Blocking Observed in Human Instrumental Conditioning

MARIANNE HAMMERL

Heinrich-Heine-University Duesseldorf, Germany

In animal research, blocking of instrumental responding is a well-known phenomenon, whereas no study has been reported that investigated blocking in human instrumental conditioning. Following the standard method used in nonhuman subjects, 48 students randomly assigned to one of three groups were exposed to a variable interval schedule (VI 10 s) in which reinforcement was delivered with a brief delay (500 ms). In the blocking condition, subjects experienced a tone stimulus during the delay (correlated group). In the two control conditions, subjects received either no tone (no-tone group) or the same number of tones as subjects of the correlated group, but the tones were independent of their behavior and reward (random group). As expected, instrumental responding was significantly lower in the correlated group than in either the no-tone or the random group. In a subsequent extinction phase, no difference in resistance was observed. A postexperimental interview revealed that subjects of the correlated group were more likely to detect the temporal nature of the reinforcement schedule than subjects of the other groups, but there was no relation to response rate. The data provide only little support for a notion of signal-induced enhanced learning, but do not challenge an interpretation in terms of associative competition between the response and the signal. © 1993 Academic Press, Inc.

Over the past several years, the traditional distinction between Pavlovian and instrumental conditioning has been called in question in many respects. Both the techniques of classical conditioning and the theoretical understanding about the underlying learning process have been applied to the analyses of instrumental learning (e.g., Dickinson, 1987, 1989; Mackintosh, 1974, 1983; Rescorla, 1987; St. Claire-Smith, 1979a, b). Ac-

This experiment is part of the author’s doctoral thesis accepted by the Faculty of Mathematics and Natural Science at the Heinrich-Heine-University Duesseldorf and honored with the “Career Award for Science 1991” granted by the City of Duesseldorf. A preliminary report of the work was presented at the 12th annual meeting of Belgian, German, and Dutch Comparative and Physiological Psychologists and Behavioral Pharmacologists in Cologne, Germany, May 1991. The author is indebted to H.-J. Grabitz for his support of the entire project. Address correspondence and reprint requests to Marianne Hammerl, Department of Psychology, Heinrich-Heine-University Duesseldorf, Universitaetsstr. 1, D-4000 Duesseldorf 1, Germany.
cording to this view, instrumental conditioning is based on the acquisition of response–reinforcer associations in a manner similar to the acquisition of stimulus–reinforcer associations in Pavlovian conditioning. One possible way to test this assumption is to investigate whether phenomena revealed in classical conditioning experiments also occur in an instrumental paradigm. For instance, St. Claire-Smith and MacLaren (1983) demonstrated in instrumental learning an effect that resembles the sensory preconditioning phenomenon (Brogden, 1939) in Pavlovian conditioning. They simply replaced one of the two neutral stimuli with a response. Rats that experienced presentations of this response–stimulus sequence, followed by pairings of the stimulus with food (Experiment 1) or shock (Experiment 2), showed enhancement (Experiment 1) or suppression (Experiment 2) of the instrumental response during a test in which responding had no programmed consequences.

Other phenomena that have been moved from a Pavlovian to an instrumental paradigm are overshadowing and blocking. Overshadowing and blocking occur in classical conditioning experiments when a compound stimulus, composed of two elements, is paired with the unconditioned stimulus. When one element in the compound is more salient than the other, the more salient stimulus may prevent the other from developing a conditioned response. This effect, termed overshadowing, was first described by Pavlov (1927). A similar effect, named blocking (Kamin, 1969), occurs when one stimulus of the compound has undergone conditioning before being placed in that compound. The previously conditioned element reduces the ability of a novel stimulus to elicit the conditioned response.

There have been different attempts to investigate overshadowing and blocking in an instrumental context. First results were obtained in a discriminative learning paradigm. When the discriminative stimulus was a compound of two elements, the more salient stimulus overshadowed the less salient (e.g., Mackintosh, 1976; Miles, 1969; Miles & Jenkins, 1973) and a preconditioned element blocked a novel stimulus (e.g., Mackintosh & Honig, 1970; Neely & Wagner, 1974) in developing discriminative properties. These experiments, however, do not fully meet the requirements of the research strategy outlined above (moving phenomena from a Pavlovian to an instrumental paradigm). The elements used to demonstrate overshadowing and blocking are still stimuli.

In order to show overshadowing or blocking in instrumental learning, at least one of the two elements should be replaced with a response. Mackintosh and Dickinson (1979) have provided a rather formal application of the classical blocking procedure to instrumental conditioning using forced wheel-running and a tone stimulus as elements. Another, perhaps more convincing (as stated by Mackintosh, 1983), version of the instrumental blocking paradigm is the signaling reward procedure, either
as a discrete-trial procedure (Williams, 1975, 1978) or a free-operant procedure (Pearce & Hall, 1978; St. Claire-Smith, 1979b). The primary finding of these studies has been attenuation of instrumental responding on a variable interval (VI) schedule when reinforcement is preceded or accompanied by a brief light or tone stimulus, relative to response rates seen when reinforcement is unsignaled. Subsequent work supported this result (e.g., Hall, 1982; Roberts, Tarpy, & Cooney, 1985; Roberts, Tarpy, & Lea, 1984; Schachtman & Hall, 1990; Schachtman, Reed, & Hall, 1987; Tarpy, Roberts, Lea, & Midgley, 1984).

In contrast to classical conditioning, the distinction between overshadowing and blocking is difficult to maintain in the instrumental context because the signaling stimulus does not have a prior history of pairing with the reinforcer. Instead, the stimulus is persistently paired with reward within the training phase itself. Hence, both terms are found in the literature. We prefer the term “blocking” because it does not seem very likely that response attenuation produced by the reward-related stimulus is a result of a difference in salience between the stimulus and the response. Rather, the effect seems to be caused by a difference in the prediction of the reward (Tarpy, Lea, & Midgley, 1983, Experiment 1): The stimulus is a better predictor of reinforcement than the instrumental response because the response, but not the stimulus, sometimes occurs without the reinforcer.

Why the signaling of reward results in a decrease in the rate of responding in VI schedules has been explained in several ways. According to the position of Pearce and Hall (1978) and St. Claire-Smith (1979b, 1987), the reduced response rate is caused by the formation of a stimulus–reinforcer association that blocks the development of the response–reinforcer association. In contrast, the efficiency theory proposed by Tarpy and Roberts (Roberts et al., 1984; Tarpy et al., 1983; Tarpy & Roberts, 1985) states that the low response rate is evidence of a strong response–reinforcer association. They suggest that the signal enhances the sensitivity of the subject to the time-dependent nature of the VI schedule and thus promotes the formation of efficient (i.e., less frequent) responding. Another explanation is based on the notion that sign-tracking behaviors to the reinforcement signaling cue interfere with the instrumental response (Iversen, 1981; Reed, 1992). According to the account of Williams (1982; Williams & Heyneman, 1982), the presentation of a signal prevents the development of “response bursts” that will otherwise occur when reinforcement is delayed for the duration of the signal. The interpretation of Williams refers to the experiments mentioned above in which brief delays of reinforcement were used. In procedures in which the delay was longer, Williams demonstrated that the signal has different effects (attenuation or enhancement of responding) depending on the signal location in the delay-of-reinforcement interval (Williams, Preston, & de Kervor, 1990).
The findings and theoretical explanations described above are based on animal experiments, whereas no study has been reported that investigated the blocking effect in human instrumental conditioning. One reason for this lack may be the unsuccessful attempts to demonstrate blocking in human Pavlovian conditioning (e.g., Lovibond, Siddle, & Bond, 1988; but see Martin & Levey, 1991). In conditioning experiments in which the dependent variable is not an elicited response (e.g., electrodermal conditioning) but a causal judgment, blocking has been found more easily (Dickinson, Shanks, & Evenden, 1984; Shanks, 1985). However, presumably, there are some logical errors in the experimental task of Shanks and Dickinson (i.e., the ambiguity of the experimental stimuli) that do not call into question their demonstration of blocking but do call into question their explanation of the effect (see Maier, 1989, for a detailed critique). Siegel and Allan (1985) demonstrated blocking in humans using a visual perception phenomenon (orientation-contingent color after effect).

The purpose of the present study was to provide a blocking procedure for human instrumental behavior that resembles the standard conditioning methods used in nonhuman subjects. According to the animal experiments mentioned above, the subjects were exposed to a VI schedule in which the reinforcer was delivered with a brief delay (500 ms). In the blocking condition, a tone stimulus occurred in the delay interval between the response and the reinforcement. In two control conditions, subjects received either no tone or the same number of tones as subjects of the blocking condition, but the tones were independent of their behavior and reward. It was expected that the rate of responding in the blocking condition would be lower than that in either of the control conditions.

Following the suggestions of Tarpy and Roberts (1985; Roberts et al., 1984), in addition to examining the instrumental responding, we examined the efficiency of this activity (i.e., response rate compared with obtained reinforcement) and the resistance of responding in the face of changed contingencies (i.e., extinction). These measurements along with the subjects' verbal descriptions of the relationships between the experimental events (response, reward, and tone stimulus) had the purpose of permitting first insights into the underlying process of the expected blocking effect in human instrumental responding.

**METHOD**

*Subjects*

Fourty-eight students of the Heinrich-Heine-University Duesseldorf (24 females, 24 males) majoring in subjects other than psychology voluntarily participated in the experiment. All subjects were right-handed and ranged in age from 19 to 32 years ($M = 23.63$, $SD = 3.42$). They were told
that they would earn DM 7.00 (approx. $4.00) for 1 h of participation with the potential for earning more in the session. No subjects reported any previous experience with behavioral research.

Apparatus

Figure 1 shows a sketch of the experimental apparatus. It was located on a table in a darkened room that measured 3.5 by 4.5 m. Each subject was seated in a chair in front of the monochrome monitor. On the right-hand side, there was a wooden box that resembles a slot machine. It measured 16.4 cm long by 18 cm wide by 31 cm high. The instrumental response consisted of putting a metal coin (3.9 cm in diameter) into the upper slot of the box. Inside the box, the coin fell down (guided by a rail) to the bottom of the box where it appeared in an ejection rail shortly after it had been inserted into the upper slot. Each insertion was registered by a light barrier positioned inside the box.

A speaker was mounted under the table invisible to the subjects and supplied a 240-Hz tone measuring approximately 55 dB(A) (re 20 μN/m²). Experimental events were controlled and recorded by a Commodore personal computer (PC 10) located in an adjacent room.

Procedure

Subjects were run individually. The instructions for all subjects were as follows:

Hello! You are now in our Psycho-Lab. In front of you, there are a monitor and a wooden box, which is similar to a slot machine. In this experiment, your task is to gain points. Everytime you receive a point it will appear on the screen. The way to obtain points is by putting the metal coin into the upper slot of the machine. The number of points you gain will determine the amount of additional money
you earn. Thus, you get your payment for participation (DM 7.00) plus a monetary bonus. This is a stress test; therefore, the main thing is to put the coin into the slot quite often. The session will not be finished until the experimenter informs you. Do you have any questions?

The statement "This is a stress test; therefore, the main thing is to put the coin into the slot quite often" had the purpose of masking the goal of the study and minimizing subject-produced self-instructions, which had been observed in a pilot study using instructions without a masking sentence. Any questions asked by a subject were responded to by repeating the appropriate section(s) of the instructions. The experimenter then left the room for the remainder of the session.

The experiment consisted of three sequential phases: an acquisition phase (30 min), an extinction phase (15 min), and a postconditioning phase measuring subjects' explicit knowledge of the relationships between the experimental events.

Acquisition phase. For all subjects, instrumental responses were reinforced according to a VI 10-s schedule (range 3 to 27 s). Each time the programmed schedule of reinforcement was satisfied, one point was delivered. When a point was earned, a small twinkling square located in a "point tub" (see Fig. 1) appeared on the monitor screen for 3 s. Thereafter, the twinkling turned into a permanent display, and an additional counter presented on the upper part of the screen was updated. At the end of the session, the points were exchanged for money. The amount of additional money ranged from DM 3.70 (approx. $2.10) to DM 4.57 (approx. $2.60). There was a delay of 500 ms between the critical response and the delivery of reinforcement. The timing of the delay was initiated from the point the coin had passed the light barrier. The subjects were randomly assigned to one of three groups, with the restriction that each group contained an equal number of females and males (n = 8). In the blocking condition, the delay was filled with a tone stimulus (correlated group). In the two control conditions, subjects received either no tone (no-tone group) or the same number of tones as subjects of the correlated group, but the tones were independent of their behavior and reward (random group).

Extinction phase. After the acquisition phase, all subjects went directly into 15 min of extinction. In all three groups, reinforcers were no longer delivered. In the correlated and random groups, the tone stimulus was also omitted.

Postconditioning phase. When the extinction phase ended, all subjects were requested to describe the manner in which points were earned during the acquisition phase. In addition, subjects of the correlated and random groups were asked about the relation between occurrence of the tone stimulus and delivery of reinforcement.
Data Analysis

Instrumental responses were tested statistically using Mann–Whitney U tests (one-tailed) with \( n_1 = n_2 = 16 \). A rejection level of \( p > .05 \) was adopted for all comparisons. To assess resistance to extinction, the extinction data were transformed to percentage-of-baseline scores: The number of absolute instrumental responses for each subject for each of the three blocks of the extinction phase was divided by the subject’s final acquisition score (Block 6). Subjects’ schedule descriptions were analyzed by a modified version of the Fisher exact probability test for the case of more than two samples (Krauth, 1988).

RESULTS

Instrumental Responding

Acquisition phase. The mean numbers of instrumental responses made during the six blocks of the acquisition phase are presented in Fig. 2. As can be seen, the subjects of the correlated group showed the lowest rate of instrumental responding. Overall instrumental responding (see Table 2) was significantly lower in the correlated group than in either the no-tone (\( U = 41 \)) or the random (\( U = 76 \)) group. The control groups did not differ significantly (\( U = 112 \)). Figure 2 also shows that the difference between the correlated group and the other groups was present from the very first block of the acquisition phase. Further analyses using data from the first 5 min revealed that the groups did not differ at the beginning of
the experiment, but differed significantly from the 3rd min on (correlated group vs no-tone group: $U = 64$, correlated group vs random group: $U = 82$).

**Extinction phase.** Figure 2 also presents the mean numbers of instrumental responses made during the three blocks of the extinction phase. Overall instrumental responding was significantly lower in the correlated group than in either the no-tone ($U = 58$) or the random ($U = 80$) group. There was a noticeable, but nonsignificant, difference in responding between the no-tone and the random groups ($U = 100$). Extinction resistance data are displayed in Table 1. Inspection of these data reveals that all groups showed a decline in the relative rate of responding. At the end of the extinction phase (Block 9), the three groups responded at nearly the same rate, namely at three-fourths of the baseline level. The groups did not differ significantly at any block.

**Efficiency of Responding**

Table 2 also presents the mean numbers of reinforcers for each group. Inspection of these data reveals that all groups obtained nearly the same amount of reinforcement, whereas their instrumental responding differed

<table>
<thead>
<tr>
<th>Group</th>
<th>Responses</th>
<th>Reinforcers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlated</td>
<td>$M$</td>
<td>1119.31</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>354.78</td>
</tr>
<tr>
<td>Random</td>
<td>$M$</td>
<td>1340.38</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>347.63</td>
</tr>
<tr>
<td>No-tone</td>
<td>$M$</td>
<td>1453.31</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>156.39</td>
</tr>
</tbody>
</table>
as described above. The correlated group showed the lowest number of emitted responses, and thus, this group was more efficient than the other groups.

Postexperimental Interview

In the postconditioning phase, all subjects were requested to describe the manner in which points were earned during the acquisition phase. Their answers were classified into time, ratio, or other descriptions (see Table 3). Four subjects of the correlated group, none of the subjects of the random group, and 1 subject of the no-tone group correctly described earnings as based on time passage (e.g., “I put the coin into the slot about every 10 s.”). Six subjects of the correlated group, 7 subjects of the random group, and 11 subjects of the no-tone group supposed that earnings were based on a varying number of coin insertions (e.g., “I put the coin into the slot 10 to 15 times.”). Answers classified as “other description” either were that the computer delivered the points randomly or were without informational value (e.g., “I don’t know in which way I gained points.”). In the correlated, random, and no-tone groups, 6 subjects, 9 subjects, and 4 subjects, respectively, answered in this way. Statistical analysis revealed that these results fell short of significance ($p = .055$).

The mean numbers of instrumental responses in the acquisition phase for the subjects of the correlated group who reported time, ratio, or other descriptions were 820 ($SD = 465.11$), 1387.83 ($SD = 142.74$), and 1050.33 ($SD = 252.05$), respectively. Instrumental responding was significantly higher in the ratio-description subgroup than in either the time-description ($U = 1$) or the other-description ($U = 4.5$) subgroup. The latter two subgroups did not differ significantly ($U = 9$).

In addition, subjects of the correlated and random groups were requested to describe the relation between the occurrence of the tone stimulus and the delivery of reinforcement. All 16 subjects of the correlated group correctly answered that the tone only occurred in the context of a point gain that it preceded. In the random group, 10 subjects accurately described the uncorrelated relationship between the occurrence of the
tone and point earning, whereas 6 subjects assumed that there was a correlation between tone and point gain.

DISCUSSION

The results of the present study show that blocking of instrumental responding is demonstrable in humans. In agreement with findings from animal experiments, the rate of responding is retarded when each occurrence of the reinforcer is preceded by a stimulus better correlated with reinforcement than the response. Surprisingly, even the very early beginning of the detrimental influence of the reinforcement signaling cue resembles observations reported in animal literature (e.g., Pearce & Hall, 1978, Experiment 2; Roberts et al., 1984, Experiment 1). Some conditioning trials seem sufficient to produce blocking of instrumental responding. This finding is similar to results revealed in classical conditioning experiments, according to which evidence of blocking and overshadowing following a single training trial was obtained (Balaz, Kasprow, & Miller, 1982; James & Wagner, 1980; Mackintosh, 1971; Mackintosh & Reese, 1979; Revusky, 1971). One-trial learning is not unusual in classical conditioning. Additional examples are taste-aversion learning (Garcia & Koelling, 1966) and latent inhibition (Devietti, Blair, & Schlesner, 1989).

The high rates of responding (up to 248 responses per 5 min) shown by the three groups of the present study are probably caused by a particular formulation given in the instructions to minimize subject-produced self-instructions (i.e., the masking sentence). The demand "to put the coin into the slot quite often" might also be the reason why most subjects supposed that reinforcement had been ratio-based. Despite these possible effects both on overall response rate and on subjects' verbal descriptions, there is still no instruction-induced insensitivity to influences of the experimental contingencies. Signaling the occurrence of reinforcement leads to a significantly lower response rate relative to control conditions and to few correct schedule descriptions (i.e., time description).

As outlined in the introduction, four different accounts apply to the decrease in response rate when reinforcement is signaled. According to the explanation of Iversen (1981) and Reed (1992), the reinforcement signaling cue evokes sign-tracking behaviors that interfere with instrumental responding and thus reduces the number of counted responses. In the present study, the signaling cue was a tone stimulus that was difficult to locate and approach. Furthermore, it would have been possible to perform the instrumental response even if orienting responses occurred.

Another explanation is based on the assumption that delayed reinforcement leads to "response bursts" which do not appear when the delay is filled with a signal (Williams, 1982; Williams & Heyneman, 1982). In the present study, the delay interval (500 ms) was shorter than the time
required to perform the instrumental response (at least 800 ms). Therefore, bursts of responses were not seen.

The two remaining accounts differ with respect to their assumptions on the strength of the response–reinforcer association. According to the position of Pearce and Hall (1978) and St. Claire-Smith (1979b, 1987), the low rate of responding is a result of poor learning (i.e., a weak response–reinforcer association), whereas Tarpy and Roberts (Roberts et al., 1984; Tarpy et al., 1983; Tarpy & Roberts, 1985) suppose that reduced response rates in VI schedules are an evidence of efficient behavior and thus a result of a strong response–reinforcer association. In agreement with the assumptions of Tarpy and Roberts, the present findings show that subjects of the correlated group were more efficient than subjects of the other groups. However, a suppression of overall response rate improves the relation between input (number of responses) and output (number of reinforcers) in VI schedules nearly automatically, inasmuch as reinforcement in VI schedules does not depend primarily on the number of responses the subject makes. A more convincing example of efficient responding is reported by Tarpy and Roberts (1985). In their experiment, the delivery of reinforcement was signaled in temporally based schedules that were not VI schedules. Using a differential reinforcement of low rates (DRL) schedule and a differential reinforcement of high rates (DRH) schedule, respectively, they were able to show that signaled reward generates lower responding on the DRL schedule but a higher rate on the DRH schedule. Furthermore, animals of the signaled conditions were more efficient than subjects of the control groups in which reward was unsignaled.

Subjects’ verbal descriptions of the schedule in operation suggest that the reinforcement signaling stimulus may have stressed the time-based nature of the reinforcement schedule and thus permitted more subjects in the correlated group to detect the correct reinforcer contingency. However, a strong relation between the detection of the temporal nature of the schedule and the response rate is missing. Subjects who reported descriptions of the schedule other than time or ratio show a response rate that is similar to the low rate emitted by the subjects who reported a time description. Thus, the efficiency theory does not appear to apply to the present data.

Other results of the present study, namely the extinction data, are inconsistent with the efficiency theory as well. Following Nevin (1979), Tarpy and Roberts suppose that a strong response–reinforcer association produced by a reinforcement signaling stimulus becomes evident in a great resistance to change. The present results, however, show a very similar decline in response rate for all groups. The differences in absolute responding shown in the acquisition phase were also apparent during extinction, but on a lower response level. Relative responding calculated
according to Roberts et al. (1984) as percentage-of-baseline scores did not differ significantly. Hence, there is no evidence of a stronger resistance to extinction in the correlated group than in the other groups.

As mentioned above, a reinforcement signaling stimulus does not reduce responding in all circumstances. Signaling reinforcement increases response rates not only in DRH schedules but also in variable ratio (VR) schedules in the majority of the studies (Reed & Hall, 1988, 1989; Reed, Schachtman, & Hall, 1988a, b; Schachtman & Reed, 1990). The latter finding contradicts both the efficiency theory and the position that the signal weakens the response–reinforcer association. According to efficiency theory, the stimulus is expected to have no effect on responding when reinforcement is programmed on a ratio schedule because ratio schedules do not incorporate a temporal contingency that may be stressed by the signal. On the basis of the competing position, the signal is expected to weaken the response–reinforcer association in VR schedules as well, and thus, a reduced response rate should occur. In order to explain both increased responding in VR schedules and decreased responding in VI schedules when reinforcement is signaled, theories which interpret these effects in terms of conditioned reinforcement or “marking” have been worked out (Reed, 1989; Williams et al., 1990). It must be added that the “marking” interpretation can be considered as modified version of the original efficiency theory (Reed, Schachtman, & Hall, 1991).

In summary, the present study provides a procedure for the investigation of blocking in human instrumental behavior. Responding is retarded when each occurrence of the reinforcer is preceded by a stimulus better correlated with reinforcement than the response. The results reported here, coupled with those of the studies described above, support interpretations of instrumental conditioning which emphasize learning about the consequences of responding. However, the question that remains is whether the decrease in human instrumental responding is caused by poor or enhanced learning about the response–reinforcer relation. For the present, there is only little support for the argument of enhanced learning, whereas the data provide no proof against an interpretation in terms of associative competition between the response and the signal, analogous to the blocking effect obtained in classical conditioning. Thus, a weak response–reinforcer association seems to be responsible for the present results. Investigation of blocking in human instrumental responding is just beginning, whereas research with nonhuman subjects provides a great deal of ideas and empirical findings with respect to this problem.

REFERENCES


Received January 1, 1992
Revised March 24, 1992