Categorization, attitude and memory for the source of attitude statements

J. RICHARD EISER
University of London
JOOP VAN DER PLIGT
University of London
MICHAEL R. GOSSOP
Bethlem Royal and Maudsley Hospitals, London

Abstract

In a study of factors influencing recognition-memory for the sources of attitude statements, a final sample of 107 subjects, aged 15–16, first rated their agreement with 24 statements concerning drug-use, 12 of which were attributed to one, and 12 to another, fictitiously named newspaper. Later, the statements were resown to subjects with half the names altered, and subjects had to indicate which names were correct (i.e., unaltered). Discrimination sensitivity was very significantly higher in a condition where the initial relationship between the sources and the statements was systematic, so that the 12 most pro-drug statements were attributed to one newspaper and the 12 most anti-drug statements to the other, than in two conditions where the initial relationship was random, in which discrimination was at chance level. In the first of these conditions, subjects were also more likely to claim that the attributed source was correct if they had previously agreed with the statement. Overall, subjects were more accurate in discriminating correct and incorrect sources for statements to which they had previously given a more moderate, or a more negative response on the agreement scale.

INTRODUCTION

One of the most basic assumptions of cognitive psychology, and of the applications of cognitive theory in social psychology, is that individuals will attempt to simplify stimulus information through the process of categorization (Bruner, Goodnow and Austin, 1956). Stimulus sequences which can be ‘chunked’ (Miller, 1956) or which suggest some apparently meaningful framework (e.g., Epstein, 1961) are more easily recalled than disconnected sequences of the same length. Stimuli which fall into distinct classes may also be judged as more different from each other in terms of their perceived positions on some attribute than stimuli of equivalent magnitude on which no such classification is superimposed (Tajfel, 1959; Tajfel and Wilkes,
A necessary condition for this latter effect appears to be that the superimposed classification should be systematically related to the attribute being judged (Eiser and Stroebel, 1972; Tajfel, 1959) so that it may provide an aid to discrimination.

Within social psychology there is now considerable evidence that superimposed classifications increase discrimination in judgement tasks. This classification may take the form of group membership either of hypothetical stimulus persons (Tajfel, Sheikh and Gardner, 1964), or of others with whom the subject believes he is interacting (Tajfel, 1978); or, in attitudinal judgements, it may take the form either of the supposed source of attitude statements or of the subjects’ own levels of agreement with the statements (Eiser, 1971; Eiser and Mower White, 1975). For example, subjects in the Eiser (1971) study rated a series of statements in terms of their permissiveness or restrictiveness on the issue of the non-medical use of drugs. When the more permissive statements were presented as coming from one newspaper and the more restrictive statements from another, subjects’ ratings were more polarized (subjects discriminated more between the statement groups) than in a control condition when subjects were not given information about the supposed source of the statements. Depending on the response language used, subjects may also discriminate in their judgements between statements they find acceptable and unacceptable.

The question of how subjective classifications influence memory and recognition, has received rather less attention from social psychologists over recent years. Early research suggesting an influence of attitude on recall of attitude statements (Levine and Murphy, 1943; Jones and Kohler, 1958) has been disputed (Greenwald and Sakumura, 1967). Related research on the influence of ethnic attitudes on a person’s ability to identify another ethnic group membership (Lindzey and Rogolsky, 1950) has generally confounded discrimination sensitivity (the ability to assign stimuli correctly to different categories) with response bias (the tendency to assign more stimuli to one particular category regardless of their specific features). An important exception in the ethnic identification field is the study by Dorfman, Keeve and Saslow (1971), which found a relationship between anti-Semitism and sensitivity of discriminating Jewish from non-Jewish facial photographs.

Two more recent studies have investigated the influence of subjects’ attitudes on their recognition memory for attitude statements whilst distinguishing sensitivity and response bias effects. Eiser and Monk (1978) had subjects rate a series of attitude statements for agreement, and then replaced some of these with slightly reworded statements from a parallel version of the attitude scale. When asked to say whether the wording of the statements had or had not been changed, subjects were more likely to claim that the wording of a statement was correct (‘same as before’) if they had previously agreed with the statement, and were more accurate in discriminating changes from original wordings in statements to which they had given a moderate rather than an extreme agree or disagree response. Upmeyer and Layer (1974) presented subjects with favourable and unfavourable statements describing a politician whom they evaluated favourably (Brandt) or unfavourably (Barzel or Strauss). The politicians’ names were then changed on some items, and subjects were required to identify any such changes. The use of politicians’ names improved discrimination as compared with a control condition where the labels X and Y were used instead. There was evidence of an effect of subjects’ attitudes on
response bias, in that subjects were more likely to recall unfavourable than favourable statements as being characteristic of Strauss rather than Brandt. According to Upmeyer and Layer, such ‘output accentuation’ depends on a ‘rule of correspondence’, i.e. evaluative and descriptive consistency between the content of the statements and the superimposed labels.

Since such evaluative and descriptive consistency appears a necessary condition of judgemental accentuation (Eiser and Mower White, 1974, 1975), there would seem a case for attempting to draw the two research paths (judgement and recognition-memory) closer together. In the judgement tradition, it is typically argued that superimposed classifications provide an aid to discrimination, i.e. they help to reduce cognitive overload through highlighting the redundancy inherent in stimulus information. Upmeyer and Layer argue that the main effects of superimposed classifications on recognition-memory operate at the level of output rather than input. However this argument seems to run counter to the implications of research on ‘chunking’, and is not in fact established by their study, which does not involve a direct comparison to determine subjects’ recognition-memory for a superimposed classification under conditions when it is either systematically or randomly associated with a discriminable attribute of the stimuli presented.

In judgement tasks, a superimposed classification systematically related to the positions of the stimuli along the dimension of judgement leads typically to an accentuation of judged interclass differences, whereas one only randomly related to that dimension does not (Eiser and Stroebe, 1972; Tajfel, 1959). The present study therefore tests the hypothesis that a systematic classification will be more accurately discriminated in a recognition-memory task than will a random classification, using attitude statements as stimuli. Also, this study investigates whether subjects may be more prepared to accept as correct an association between a stimulus series and a superimposed classification which is systematic rather than random.

An important difference from the Upmeyer and Layer (1974) design is the fact that the classification used does not, as in their study, have established associations of value for the subjects. The question of subjects’ own attitudes therefore is confined to their evaluation of the statements, and the issue of their evaluations of the superimposed labels does not arise. In this respect, the present study provides an opportunity for seeing how far the Eiser and Monk (1978) results can be generalized to a situation when what subjects have to remember is not the precise wording of the statements, but the sources to which they are attributed.

**METHOD**

A total of 112 pupils, aged 15–16 years, from two comprehensive schools in S.E. London served as subjects, of whom five were excluded for incomplete responses. Approximately two-thirds of the subjects were girls. They were administered a question during a normal classroom period for voluntary and anonymous completion (some two or three additional pupils declined to participate after seeing the questionnaire). The questionnaire was in two separate parts, and was handed to subjects with part 2 enclosed inside a large envelope and part 1 clipped to the outside. Subjects were instructed to complete part 1 before taking out part 2, both parts being put back into the envelope on completion.
Part 1 of the questionnaire contained 24 statements expressing ‘different opinions about the use of drugs, whether cannabis should be legalized, etc.’. Subjects were told: ‘These different opinions are often published in newspapers, in articles written by journalists, or in letters from members of the public. We have collected 24 such statements of opinion from two different newspapers. We can’t tell you the real names of the newspapers, since this might bias your answers, so we have made up a fictitious name for each newspaper. We have called one of the newspapers The Gazette and the other The Messenger. Underneath each statement you will find the “name” of the newspaper in which it appeared’. They were then instructed to rate their own agreement with each statement in terms of a 100 mm linear scale from ‘I disagree very much’ (scored as 0) to ‘I agree very much’ (scored as 100). They were further instructed ‘Please pay attention to the newspapers in which the statements appeared, as we’ll be asking you about your impressions of the two newspapers later on’. No more explicit warning was given that their recall for the names would be tested, nor that there could be a systematic relationship between the favourability of the statements and the newspaper names.

Part 2 consisted first of a rating task relevant to the attitude issue but not to the hypothesis examined in this report, and fulfilled the function of an intervening task which took most subjects about five minutes to complete. One question provided a self-rating measure of their own attitude ‘towards the use of drugs generally’ on a linear scale from ‘extremely opposed to drug use’ (scored as 0), to ‘extremely in favour of drug use’ (scored as 100).

After this intervening task, subjects read a new set of instructions which explained that ‘we want to find out how well you remember which statements came from which newspaper. On the next few pages, you will find the same 24 statements as before, with one of the two newspaper “names” printed underneath each statement. In some cases we have deliberately changed the “names” so that some of the statements that really come from the The Gazette are now presented as coming from The Messenger, and vice versa. What you have to do is to spot the deliberate mistakes’. Subjects then rated whether the ‘name’ under each statement was definitely wrong, probably wrong, probably correct, or definitely correct (correct meaning ‘same as before’).

The experimental manipulation consisted of the ways the ‘names’ were attached to the statements in the learning and recall phases of the experiment. Three versions of the questionnaire were distributed at random. In condition SR (N = 35) the association between the names and the favourability of the items towards drug use (based on pilot ratings) was systematic in part 1 and random in part 2. That is, in part 1, the 12 most pro-drug items were attributed to The Gazette and the 12 most anti-drug items to The Messenger, and in part 2, each name was attached to 6 pro and 6 anti items at random (so that each name was ‘correct’ 50 per cent of the time). In condition RR (N = 38) both parts had 6 pro and 6 anti items assigned to each ‘newspaper’ at random, subject to the restriction that 50 per cent of the part 2 names were ‘correct’. Part 2 was identical in conditions SR and RR. Finally, condition RS (N = 34) was the reverse of the first condition, with a random association of names and favourability in part 1, and a systematic association in part 2, where again 50 per cent of the names were ‘correct’.

Our main hypothesis regarding the effects of this manipulation was that recognition memory for the original names paired with each statement would be more
accurate in condition SR than in the other two conditions. More speculatively, it was predicted that a greater percentage of names would be judged as correct (there would be a ‘response bias’ effect) in condition RS than in condition RR. In addition, extrapolating from the Eiser and Monk (1978) results, it was predicted that, over all conditions, subjects would be more likely to judge names as correct if they agreed with the statements, and that discrimination would be more accurate for items previously receiving less extreme ratings on the agreement scale.

RESULTS

Treatment of results

Recognition accuracy (discrimination sensitivity) was measured from subjects’ part 2 ratings by the $A$-Index, or $\hat{A}$ (Brown, 1974; Brown and Routh, 1970; Eiser and Monk, 1978). In fact $\hat{A} = 2R - 1$, where $R$ is the likelihood that any target (correct name) will be judged as more correct than any distractor (incorrect name) and is thus the area under the ROC curve (McNicol, 1972). $\hat{A}$ takes a value of +1 for perfect discrimination (all targets rated as more correct than all distractors), 0 for discrimination at chance level, and minus scores, with a minimum of $-1$, for discrimination at worse than chance.* In addition, two further scores were derived from subjects’ part 2 ratings: a confidence score, representing the proportion of items rated as either definitely wrong or definitely correct, and a bias score, representing the proportion of items rated as probably correct or definitely correct.

In addition to subjects’ self-ratings of their own position, two measures of attitude were derived from their part 1 ratings. The first was a Likert-type score (mean agreement for pro minus anti items), and the second the P(A) measure proposed by Monk and Eiser (in press), representing the consistency of preference for pro over anti items.†

Attitude differences between conditions

There were no significant attitude differences between conditions. The means of P(A), which can range from 0 (extremely anti) to 1 (extremely pro) with a no preference midpoint of 0.5, were 0.51, 0.45, and 0.50 in conditions SR, RR and RS respectively. Overall, P(A) correlated 0.946 with the Likert score and 0.202 with the self-rating: the Likert–self-rating correlation was 0.186.

Effects of the experimental manipulation

Table 1 presents the means for the three conditions of $\hat{A}$, the absolute value of $\hat{A}$ (i.e. averaged without regard to sign), the measure of confidence, and the measure

*Stated generally, if one is using a rating scale containing $K$ categories, from 1 (item ‘definitely wrong’) to $K$ (item ‘definitely right’), then an individual subject’s $R$ score is given by the following formula:

$$ R = \sum_{i=2}^{K} \left( \frac{t_i - 1}{K - 1} \right) + \frac{1}{2} \sum_{i=1}^{K} td_i, $$

where $t_i$ is the proportion of the total number of targets rated in category $i$, and $d_i$ if the proportion of the total number of distractors rated in category $i$.

†P(A) is calculated in the same way as $R$ in the above formula, substituting ‘pro’ items for targets, and ‘anti’ items for distractors, and assuming one’s rating categories are ordered from least to most agreement.
Table 1. Mean values of $\hat{A}$ (discrimination sensitivity), $|\hat{A}|$ (absolute value of $\hat{A}$, regardless of sign), confidence and response bias measures for subjects in each condition

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Planned comparisons (F with df = 1,104)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR $N = 35$</td>
</tr>
<tr>
<td>$\hat{A}$</td>
<td>0.356</td>
</tr>
<tr>
<td>$</td>
<td>\hat{A}</td>
</tr>
<tr>
<td>Confidence</td>
<td>0.427</td>
</tr>
<tr>
<td>Bias</td>
<td>0.525</td>
</tr>
</tbody>
</table>

*p < 0.001
Note: SR = classification systematic in learning phase, random in recall phase; RR = classification random in both phases; RS = classification random in learning phase, systematic in recall phase.

of bias. One-way analyses of variance were conducted on these means, with planned comparisons on conditions SR vs. RR and RS, and on RR vs RS. As can be seen, our main hypothesis, that discrimination would be highest in condition SR, was very significantly confirmed. Discrimination in conditions RR and RS was only at chance level. This effect was even clearer with $|\hat{A}|$, suggesting that a number of subjects in condition SR may have remembered that there was *some* association between the newspaper names and item favourability, but have forgotten which ‘newspaper’ was the more pro-drug. Our secondary hypothesis, that condition RS should lead to a higher bias score than condition RR, was not confirmed. None of these four dependent variables correlated significantly with any of the three measures of subjects’ overall attitudes.

Effects of previous agreement

We next compared, over all conditions, the mean part 1 agreement scores for items subsequently discriminated correctly by each subject with those for incorrectly discriminated items. Those items which were correctly discriminated in part 2 (disregarding the probably-definitely distinction) received lower levels of agreement (mean = 48.3) than those which were incorrectly discriminated in part 2 (mean = 55.8, $t = 3.58$, df = 106, $p < 0.001$). Recognition accuracy was thus associated with item unacceptability. Subjects also tended to agree somewhat less with items to which they gave a confident (i.e., ‘definitely’) response in part 2, regardless of whether that response was correct, than with items receiving a less confident (i.e., ‘probably’) part 2 response (means 45.6 and 49.3 respectively, $t = 1.81$, $p < 0.08$). Overall, there was no significant relationship between previous agreement and response bias, with the mean agreement rating being 51.5 for items attracting a ‘definitely correct’ or ‘probably correct’ part 2 response, and 49.6 for the remaining items ($t = 1.42$, ns). However in condition SR, this tendency was significant, with the means being 58.2 and 49.3 respectively ($t = 2.61$, df = 34, $p < 0.02$). This gives weak confirmation to the first of our predictions based on the Eiser and Monk (1978) results.
Effects of extremity of previous agreement

Next, instead of computing the mean agreement scores for different item categories, the mean absolute deviation of these scores from the midpoint of the scale (50) was calculated for each subject. Items which were discriminated correctly in part 2 tended to receive less extreme ratings in part 1 (mean extremity = 27.8) than those incorrectly discriminated (mean = 35.4, $t = 4.72$, $p < 0.001$). This confirms our second prediction based on the Eiser and Monk (1978) results. Greater confidence in part 2 was also associated with less extreme ratings in part 1, with items responded to confidently receiving a mean (part 1) extremity score of 27.1 and the remaining items a mean of 29.6 ($t = 2.19$, $p < 0.05$). Response bias was unrelated to extremity of agreement, with the means being 30.3 and 30.1 ($t = 0.44$) for items attracting part 2 responses of ‘correct’ and ‘wrong’ respectively.

DISCUSSION

This study shows clearly that subjects' recognition memory for the sources of attitude statements (when not warned that their memory would be tested) is significantly higher when these sources bear a systematic relation to the positions of the statements for or against the attitude issue. Where the association between source and item favourability is random, subsequent discrimination performance is at chance level. This result is independent of response bias and of response confidence, neither of which differed significantly between the experimental conditions. Thus, the kind of relationship between item positions and a superimposed classification which leads to greater accentuation of interclass differences in a judgement paradigm (Eiser, 1971) leads also to improved discrimination sensitivity in a recognition-memory paradigm. This suggests that, in both situations, the process of categorization simplifies the cognitive task with which subjects are faced.

The influence of subjects' own levels of agreement with the statements (on the particular responses required in this study) is more complex. Confirmation was obtained for the Eiser and Monk (1978) finding that subjects would be more likely to claim that an item was correct if they had previously agreed with that item, but this was the case only in condition SR, where subjects' discriminations were better than chance. On the other hand, the Eiser and Monk finding was potentially confounded by the fact that subjects in that experiment made identical motor responses when recording their agreement with the items and their judgements that the items were correct. The present result is not open to this objection. Confirmation was also obtained for the Eiser and Monk finding that items which were more accurately discriminated tended previously to have received more moderate responses on the agree–disagree scale. This was a surprising result, in view of the assumption that extreme stimuli are more ‘distinctive’ (e.g. Helson, 1964). Eiser and Monk recorded response latencies (which was impossible with our classroom procedure) and found longer latencies for moderate as opposed to extreme agree–disagree ratings as well as better discrimination for the moderate items. This suggested that items attended to for longer were better recognized later on. Such attentional differences might also possibly underlie the additional finding of this study, that recognition accuracy was related to item unacceptability. This was not
found by Eiser and Monk, although disagree responses had longer latencies. In the present study, however, there was a wider and more even spread of subjects’ attitudes, so that the effects of item acceptability are less confounded, overall, with the favourability of items towards the issue.

Previous papers (Eiser, 1971; Eiser and Mower White, 1975) have argued for a parallelism between the effects of experimentally superimposed classifications and of subjects' own categories of acceptability–unacceptability on judgements of attitudes. The present study is not specifically designed to establish that there is as close a parallelism between their effects in the recognition-memory paradigm. (We did not test, for instance, how well subjects could discriminate acceptable from unacceptable items). However, our results do suggest that both superimposed classifications and subjects' acceptance-rejection of statements can influence recognition accuracy, independently of response bias.

REFERENCES


**RÉSUMÉ**

On a étudié quelques-uns des facteurs qui influencent la mémoire (la reconnaissance) des sources des énoncés d’attitude.

Une population de 107 sujets de 15–16 ans donnaient dans un premier temps leur degré d’accord avec 24 énoncés portant sur la consommation de la drogue; on disait que 12 de ces énoncés venaient d’un journal au nom fictif et les 12 autres d’un autre journal, fictif également. Dans un deuxième temps on représentait les énoncés, la moitié des noms de journaux étant permutés. Les sujets devaient alors dire quels étaient les titres corrects ou inchangés.

La reconnaissance était meilleure lorsque la relation entre les titres et les énoncés était systématique: les énoncés favorables à la drogue étaient imputés à un journal et les énoncés défavorables à l’autre journal. Dans deux autres conditions où l’on apparaissait les items et les titres de manière aléatoire la reconnaissance n’était pas supérieure au hasard. Dans la première condition les sujets avaient aussi tendance à dire que la source présumée était la bonne si ils étaient d’accord avec l’énoncé. Dans l’ensemble les sujets étaient plus précis (reconnaissance de l’appariement initial) pour les énoncés avec lesquels ils étaient modérément d’accord ou en total désaccord.

**ZUSAMMENFASSUNG**

Untersuchte, welche Faktoren das Ausmaß beeinflussen, in dem Quellen von Attitüdenaussagen wiedergekennnt werden. Vpn einer Stichprobe (Gesamt- N = 107) im Alter von 15–16 Jahren, beurteilen zunächst ihre Übereinstimmung mit dem Gehalt von 24 Aussagen zum Drogengenuß, von denen je 12 jeweils einer von zwei erfundenen Zeitungen zugeschrieben wurden. Später wurden die Aussagen den Vpn wieder vorgelegt, wobei die Hälfte der Namen der Quelle vertrauscht wurde; die Vpn mußten angeben, welche Zeitungsnamen richtig (d.h. unverändert geblieben) angegeben waren. In der Bedingung, in der die anfängliche Beziehung zwischen den Quellen und den Aussagen in dem Sinn hochsystematisch war, daß die 12 eher den Drogengenuß befürwortenden Aussagen der einen, die 12 eher ablehnenden der anderen Zeitung zugeschrieben wurden, war das Auflösungsvermögen hochsignifikant besser als unter den zwei Bedingungen, in der die Beziehung zufällig war und das Auflösungsvermögen auf Zufallsniveau lag. In der ersten Bedingung neigten die Vpn eher dazu, die angegebene Quelle als richtig zu bezeichnen, wenn sie zuvor mit der Aussage übereingestimmt hatten. Insgesamt konnten die Vpn richtige und falsche Quellen bei solchen Aussagen besser unterscheiden, denen sie eine mittlere oder eher negative Beurteilung gegeben hatten.

Authors Address: