

*THE ROLE OF INFORMATION IN THE
EMISSION OF OBSERVING RESPONSES:
A TEST OF TWO HYPOTHESES¹*

R. N. WILTON AND R. O. CLEMENTS

UNIVERSITY OF TEXAS

Pigeons were trained on a trial procedure in a Skinner box. Each trial began with a fixed-interval schedule. Responding on this schedule produced a stimulus and a delayed trial outcome. The stimulus signalled whether the forthcoming outcome was reinforcement or nonreinforcement. Thus, the response was an observing response. When reinforcement was the outcome on 20% of the trials, response rates in the fixed interval were higher than when reinforcement was the outcome on 80% of the trials. This result is consistent with the hypothesis that observing responses are reinforced by the information associated with the stimulus signalling reinforcement. The result seems inconsistent with the hypothesis that observing responses are also reinforced by the information associated with the stimulus signalling nonreinforcement.

Using a trial procedure, Wilton and Clements (1971) studied the emission of observing responses (Wyckoff, 1952). Pecking on a fixed-interval (FI) schedule produced one of two equi-probable stimuli. Some seconds later the stimulus terminated in one of two equi-probable outcomes—reinforcement or nonreinforcement. The stimuli and outcomes could be correlated so that one stimulus (hereafter called the positive signal) was always followed by reinforcement, and the other (the negative signal) by nonreinforcement. Under these conditions the FI response was an observing response (Wyckoff, 1952). Alternatively, the stimuli and outcomes could be uncorrelated, with each stimulus being followed equally often by reinforcement and nonreinforcement. Response rates were shown to be higher when the stimuli and outcomes were correlated than when they were uncorrelated. Hence, the response was emitted more frequently when it was an observing response. In common with Hendry (1965, 1969), Wilton and Clements explained their results by supposing that the information transmitted by a correlated stimulus is reinforcing.

The hypothesis can be developed further. Hendry (1965, 1969) suggests that information

is reinforcing regardless of whether it is delivered by the positive or negative signal. On the other hand, it may be that only information in the positive signal is reinforcing, and that it alone maintains observing responses. Data from Schaub (1969) seem to have already disproved the latter hypothesis because they appear to show reinforcement by the negative signal. "Master" and yoked pigeons were trained on a schedule in which variable interval (VI) and extinction components irregularly alternated. For both components the keylight was normally white (mixed schedule), and therefore uninformative with respect to the prevailing component. However, three pecks (FR 3) by a master bird turned the key either red or green, for both the master bird and its corresponding yoked bird. The colors were correlated with the prevailing component, and hence were informative. Schaub (1969) showed that the master birds responded at a higher rate in the presence of the white light than the yoked birds, and he concluded that the informative colors, contingent on pecking for the master birds, were reinforcers. Next, Schaub changed the procedure so that only one of the colors could be produced. For instance, responses in the (white) extinction component produced the negative color (red) while responses in the (white) VI component produced no change. Responding in the (white) extinction component was highest for the master birds. Schaub concluded that the information delivered by the negative signal

¹This research was supported by Grant MH-18185-01 from the National Institute of Mental Health, United States Public Health Service, to R. N. Wilton. Reprints may be obtained from R. N. Wilton, Department of Psychology, University of Texas, Austin, Texas 78712.

(red) was reinforcing. However, this conclusion is unwarranted. Although responding in the (white) VI component failed to produce the green positive signal it did produce information, since the occurrence of three pecks without the production of red implied the operation of the VI component just as much as the green did previously. Hence, the relatively high pecking rates by the master birds in the presence of the white key, (regardless of the associated schedule being one of VI or extinction) may have been due only to the production of positive information.

What these objections amount to is that, although it was intended otherwise, Schaub's procedure was arranged so that pecking at the white key produced both the positive and the negative signals. Given this, we might try to think of some other procedure whereby we could show that just one type of information (positive or negative) is sufficient to *maintain* responding by having only it contingent upon key pecking. But this is impossible, *in principle*. With such a procedure the stimulus setting the occasion for the response (white key) will always be followed by the originally informative cue. Therefore, after a few trials the first (white) stimulus will make the second (colored) redundant, and as a result, the latter will no longer be informative (reduce no uncertainty).

It seems that Schaub's (1969) results argue neither for nor against the hypothesis that information in the negative signal is reinforcing, and therefore they fail to discriminate between the two hypotheses.

One difference between the hypotheses becomes apparent if we examine how reinforcement by information should vary as a function of the relative frequency (probability) with which the positive signal follows the observing response. In conformity with the hypothesis that information in both the positive and negative signals is reinforcing, Hendry (1969, p. 390) suggests that an observing response is reinforced in proportion to the average amount of information in the following signal—regardless of the signal being positive or negative. This amount of information is given by the expression $p \log_2 (1/p) + 1-p \log_2 (1/1-p)$, where p is the probability of the positive signal following the response, and $1-p$ is the probability of the negative signal following the response. Figure 1 shows $p \log_2$

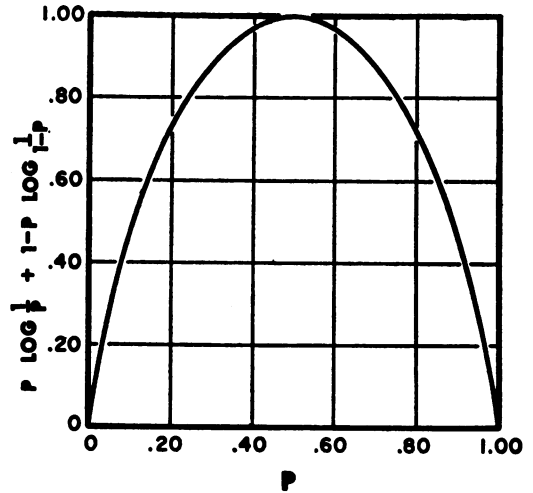


Fig. 1. The average amount of information transmitted by the positive and negative signals as a function of the probability of the positive signal (adapted from Garner, 1962).

$(1/p) + 1-p \log_2 (1/1-p)$ as a function of p . The parts of the expression $p \log_2 (1/p)$ and $1-p \log_2 (1/1-p)$ can be taken to represent the contribution to average information by information associated respectively with the positive and negative signals. Thus, if only information in the positive signal is reinforcing, then reinforcement by information is propor-

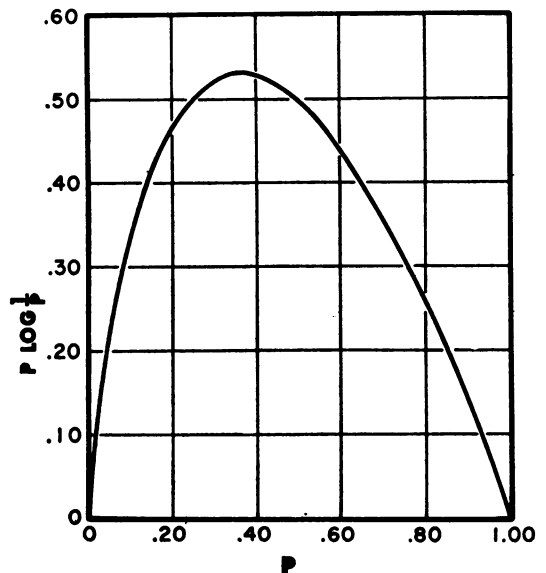


Fig. 2. The amount of information contributed by the positive signal to the average amount of information, plotted as a function of the probability of the positive signal (adapted from Garner, 1962).

tional to $p \log_2 (1/p)$, the amount of information contributed to the average by the positive signal. Figure 2 shows $p \log_2 (1/p)$ as a function of p .

Consideration of the two curves suggests that the hypotheses can be set against each other without attempting operationally to separate the effects of information from the positive and negative signals (*cf.* Schaub, 1969). As one of the curves is symmetrical about a p value of 0.5 (Fig. 1), while the other is asymmetrical (Fig. 2), different predictions are made regarding the relative amount of reinforcement by information at p values equidistant from (on opposite sides of) 0.5. Consider p values of 0.2 and 0.8 for example. The first hypothesis predicts equal reinforcement by information at these values, while the second predicts more reinforcement when $p = 0.2$.

Response rates on a fixed-interval schedule increase with reinforcement magnitude (Stebbins, Mead, and Martin, 1959). In conjunction with the above reasoning, this suggests that if an observing response were reinforced on a fixed-interval schedule, as in the Wilton and Clements (1971) study, then the two hypotheses would differ in their prediction of relative response rates at the two p levels. However, perhaps a further factor must be considered. Whenever the positive signal follows the response, food follows also. And one could assume that the food reinforces responding independently of the information—in the sense that its reinforcing effects summate with those of information. In that case one could not predict response rates only on the basis of either of the curves shown in Fig. 1 and 2. But if one then makes the reasonable assumption that food is more reinforcing when it is more frequent—for instance when it follows the response with a relative frequency of 0.8 rather than 0.2—then the following assertion still holds. The hypothesis that states that information is reinforcing regardless of its being in the positive or negative signal precludes the possibility of obtaining a higher observing response rate when $p = 0.2$ than when $p = 0.8$, while the hypothesis stating that only information in the positive signal is reinforcing does not. In the present experiment, observing responses that produced information on a fixed-interval schedule were measured at p values of 0.2 and 0.8 by varying frequency of

reinforcement (and therefore p) in a situation similar to that used by Wilton and Clements (1971).

METHOD

Subjects

Six adult homing pigeons were maintained at about 80% of their free-feeding weights. Four of the birds (103Y, 199G, 186G, 131Y) had been used in a similar experiment (Wilton and Clements, 1971). The other two (181G, 196G) were naive.

Apparatus

A sound-insulated operant conditioning chamber contained a translucent response key of 0.75 in. (2 cm) diameter that could be illuminated from behind by either a white, red, or green light. A grain magazine was centered below the key. The reinforcer consisted of a 3 sec exposure to a grain mixture. Sessions were scheduled through relays and timers contained in an adjoining room.

Procedure

The naive birds were trained to peck at the key by the auto-key peck training procedure used by Brown and Jenkins (1968). All the birds, both naive and non-naive, were then run according to the procedure described below, except that for the naive birds the time intervals used in the procedure were shortened in the first few sessions to allow them to adjust to the schedule.

A bird was placed in the chamber, and 5 min later the onset of the houselights initiated a session. Simultaneously, a trial began with the illumination of the pecking key by a white light. The first peck to occur after 15

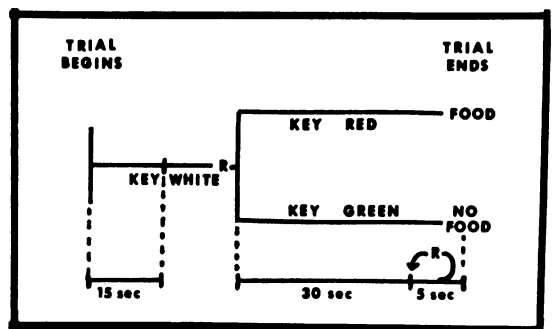


Fig. 3. A description of the contingencies in the experiment. The letter R refers to a key-peck response.

sec (FI 15-sec) changed the key to either red or green (Fig. 3). Thirty-five seconds after this peck, an outcome of reinforcement or nonreinforcement occurred—unless a response was emitted within the last 5-sec before the outcome was to occur. If such a response was emitted the outcome was delayed, and continued to be delayed until no response occurred for 5 sec. This prevented a pecking response being immediately followed by an outcome. The red and green colors, produced by the pecking response, were correlated with the two types of outcome so that one color was always followed by reinforcement while the other color was always followed by non-reinforcement. When a trial outcome was reinforcement, the food magazine was raised and a light illuminated the food. When the outcome was nonreinforcement, only the light came on. At the occurrence of an outcome the key darkened, and remained dark until the next trial 60 sec later. Twenty-five trials constituted a session.

For Birds 103Y, 186G, and 181G, the stimulus correlated with reinforcement was red, while for Birds 196G, 199G, and 131Y it was green. The probability of reinforcement was determined by a Lehigh Valley probability generator, set to give a reinforcement schedule of either 20% or 80%. Birds 131Y, 181G, and 103Y were initially run on the 20% reinforcement schedule. Birds 186G, 199G, and 196G were initially run on the 80% schedule. Sessions were conducted every day. Each bird was run for between 75 and 90 sessions; at the end of this training responding was judged stable over at least the last 25 sessions. The reinforcement schedule was then changed—from 20% to 80% or from 80% to 20%. Sixty sessions were given on the second schedule.

RESULTS

Figure 4 shows response rates over the final 25 sessions of the first experimental condition, and all 60 sessions of the second condition. Each bird responded more frequently on the 20% than on the 80% schedule. This happened regardless of which schedule was given first, and whether the birds were naive. (The data from the non-naive birds may be compared with those obtained in a previous experiment [Wilton and Clements, 1971].) In-

cluded in Fig. 4 are the actual percentages of reinforcement that each bird received over each phase. These correspond quite closely to the percentages stipulated by the setting of the probability generator. More importantly, the lower and upper percentages are about equidistant from 50%. The data from Bird 103Y deserve individual comment. Over the last 25 sessions on the 20% schedule, response rates tended to decline. Consequently, the subsequent decline on the 80% schedule could be attributed to a continuation of this process rather than the schedule change.

Following a change in schedule, response rates usually changed quite slowly, taking an average of about 30 sessions to become stable. This suggests that the relatively low response rates on the 80% schedule did not result from satiation occurring within a session. If satiation caused the rate differences between the 20% and 80% schedules, response rates would have changed abruptly in the very first session following a schedule change. Satiation would have occurred in the first (and last) session of the 80% schedule, and not in the first (and last) session of the 20% schedule.

DISCUSSION

The present data, and other data (Hendry, 1965, 1969, p. 391; McMichael, Lanzetta, and Driscoll, 1967; McMillan, 1970) suggest that a positively skewed asymmetrical curve describes the relation between observing response rate and reinforcement frequency. From this one can conclude that the reinforcing effect of information is not determined only by its amount (number of bits). It seems that the quality of the information is also important, for the asymmetrical function shows that a given amount of information in the negative signal does not have the same effect as an equivalent amount in the positive signal. The hypothesis that *only* information in the positive signal is positively reinforcing goes beyond this, and is not proven by the finding of an asymmetrical function (information from the negative signal could simply be *less* positively reinforcing). Yet such a function is consistent with this hypothesis, as is our intuitive notion that "good news" (from the positive signal) should be positively reinforcing while "bad news" (from the negative signal) should not.

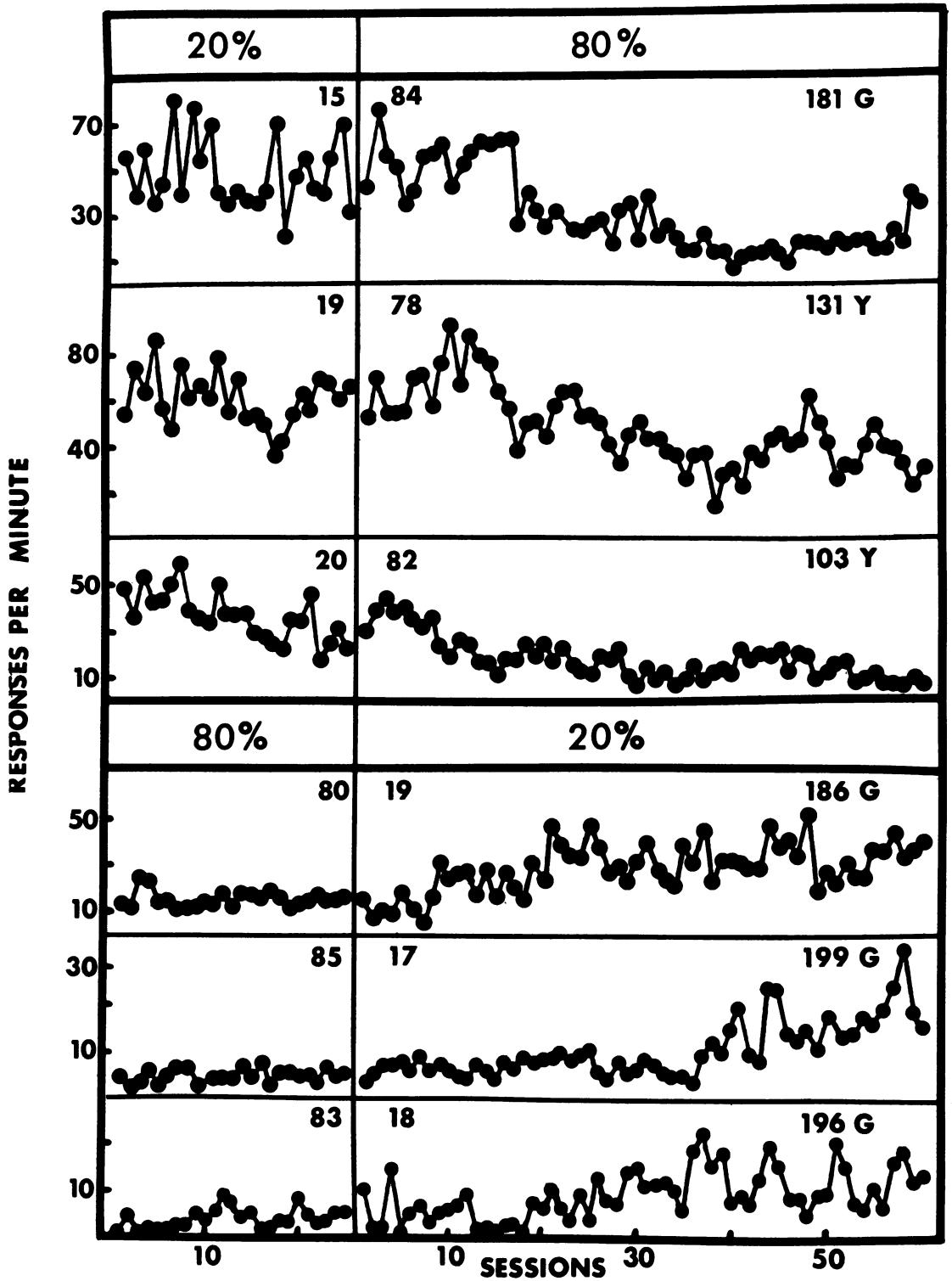


Fig. 4. Response rates in the fixed interval at 20% and 80% reinforcement (positive signal) frequencies. (Note that the ordinate scale for Birds 199G and 196G differs from that of the others.)

REFERENCES

- Brown, P. L. and Jenkins, H. M. Auto-shaping of the pigeon's key peck. *Journal of the Experimental Analysis of Behavior*, 1968, 11, 1-8.
- Garner, W. R. *Uncertainty and structure as psychological concepts*. New York: John Wiley and Sons, 1962.
- Hendry, D. P. Reinforcing value of information. NASA Technical report 65-1, Space Research Laboratory, University of Maryland, 1965.
- Hendry, D. P. (Ed.) *Conditioned reinforcement*. Homewood, Illinois: The Dorsey Press, 1969.
- McMichael, J. S., Lanzetta, J. T., and Driscoll, J. M. Infrequent reward facilitates observing responses in rats. *Psychonomic Science*, 1967, 8, 23-24.
- McMillan, J. C. *Uncertainty reduction and the maintenance of observing behavior*. Paper read at the meeting of the Psychonomic Society, San Antonio, Texas, 1970.
- Schaub, R. E. Response-cue contingency and cue effectiveness. In D. P. Hendry (Ed.), *Conditioned reinforcement*. Homewood, Illinois: The Dorsey Press, 1969. Pp. 342-357.
- Stebbins, W. C., Mead, P. B., and Martin, J. M. The relation of amount of reinforcement to performance under a fixed-interval schedule. *Journal of the Experimental Analysis of Behavior*, 1959, 2, 351-357.
- Wilton, R. N. and Clements, R. O. Observing responses and informative stimuli. *Journal of the Experimental Analysis of Behavior*, 1971, 15, 199-204.
- Wyckoff, L. B. The role of observing responses in discrimination learning. *Psychological Review*, 1952, 59, 431-442.

Received 16 December 1970.