decided for each item separately by reference to a table of random numbers. The coded words were incomplete but contained sufficient letters for an unequivocal solution. The code was in every case arbitrary, *i.e.* not based on any rule of alphabetical sequence, and item-solution was dependent on structural features of the sets of words. In this respect the items were nonverbal rather than verbal in nature. Take, for example, Item *i*:

F R O N D	- V - - C	
F I N A L	- - - C -	
F A D E S	- - J - V	
F L O U T	- - - J -	
F I R T H	- - - - -	

The coded letters occur in pairs. To solve the item, it is necessary to consider the positions in which these letter-pairs occur. Suppose, for example, one begins with the *V*s. *V*s occur in the first and third coded words, in second and fifth positions respectively. Hence the first and third coded words must represent a pair of words one of which has a certain letter in the second position while the other has this same letter fifth. Scrutiny of the uncoded words reveals that *V* can only stand for *L*, the first coded word being FLOUT and the third FINAL. If - *V* - - *C* represents FLOUT, *C* stands for *T*. Hence the second coded word must be FIRTH. And since - - *J* - *V* represents FINAL, *J* stands for *N*. The fourth coded word must therefore be FROND. The fifth coded word, for which no letters are given, must by default be FADES. It is the only word left. The subject is required to write the uncoded counterpart of each coded word in the spaces on the right hand side. For this item the coded letters are in toto just sufficient for item-solution. All need to be used, although not necessarily in the sequence outlined above. It may also be noted that none of the coded letters opens up any false leads, there being only one possible identification for each. Thus as far as the identification of letter-pairs is concerned, the probability of starting out on a false lead is zero.

Take now Item *ii*:

F I L C H	- X Q D -	_____
S A L V O	- - Q - -	_____
I S L E T	- - Q - D	_____
B E L O W	M - - X -	_____
F A L S E	- M Q - -	_____

Here there are two initially plausible identifications for each letter-pair and a probability of 0.50 of starting out on a wrong path. X, for example, could stand for S, the first coded word representing ISLET and the fourth FALSE. X could alternatively stand for E, with the first coded word BELOW and the fourth ISLET. To decide between these two initially plausible alternatives one must work further. There is also the added complication of letters over and above those sufficient for solution, viz. the Qs. The Qs are irrelevant to solution, since knowing that Q stands for L is of no help in deciding which coded word is which (all five uncoded words having an L in third position).

Take now Item *iii*:

R A I L S	J - W - B	_____
A R E A S	O F X W -	_____
I S L E S	F - - P B	_____
S C A N S	- J - X B	_____
M I N U S	- O P - B	_____

Here there are four initially plausible identifications per letter-pair, whichever letter-pair one begins with. Thus J could stand for R (first coded word RAILS and fourth AREAS) or for A or I or S. For items like this one there is thus a probability of 0.75 of starting out on a wrong path. This item also has coded letters *relevant* to solution, over and above the minimum sufficient. Not all six pairs of coded letters need be used; the solution is overdetermined. The item also contains supraminimal letters *irrelevant* to solution, this time in the fifth position (the Bs).

A 12-item test was constructed, in accordance with the plan set out in Table I. All items with supraminimal data not relevant to solution were analogous to Items *ii* and *iii* above in that the irrelevant data consisted of four letters all the same and all in the same ordinal position. The ordinal position in which the irrelevant letters occurred was varied from item to item at random. (It has been demonstrated previously that irrelevant letters affect item-solution similarly irrespective of their ordinal position.)³ In the items with supraminimal data relevant to item-solution, the supraminimal data always consisted of three additional letter-pairs as in Item *iii* above. Copies of the test plus the four sample items that prefaced it are available from the author.

Subjects and procedure. The Ss were eight undergraduate students in the Faculty of Arts at the University of Edinburgh. Each was tested individually. The order of the items was decided for each S separately by reference to a table of random numbers. This was to enable a Subjects \times Treatments design, allowing the assessment of interaction effects between any two or all three main effects (presence *vs.* absence of supraminimal data relevant to solution; presence *vs.* absence of supraminimal

³ Campbell, *loc. cit.*

TABLE I
 MEAN SOLUTION-TIMES AND SDs (IN SEC.) FOR EACH TEST ITEM AND FOR
 DIFFERENT GROUPS OF ITEMS ACCORDING TO AMOUNT AND KIND OF DATA AND
 PROBABILITY OF STARTING OUT ON A WRONG PATH (p)

Supraminimal data not relevant to item-solution		Supraminimal data relevant to item-solution					
		Absent			Present		
	Item	p	Mean solution time and SD	Item	p	Mean solution time and SD	
Absent	1	0.00	50.25 (SD 19.46)	7	0.00	60.50 (SD 13.23)	
	2	0.50	81.75 (SD 25.96)	8	0.50	110.13 (SD 54.80)	
	3	0.75	172.25 (SD 91.90)	9	0.75	194.62 (SD 105.20)	
	4	0.00	63.13 (SD 18.53)	10	0.00	75.25 (SD 16.71)	
	5	0.50	78.00 (SD 35.16)	11	0.50	168.13 (SD 80.50)	
	6	0.75	233.25 (SD 77.90)	12	0.75	279.37 (SD 85.11)	
Combined means	Items 1-6		113.11 (SD 85.36)	Items 7-12		148.00 (SD 104.50)	
Combined means:	Items 1, 4, 7, 10 ($p=0.00$)		Mean = 62.28; SD = 20.83				
	Items 2, 5, 8, 11 ($p=0.50$)		Mean = 109.50; SD = 64.36				
	Items 3, 6, 9, 12 ($p=0.75$)		Mean = 219.88; SD = 90.37				
Supraminimal data relevant to item-solution						Items 1-3; 7-9 111.68 (SD 83.26)	
						Items 4-6; 10-12 149.52 (SD 80.71)	
Combined means							

data not relevant to solution; and probability of starting out on a wrong path). Each item was presented to *S* on a separate page and each was separately timed. If *S* offered a solution that was incorrect, he was asked to revise it. The time for revision was added to the time for initial solution. Qualitative information was derived from observing behaviour and verbal comments in the test situation and also by questioning *Ss* on completion of the testing as to how they set about their task.

RESULTS

Mean solution-times and *SDs* for each test item and for different groups of items are shown in Table I. The results of the appropriate analysis of variance appear in Table II. In view of the lack of homogeneity of variance

TABLE II
ANALYSIS OF VARIANCE
(R=supraminimal data relevant to solution. I=supraminimal data irrelevant to solution. P=probability of starting out on a wrong path.)

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
R	1	29,225.26	29,225.26	13.48*
I	1	34,542.09	34,542.09	13.93*
P	2	418,645.77	209,322.89	42.70†
<i>Ss</i>	7	151,006.99	21,572.43	
R× <i>Ss</i>	7	15,173.99	2,167.71	
I× <i>Ss</i>	7	17,359.16	2,479.88	
P× <i>Ss</i>	14	68,625.73	4,901.84	
RI	1	5,089.59	5,089.59	2.78
RP	2	9,245.02	4,622.51	3.78
IP	2	15,356.44	7,678.22	1.90
RI× <i>Ss</i>	7	12,802.00	1,828.86	
RP× <i>Ss</i>	14	17,134.49	1,223.89	
IP× <i>Ss</i>	14	56,628.07	4,044.86	
RIP	2	3,671.69	1,835.85	1.00
RIP× <i>Ss</i>	14	25,669.45	1,833.53	
Total	95	880,175.74		

* Significant at 0.02 level.

† Significant at 0.001 level.

between items, a more rigorous level of significance was adopted in assessing the *F* ratios than is indicated by the standard tabled values. Hence while the *F* ratio for supraminimal data relevant to item-solution is shown in the tables to be significant at the 0.01 level, it was here claimed as significant only at the 0.02 level.⁴ Since items with supraminimal relevant data were significantly more difficult than comparable items without, Hypothesis 1 was confirmed. The *F* ratio for supraminimal data irrelevant to solution was also significant at the 0.02 level, and Hypothesis 2 was likewise confirmed. The *F* ratio for probability of starting out on a wrong path was significant at the 0.001 level, and *t* tests then revealed significant differences

⁴ See Q. McNemar, *Psychological Statistics*, 1962, 252.

in solution-times only between items with a probability of 0.00 and 0.75 of starting out on a wrong path ($p < 0.01$) and between items with a probability of 0.50 and 0.75 of starting out on a wrong path ($p < 0.01$), but not between items with a probability of 0.00 and 0.50 of starting out on a wrong path ($p > 0.05$).⁵ Hypothesis 3 was thus confirmed only in part. None of the F ratios for the interaction-terms was significant at the 0.05 level.

DISCUSSION

That items with supraminimal relevant (redundant) data were more difficult than those without may at first seem surprising. It is well established, for example, that redundant instances facilitate, not impede, concept-attainment.⁶ The type of redundancy used in these concept-attainment experiments does not, however, lead to an increase in the absolute amount of data present and is thus not comparable to the redundancy of the code-items. Any interpretation of item-difficulty merely in terms of the absolute amount of data present (number of coded letters) is not, however, tenable. Items *without* supraminimal relevant data, but with a high probability of starting out on a wrong path, were consistently more difficult (*i.e.* took longer to solve) than items *with* supraminimal relevant data but without false leads. It was not the sheer amount of data that was important so much as what these data were and, in particular, whether they opened up false leads. The absolute amount of data present only became crucial as a determinant of item-difficulty when the probability of starting out on a wrong path was held constant.

With regard to supraminimal data not relevant to item-solution, the literature suggests that irrelevant data may impede performance to the extent that the relevant and irrelevant stimuli are qualitatively alike and hence not readily discriminable one from the other. While the irrelevant letters in the code-items were qualitatively similar to the relevant letters in that all were letters of the alphabet, there was nonetheless a ready basis for discriminating between the two sets of letters in that the relevant

⁵ The variances for the different groups of items were not here homogeneous. It has been demonstrated, however, (see McNemar, *op. cit.*, 106) that the t test remains 'robust' under such conditions. In addition the further precaution was taken of claiming the t values significant only at the 0.01 level although the tabled values indicated them to be significant at the 0.005 level.

⁶ L. E. Bourne and R. C. Haygood, The role of stimulus redundancy in concept identification, *J. exp. Psychol.*, 58, 1959, 232-238; Bourne and Haygood, Supplementary report: Effect of redundant relevant information upon the identification of concepts, *J. exp. Psychol.*, 61, 1961, 259-260; J. Bruner, J. J. Goodnow, and G. A. Austin, *A Study of Thinking*, 1956; Haygood and Bourne, Forms of relevant stimulus redundancy in concept identification, *J. exp. Psychol.*, 67, 1964, 392-397.

letters occurred in pairs while the irrelevant letters were in fours. The irrelevant letters were also distinctive by being all in the same ordinal position. Although Ss noticed this, they tended to fail to appreciate its implications. Perhaps the irrelevant letters were 'perceptually obvious' but not 'obviously irrelevant.' One S said he always began with the four letters all the same since they were 'definitely fixed' and the 'surest of the lot,' but while this was true it was not at all helpful. Only four of the eight Ss realised that knowing the identity of irrelevant letters did not advance solution in any way.

In all the items in the present study, every letter-pair in any one item had the same number of initially plausible identifications (one, two, or four). For the items with two, and especially four, initially plausible identifications per letter-pair, the Ss tended to find this fact 'annoying.' They would start out, for example, with the quite reasonable assumption that some letter-pairs would be 'better bets' (*i.e.* have fewer initially plausible starting points) than others. That this was not so was a source of frustration. The solution of items with a number of initially plausible identifications per letter-pair and the speed of such solution depended very much on a "fortuitous factor of initial choice,"⁷ *i.e.* happening not to start out on a false lead. The more false leads S pursued prior to discovering the correct identification for any letter-pair, the longer his solution-time would be. The items conformed very nicely in fact to the notion espoused by Newell, Shaw, and Simon,⁸ that the problem-solving process consists of a search for a solution in a 'space' of possible solutions and that the fewer the number of possibilities to be searched the easier the problem and the more rapidly it can be solved.

SUMMARY AND CONCLUSIONS

Eight undergraduate students at the University of Edinburgh were presented with code-items incorporating supraminimal data both relevant and not relevant to item-solution. The Ss tested were not particularly adept at separating data relevant to the solution of a problem from data not relevant to solution. Nor were the Ss particularly adept at separating data relevant and necessary to problem-solution from data relevant but superfluous. Hence supraminimal data, other things being equal, were found signifi-

⁷ E. R. John, Contributions to the study of the problem-solving process, *Psychol. Monogr.*, 71, 1957, No. 447.

⁸ A. Newell, J. C. Shaw, and H. A. Simon, Elements of a theory of human problem solving, *Psychol. Rev.*, 65, 1958, 151-166; Newell, Shaw, and Simon, The processes of creative thinking, in H. E. Gruber (ed.), *Contemporary Approaches to Creative Thinking*, 1962, 63-119.

cantly to impede solution of the items. This was equally true both for supraminimal data relevant to solution and for supraminimal data not relevant to solution. The more data present, whether relevant or not, the more complex the item as a perceptual display. The Ss then took longer to select an appropriate starting point.

The absolute amount of data, however, was crucial only when the probability of starting out on a wrong path was held constant. An item with supraminimal data but without false leads was in fact *less* difficult on the whole (*i.e.* took less time to solve) than an item without supraminimal data but with numerous false leads. The greater the number of false leads (initially plausible starting points), the greater the probability of starting out on a wrong path and hence the more difficult the item. This was true, however, only in group terms. The items were such that a fortuitous factor of initial choice could lead to rapid success. Hence an individual, even on an item with an inordinate number of initially plausible starting points, could by chance choose the correct starting point straight off.

The present research dealt with only one type of problem-situation, albeit in some detail. The extent to which the findings can be generalised to other types of items, to problem-situations other than test items, and to problem-situations beyond the confines of the laboratory can only be settled empirically, by further research. It may be concluded nonetheless that the findings of the present study, in relation to amount and kind of data and the presence of false leads, threw at least some light on some general aspects of problem-situations that pose difficulties to the problem-solver.