

CLASSICAL CONDITIONING OF HUMAN 'EVALUATIVE' RESPONSES

A. B. LEVEY

Medical Research Council Clinical Psychiatry Unit, Graylingwell Hospital, Chichester,
Sussex, England

and

IRENE MARTIN

Institute of Psychiatry, University of London, England

(Received 15 October 1974)

Summary—Human behaviour includes an important component which may conveniently be called evaluative. Using postcard reproductions as stimulus materials, 10 volunteer subjects selected the two pictures most liked and the two most disliked. These were then used as UCSs with appropriate controls from the neutral category. The conditioning hypothesis—that a neutral stimulus followed by a positively or negatively valued stimulus will acquire the evaluative weight of the second stimulus—was supported at a highly statistically significant level. The effect of negative evaluation was demonstrably stronger than that for positive evaluation, a result consistent with our knowledge of aversive conditioning. The possibility is discussed that evaluation of the UCS by the subject, shown to be a sufficient condition for learning, may also be the only necessary condition. This would imply a model of conditioning based on affective evaluation rather than on response production.

Contemporary views are critical of the dominating zoöcentric influences on conditioning theory, and of their limited applicability to human behaviour. They question the assumption that the muscle and gland responses of conditioning studies are necessarily more basic and reliable than complex human behaviour. Grant (1964) argues that such characteristically human behaviours as verbal and cognitive activity are in fact more reliable, systematic and predictable than the socially and behaviourally irrelevant primitive reflex behaviours.

Traditionally the effects of cognitive factors have been studied in terms of how verbal instructions, set, cognitive disparity, etc. interfere with (facilitate or depress) underlying conditioning phenomena. Unfortunately they have also become linked with the problem of 'awareness' and its relationship to verbal report. Awareness of CS-UCS relationships in autonomic conditioning is one current controversy (Grings, 1973; Lockhart, 1973); awareness of the demand characteristics of the experimental situation in attitude conditioning is another (Insko and Oakes, 1966). The 'troublesome variables' are for the most part subject variables which are not definable in terms of stimulus manipulations (Grings, 1965) and which, in the case of verbal reports of 'awareness', give rise to considerable measurement difficulties (Page, 1969). Preoccupation with such cognitive-verbal problems has tended to deflect interest from the simpler, more reliable assessment of *affective* experience.

Of potentially greater relevance to human conditioning studies is the use of verbal reports as summaries or representations of natural affective reactions to active stimuli. A characteristic human reaction to a stimulus or stimulus complex is to *evaluate* it as good-bad, dangerous-safe, desirable-undesirable, etc. at a level which is typically (1) well removed from strong emotion (intense autonomic response) and (2) well before actual motor behaviour (real avoidance or approach). This evaluative response to a situation can be differentiated both from its concomitants (e.g. physiological change) and from emotional behaviour. It is frequently described as a direct, unmediated type of affective rather than cognitive appraisal: "an object or situation is perceived, appraised, and liked or disliked" (Arnold, 1960); "the issue is how much am I in danger from a situation . . . the process is often nearly instantaneous. . . and an individual need not be fully aware of the evaluations he is making, or of the factors that enter into them. . ." (Lazarus, 1966). Young

(1967) has repeatedly drawn attention to the role of affective processes in motivation and learning, maintaining the view that behaviour is regulated by the subjective experience of pleasantness/unpleasantness, i.e. a hedonic dimension. The human evaluative process can be assumed to involve a wide range of mechanisms, perceptual, emotional, experiential, symbolic, etc., operating within such a dimension.

The evaluative response may be construed as a critical response to stimuli. As such, a conditioned evaluative response would fit the suggestions of Grings (1973) and Levey and Martin (1966) that the CR 'prepares' for responding, and accomplishes part of the 'work' of the motor or glandular response which it can be assumed to precede. It is possible that in general human experience the UCR for a noxious UCS may always be evaluative: a non-verbal response the verbal equivalent of which would be "this is bad, dangerous, hurtful, to be avoided". This would imply that the conditioning of evaluative reactions is a fairly basic human experience. Certainly it must be more common than terror and aggression or flight—the models of zoöcentric conditioning theory. At the same time it avoids, as a dependent variable, some of the difficulties of cognitive behaviour *per se*, though it must involve dependent cognitive components. Subjects are readily able to make reports of the kind "I like it" or "I dislike it", and many years of study have shown these forms of self report to be thoroughly reliable and useful (Osgood, 1971; Young, 1967). Modern scaling techniques have shown evaluative behaviour to be systematic and reproducible. The human evaluative response, therefore, appears to fulfil the requirements of a dependent variable in conditioning studies.

The relevance of the classical conditioned response to human behaviour has of course been gradually extended from simple reflex acts to more complex and specifically human forms of responding. The classical conditioning technique has been used, for example with noxious or aversive stimuli to induce parallel changes in the semantic significance of neutral stimuli (Staats and Staats, 1957) as well as changes of attitude and opinion (Razran, 1938; Staats and Staats, 1958). Underlying these and similar conditioning studies however, is the implication that what may be transferred to the neutral CS is simply the subject's own evaluation of the UCS, rather than a specific physiological reaction or behavioural act. We therefore undertook the following experiment to determine whether the simplest form of evaluative response, 'like' or 'dislike', can be directly conditioned to a neutral stimulus (CS) in the absence of a physiologically provocative stimulus (UCS).

METHOD

Subjects

The subjects (Ss) of this study were 10 normal adults, 5 males and 5 females, age range 21–40 yr recruited as unpaid volunteers and unaware of the purpose of the experiment. They were tested individually.

Stimulus materials

Stimulus materials were 50 unfamiliar postcard reproductions (10.5 × 15.0 cm) of paintings and scenic photographs, in colour and black and white, selected to offer a wide range of subjective preference.

Design and procedure

Each S was first instructed to sort the pictures into three categories: 'liked', 'neutral', or 'disliked'. He was given a second opportunity to remove from the neutral pile any further pictures for which he felt even minimal liking or disliking. This was done in order to increase the stability of the scale. He was then asked to choose from each of the appropriate categories the two pictures most liked and the two most disliked. From the S's neutral category the experimenter then chose a matching picture for each of the most liked and most disliked stimuli, on the basis of the similarity of form, content and colour. On the same basis the two neutral pictures which most closely resembled the experimental stimuli were chosen as a control pair. The matching of stimuli was intended to ensure that differences in complexity and discriminability known to influence the liking of unfamiliar

stimuli (Zajonc *et al.*, 1972) would not affect the results. The resulting five pairs were arranged in the following sequences:

NEUTRAL-LIKED, NEUTRAL-DISLIKED, NEUTRAL-NEUTRAL, DISLIKED-NEUTRAL, AND LIKED-NEUTRAL.

In the experiment proper, the *S* was seated at the viewing aperture of a three-field tachistoscope and given a push-button switch which triggered exposure of a pair of pictures in rapid succession. The *S* was instructed to press the switch whenever he wished to view the pictures. After each series of twenty self-presentations the stimulus pair was replaced with a new pair, until all five pairs had been viewed. The order of presentation of the pairs was determined by one of two Latin squares to which *Ss* were assigned at random. For Square I the durations of the stimulus exposures were 300 msec followed without overlap by 1000 msec, and for Square II 1000 msec followed by 1000 msec, again without overlap. The illumination of the viewing field was kept constant between exposures at the level at which stimuli were viewed.

Following the stimulus presentations, the *S* was asked to sort the ten pictures into two equal categories, 'liked' and 'disliked'. This forced choice technique was used in the expectation that experimental effects might be fairly weak. In addition, however, the *S* was also asked to arrange the picture in order of preference, and then to assign each of them a rating from +100 (maximum liking) to -100 (maximum disliking) within his own subjective range. This type of scaling is readily acceptable to untrained *Ss* and is able to yield normal variate distributions which were observed in the present study. Finally the experiment was discussed with each *S* on an open-ended basis in an attempt to determine his perception of its purpose.

The conditioning hypothesis requires that a neutral stimulus followed by a positively or negatively valued stimulus will acquire the evaluative weight of the second stimulus. Since backward conditioning is known to be relatively ineffectual, a second hypothesis was that a neutral stimulus preceded by a positively or negatively valued stimulus will show no consistent change. The control pair, neutral followed by neutral, was of course expected to display only random effects.

RESULTS

The results of the experiment are presented in Table 1, which shows the average ratings assigned to the first stimulus of each of the five pairs. Also shown in the table is the frequency of assignment of these stimuli to liked (+) or disliked (-) categories. The difference among ratings was shown to be statistically significant by analysis of variance ($F_{4,24} = 40.29, p < 0.001$) and the least significant difference required at the 0.001 probability level was 14.5, indicating that all of the means differed significantly from one another. There were no significant effects due to order of presentation or duration of stimuli.

Table 1 also gives the average post-test ratings assigned to the second stimulus of each pair. These values show that originally liked and disliked pictures differed significantly from one another and from the three neutral stimuli. Of these latter, the value of -29.8

Table 1.

Average ratings and preference assigned to the five stimuli presented first in each pair (N = Neutral, D = Disliked, L = Liked)					
Rating	NL	ND	NN	LN	DN
Preference	16.0	-31.6	31.5	69.0	-71.4
+	8	1	10	10	0
-	2	9	0	0	10
Average ratings assigned to the five stimuli presented second in each pair					
Rating	NL	ND	NN	LN	DN
	+73.0	-74.0	18.1	-1.8	-29.8

obtained following DISLIKED-NEUTRAL pairing is the highest, and while not statistically significant from the other neutral ratings it does suggest a weak backward conditioning effect occurring in some but not all Ss.

The results of interest to the conditioning hypothesis are the ratings of the first stimulus of the NEUTRAL-LIKED pair and the NEUTRAL-DISLIKED pair. The effect of negative evaluation was clearly stronger than that for positive evaluation, and this is consistent with our knowledge of aversive conditioning. It can be seen from Table 1 that the average ratings of the liked and disliked stimuli presented second in these pairs (+73.0 and -74.0 respectively) are comparable to the ratings of liked and disliked stimuli presented in the first position, and indicate that the difference was not due to changes in the evaluation of the active stimuli themselves. Thus the classical conditioning hypothesis was clearly supported.

No S was aware of consistent relationships between or among the stimuli following paired presentation. In general they were aware that some change had occurred in their ratings, and one or two commented on these changes with surprise. For the most part they attributed such changes to the effects of contrast, almost as if they were subscribing to an adaptation level type hypothesis. Yet it is interesting that contrast effects not only do not explain the results obtained, but if applied consistently would yield quite different results. Thus the Ss as a group were unaware of the exact nature of the changes induced by classical conditioning, and, further, tended to misremember or misrepresent them.

The result for the NEUTRAL-NEUTRAL pair, intended as a control, was entirely unexpected. Indeed the first stimulus of this pair showed the strongest experimental effect, while the second stimulus elicited only random responses. This result does nothing to undermine the conditioning results, which depend on the differentiation of positive and negative evaluations of the neutral stimuli, and to a lesser extent on the absence of consistent backward conditioning. It has long been known that repeated presentations of unfamiliar neutral stimuli result in increased liking (Meyer, 1903). This explanation does not apply in the present case, however, since the second stimulus of the pair showed no change. Nor can it readily be attributed to a conditioning effect, for the same reason. It is perhaps best left for the present as an intriguing puzzle.

Finally, it was possible to re-test 7 of the original 10 Ss about 18 months later, using the same post-test procedure, i.e. forced dichotomy, ranking and rating. Only the results for transfer of negative evaluation to the neutral stimuli remained significant for this group, though the positive evaluation was in the original direction. Subjects tended to remember having previously seen 3 or 4 of the stimulus cards from their set but the items remembered bore no systematic relationship to the conditioning effect.

DISCUSSION

The present results have shown that a non-verbal evaluative response, unique to the individual concerned, can be reliably conditioned in a few trials. Thus they support an early suggestion by Razran (1954) that what is conditioned in human subjects is 'general affectivity'. The idea that an affect can be transferred through conditioning is novel so far as direct empirical testing is concerned, though highly concordant with popular impression. Mowrer (1939) and Miller (1948) were among the earliest workers to introduce affective experience, especially anxiety and frustration, as necessary components of the conditioning process. The major concomitants of emotions were held to be physiological change, yet by general agreement autonomic change merely identifies the presence of affect, not whether such affect is positive or negative, nor its specific quality. An emotion cannot be inferred from physiological arousal alone. It is doubtful to what extent theories of conditioning and learning can be relevant to human emotions and behaviours until the phenomenon of subjective affective experience is considered in their formulations.

The present conditioning results suggest that the CS has acquired some of the stimulus characteristics of the active stimulus, and this is consistent with a model of the classical conditioning process outlined elsewhere (Martin and Levey, 1969). Obviously there are many stimulus characteristics which can be shared or transferred in the conditioning situation. The proposition that the CS comes to share the characteristics of the UCS and *vice*

versa can be more closely tested when it is possible to specify the term 'characteristics of' with greater precision. This testing is possible within the present paradigm since many studies point to the existence of at least two and more likely three dimensions of affective appraisal. Generally these are interpretable in terms of the semantic differential factors—evaluation, potency and activity (Osgood, 1971). Of these three factors, evaluation is typically the strongest and occurs in many data analyses.

Attempts have been made to delineate other relevant response dimensions in therapeutic situations. In the treatment of alcoholics, for example, the assumption that a CR of 'anxiety' would be expected when electrical aversive stimuli are employed in the conditioning process is being questioned (Rachman and Teasdale, 1969). Other dimensions seem to be relevant, e.g. harmless/dangerous, repulsive/distasteful (Hallam, Rachman and Falkowski, 1972; Costello, Rice and Schoenfeld, 1974), while in the treatment of sexual deviations the transfer of an attractive/repugnant evaluation is more appropriate (Beech, Watts and Poole, 1971).

Theoretically, *response production* has been seen as a necessary condition for the conditioning model, and animal investigators have by convention looked for an overt response such as a reflex or an operant. As an alternative model, we would suggest that *evaluation* rather than response production is the necessary condition. Of course, many evaluations would carry with them their own associated responses. Cues that provoke evaluation may, for example, be of three types:

- (i) Those having a 'pre-wired' reflex reaction (e.g. withdrawal). In Seligman's terminology, these would be biologically prepared responses. (Seligman and Hager, 1972).
- (ii) Those with a learned or unlearned hierarchy of responding attached to them.
- (iii) Those with no detectable responses attached to them.

Thus it is possible to propose a logically distinct model of conditioning in which the emphasis is no longer on the specific motor or autonomic response as 'the UCR' but on the process of affective appraisal which we have called the evaluative response. Such a dependent variable has many advantages: it is characteristically human in the sense that it is complex, adaptive, stable and *individual*, the latter being the distinguishing feature of explicitly human behaviour, i.e. idiosyncratic as well as predictable. If, as a consequence of evaluation, the subject were to undertake some consistent action, that action would also occur to the neutral stimulus and would be regarded as a conditioned response. One of the implications of such a view of conditioning would be that it allows for the plasticity of responding which is characteristic of human behaviour.

REFERENCES

- ARNOLD M. (1960) *Emotion and Personality*. Columbia University Press, New York.
- BEECH H. R., WATTS F. and POOLE A. D. (1971) Classical conditioning of a sexual deviation: A preliminary note. *Behav. Therapy* **2**, 400-402.
- COSTELLO R. M., RICE D. P. and SCHOENFELD L. (1974) Measurement of subjective responses to alcohol and non-alcohol slides by alcoholic respondents. *Behav. Res. & Therapy* **12**, 35-40.
- GRANT D. A. (1964) Classical and operant conditioning. In: *Categories of Human Learning* (Ed. MELTON A. W.). Academic Press, New York.
- GRINGS W. W. (1965) Verbal-perceptual factors in the conditioning of autonomic responses. In: *Classical Conditioning* (Ed. PROKASY W. F.). Appleton-Century-Crofts, New York.
- GRINGS W. W. (1973) Cognitive factors in electrodermal conditioning. *Psychol. Bull.* **79**, 200-210.
- HALLAM R., RACHMAN S. and FALKOWSKI W. (1972) Subjective, attitudinal and physiological effects of electrical aversion therapy. *Behav. Res. & Therapy* **10**, 1-13.
- INSKO C. A. and OAKES W. F. (1966) Awareness and the 'conditioning' of attitudes. *J. pers. soc. Psychol.* **4**, 487-496.
- LAZARUS R. S. (1966) *Psychological Stress and the Coping Process*. McGraw Hill, New York.
- MEYER A. B. and MARTIN I. (1968) Shape of the conditioned eyelid response. *Psychol. Rev.* **75**, 398-408.
- LOCKHART R. A. (1973) Classical conditioning and the cognitive processes. A symposium. *Psychophysiology* **10**, 74-122.
- MARTIN I. and LEVEY A. B. (1969) *Genesis of the Classical Conditioned Response*. Pergamon Press, Oxford.
- MEYER M. (1903) Experimental studies in the psychology of music. *Am. J. Psychol.* **14**, 456-476.
- MOWRER O. H. (1939) A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.* **46**, 553-565.
- MILLER N. E. (1948) Studies of fear as an acquirable drive. I. Fear as motivation and fear reduction as reinforcement in the learning of new responses. *J. exp. Psychol.* **38**, 89-101.
- OSGOOD C. E. (1971) Exploration in semantic space: A personal diary. *J. soc. Issues* **27**, 45-64.

- PAGE M. M. (1971) Post experimental assessment of awareness in attitude conditioning. *Educ. Psychol. Meas.* **31**, 891-906.
- RACHMAN S. and TEASDALE J. (1969) *Aversion Therapy and Behaviour Disorders: An Analysis*. Routledge, Kegan Paul, London.
- RAZRAN G. H. S. (1938) Conditioning away social bias by the luncheon technique. *Psychol. Bull.* **35**, 693.
- RAZRAN G. H. S. (1954) The conditioned evocation of attitudes (cognitive conditioning?) *J. exp. Psychol.* **48**, 278-282.
- SELIGMAN M. E. P. and HAGER J. L. (Eds.) (1972) *Biological Boundaries of Learning*. Appleton-Century-Crofts, New York.
- STAATS C. K., and STAATS, A. W. (1957) Meaning established by classical conditioning. *J. exp. Psychol.* **54**, 74-80.
- STAATS A. W. and STAATS C. K. (1958) Attitudes established by classical conditioning. *J. abnorm. soc. Psychol.* **57**, 37-40.
- YOUNG P. T. (1967) Affective arousal: some implications. *Am. Psychol.* **22**, 32-39.
- ZAJONC R. B., SHAVER P., TAVRIS, C. and VAN KREVELD D. (1972) Exposure, satiation and stimulus discriminability. *J. pers. soc. Psychol.* **21**, 270-280.