

VALUE TRANSFER IN DISCRIMINATIVE CONDITIONING WITH PIGEONS

MARTINA SIEMANN, JUAN D. DELIUS,
DANIELA DOMBROWSKI, and STEFANIE DANIEL
Universität Konstanz

The existence of a direct value transfer between stimuli in the context of instrumental discrimination learning is demonstrated. In a first experiment pigeons learned to discriminate instrumentally four successively presented target stimuli. Pecks to them were reinforced with graded amounts of reward and penalty (A++, B+, C-, D- -). These stimuli were accompanied by four different neutral stimuli (N_a , N_b , N_c , N_d) but responses to these had no consequences. After discrimination of the target stimuli had been established the neutral stimuli were presented in pairs and in the absence of the target stimuli. These tests revealed a graded preference in accordance with the value transfer hypothesis. In a second experiment pigeons were taught to discriminate two target stimuli A+ and B-. Each of these was again accompanied by a neutral stimulus (N_a , N_b). Pigeons developed a strong preference for N_a . A discrimination reversal affecting the target stimuli (B+, A-) led to a strong reversed preference for N_b . A third experiment employed a similar design but in it one group of pigeons, unlike the control group, was prevented from pecking the neutral stimuli by transparent barriers. As both groups preferred the reward-related neutral stimulus with approximately equal strength it is concluded that classical conditioning, rather than pseudo-operant conditioning was responsible for the value transfer observed. A modification of an algebraic instrumental conditioning model incorporating an adventitious classical conditioning element is shown to emulate the value transfer obtained. The role of classical conditioning in instrumental discrimination learning is discussed.

A value transfer theory was proposed by Fersen, Wynne, Delius, & Staddon (1991) to account for a remarkable result they had obtained in an experiment with pigeons. These had learned to discriminate five different stimuli presented in overlapping pairs A+B-, B+C-, C+D- and

The research was supported by the Deutsche Forschungsgemeinschaft. We are grateful to Dipl. Ing. L. Xia for technical assistance, to Dr. A. W. Still (Durham Univ.), Prof. S. E. G. Lea (Exeter Univ.), and Prof. J. E. R. Staddon (Duke Univ.) for comments on an earlier draft. G. Latini helped with the English and A. McLintock with manuscript preparation.

D+E-, with responses to the + stimuli yielding a reward and responses to the - stimuli yielding a penalty. When the birds had mastered these discriminations they were tested with the pair B D in extinction trials. It was found that they responded to Stimulus B markedly more often than to Stimulus D. This was surprising because both stimuli could be expected to have acquired identical response eliciting values as each of them had been scheduled to be positive in one of the training pairs in which it occurred and negative in the other (see also Siemann, 1993). The direct value transfer explanation for this transitive responding runs as follows: Stimulus A, scheduled always to be rewarded during training, accumulates a high associative value, say $V_a = k$. Stimulus E, scheduled never to be rewarded retains a low value, say $V_e = 0$. The middle Stimuli B, C, and D, scheduled to be evenly rewarded and penalized across their occurrences, accumulate an intermediate value, say $V_b = V_c = V_d = k/2$. Because we are only concerned with ordinal predictions the precise magnitude of the k value is not critical. What is important, however, is that the theory postulates that when Stimulus X+ is chosen and rewarded, a small fraction of its value transfers instantly to its companion Stimulus Y- according to the update $V_y \leftarrow \alpha \cdot V_x + V_y$. Specifying a $0 < \alpha < 0.5$ transfer parameter this simple algorithm yields an orderly descending series of stimulus values; with $k = 2$ and $\alpha = 0.1$, for example, the values for the series mentioned above come out as $V_a = 2.000$, $V_b = 1.200$, $V_c = 1.120$, $V_d = 1.112$, and $V_e = 0.112$. The test preference for B is explained by the higher value that this stimulus accordingly attains in comparison to Stimulus D.

As simple and efficient as this direct value transfer theory seemed to be, it was soon criticized by Couvillon and Bitterman (1992) for appealing to an unproven and unnecessary transfer mechanism. They showed that a more conventional, though more complex conditioning theory (Rescorla & Wagner, 1972) provided an excellent emulation of the Fersen et al. results. This in spite of the fact that the model assumes that choices of Stimulus X+ and the choices of Stimulus Y- out of the stimulus pair X+Y- result in both increases of the value V_x and decreases of the value V_y respectively without any specified transfer between them. Wynne, Fersen, and Staddon (1992; see also Wynne, 1995) indicated that some other similar but factually less complicated models than Couvillon and Bitterman's were equally capable of making the appropriate predictions. Subsequent research has shown that one of these models, based on Luce's (1959) beta operator and involving only two free parameters, is indeed quite successful in emulating a varied set of relevant data (Siemann, 1993, 1994; Siemann & Delius, 1994; Werner, Köppl, & Delius, 1992).

It is the case, as Wynne et al. already have noted, that all these conventional conditioning models unintentionally embody an indirect form of value transfer. The mechanism that brings about this transfer in transitive responding designs is an almost inevitable choice/reinforcement biasing process (Siemann, Delius, & Wright, in press). In

the typical design the Stimuli B, C, and D are apparently scheduled to be rewarded and penalized equally often. But this becomes fictive as soon as discriminative learning begins to take effect. Consider the stimulus pair A+B-: Because A is consistently rewarded it is increasingly chosen, meaning that B is decreasingly chosen and thus rarely penalized. This in turn encourages choices of B in the B+C- pair, where they are then rewarded. Soon choices of B are overall more often rewarded than penalized. Now consider pair D+E-. As E is consistently penalized, D is increasingly chosen and rewarded. That, however, favors penalized choices of D in the C+D- pair. Choices of D are therefore more evenly rewarded and penalized than those of B. In value terms Stimulus B can thus be seen as indirectly gaining from being paired with A, which is always rewarded and Stimulus D as indirectly losing from being paired with E, which is always penalized; Stimulus C is obviously intermediate in this emerging choice/reinforcement balancing process. The net effect of this process is a $V_a > V_b > V_c > V_d > V_e$ ranking of stimulus values (Siemann & Delius, 1996b).

The indirect value transfer brought about by the above mechanism across successive training trials is obviously based on a more intricate process than the direct within-trials transfer envisaged by Fersen et al. (1991). Their model assumed that the repeated presentations of a stimulus pair X+Y- induced an increasingly likely selection of Stimulus X, and that the reward that followed such choices fractionally and directly benefited the associative value V_y of the copresent Stimulus Y. Some time ago, in a brief preliminary note, we presented experimental evidence suggesting that despite the criticism leveled at it, the direct value transfer might still be a mechanism that has some truth in it (Siemann, Daniel, Dombrowski, & Delius, 1993). A successive rather than a simultaneous discrimination of instrumental stimuli was employed because this procedure allowed an uncomplicated addition of separate stimuli as recipients of transferred value, and because it entirely prevented the confusing intervention of the aforementioned indirect form of value transfer. The present article now describes and discusses the results of the relevant experiments in appropriate detail but also presents those of a further experiment pertinent to the issue.

At this point it is fair to mention that others have meanwhile presented independent evidence suggesting that direct value transfer does indeed occur (Zentall & Sherburne, 1994; Steirn, Weaver, & Zentall, 1995) and could probably play a role in transitive responding designed experiments. In fact, the simultaneous stimulus discriminations context they chose corresponds more closely with that employed in the transitive responding experiments that originally gave rise to the value transfer issue. Later, in the General Discussion there will be opportunity to enlarge on some of these findings. The present report pursues two interrelated aims. One is to demonstrate the existence of direct value transfer in a relatively transparent discrimination learning context. The other aim is to develop and advocate a specific conditioning mechanism that could be the basis of such value transfer.

Experiment 1

A corollary of the Fersen et al. (1991) hypothesis is that the size of the value transferred from an instrumental stimulus to an accompanying stimulus should relate to the strength of the reinforcement scheduled for responses to the instrumental stimulus. The present experiment accordingly involved several instrumental, or target stimuli, pecks directed at which yielded rewards and penalties of different magnitudes, as well as corresponding accompanying, or neutral stimuli, pecks aimed at which had no scheduled consequences. The varying qualities and amounts of reinforcement were expected to yield graded value transfers to these latter stimuli, which would be revealed in tests which directly compared their response eliciting potency.

Method

Six domestic pigeons (*Columba livia*) of local homing stock were used. They were housed in individual cages in a well-ventilated room with a 12-hr light on/ 12-hr light off regime and maintained food deprived at 80% of their normal weight. The pigeons had previous conditioning experience but not with the stimuli and procedures used here.

A special panel attached to the pigeons' home cages for the duration of the experimental sessions served as the conditioning apparatus (Fig. 1). It was hung outside and faced a 10-cm x 5.5-cm opening of the cage front wall, replacing the usual feeding trough. The main element of the panel was a horizontal platform (10 cm x 6.3 cm) that incorporated two pecking keys (2.5 cm diameter, centers 5 cm apart). The keys had 3 mm high rims which made them dish-shaped. When activated two separate overhead solenoid operated dispensers would discharge millet grains directly onto the corresponding key through tubing. Stimuli were produced using two miniature 7 x 5 red light emitting diode matrices mounted directly below the transparent keys. A personal computer (Commodore 40) with an interface card (Computer Board) generated the stimuli, recorded the responses, and issued the reinforcements (Xia, Delius, & Siemann, 1996).

The pigeons were trained to discriminate four target stimuli A++, B+, C-, and D- - using a successive instrumental discrimination procedure. The various symbols indicate that responses to the corresponding stimuli led to differently sized rewards and penalties. Each of the target stimuli was accompanied by a neutral stimulus (N_a , N_b , N_c , and N_d). Pecks directed at these stimuli had no consequences. The actual stimulus shapes were differently allocated to two equally sized subjects Groups I and II as a control for any spontaneous preferences (Fig. 1). In the event both groups yielded equivalent results, they were pooled for the purposes of evaluation.

Trials began with the presentation of a given target stimulus below one key and the corresponding neutral stimulus under the other key. The allocation of target and neutral stimuli to the left and right keys was

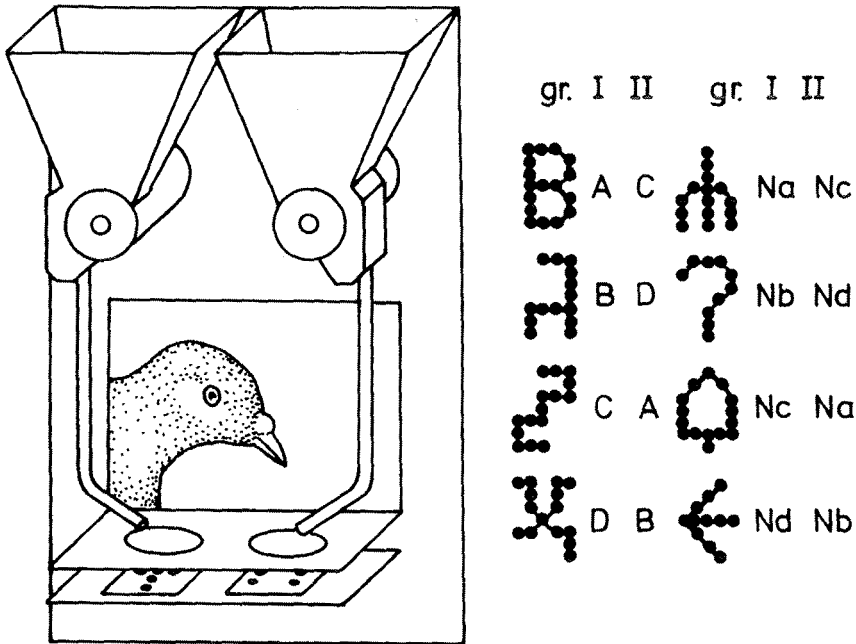


Figure 1.. Conditioning platform and stimuli employed in Experiment 1. The allocations of the stimuli to two subject groups is indicated.

quasi-random (Gellermann, 1933). Trials lasted until specific response requirements had been completed (see below) and reward or penalty had been issued, or else until 10 sec had elapsed. The on-key millet rewards arising from either a single (2 to 4 grains) or else a triple (6 to 12 grains) operation of the feeder corresponding to the target stimulus were followed by a 2-s feeding-time. Time-out penalties involved a lighting-up of all the elements of the corresponding stimulus matrix for 3 sec or alternatively 9 sec. A repeat trial allowed a correction choice. Trials were separated by a 2-s interval. Sessions were daily and consisted of 80 trials (not counting repeat trials), 20 with each stimulus pair, presented in a quasi-random sequence. During the first five training sessions the response requirement for reward was gradually increased from FR1 to FR8 pecks, the latter schedule being retained from then onwards. An FR1 schedule for penalty operated throughout. After the subjects had reached a criterion of 60% trials correct, a random 10th of the trials was no longer reinforced as a preparation for the feedback-free test trials. When the pigeons reached a criterion of at least 70% trials correct (not counting the outcomes of repeat trials) on each of the four target stimuli on two consecutive sessions, four test sessions were carried out. During these test sessions the six possible neutral stimulus combinations ($N_a N_b$, $N_a N_c$, etc.) were each presented twice interspersed among the normal training trials, with the left-right position of the member stimuli

being balanced. Pecks to the neutral stimuli were counted but continued to have no scheduled consequences. Thus every test session consisted of a total of 92 trials. Test trials lasted 10 s unless they were prematurely terminated by a total of 10 pecks.

Results and Discussion

Four pigeons required between 21 and 25 training sessions to reach a criterion of at least 70% correct trials with each of the four target stimuli. Two pigeons did not manage to reach this criterion within the predetermined maximum 25 training sessions and were excluded for expediency reasons. The successful birds exhibited a very good average discrimination on each of these stimuli during the subsequent test session (A+ +: 81%; B+: 84%; C-: 96%; D- -: 92% trials correct).

Five of the six test pairs included at least one neutral stimulus that had accompanied a rewarded target stimulus. Each of these five test pairs yielded, on average, more pecks to the neutral stimulus that had been paired with the most positively rewarded (most highly valued) target stimulus: N_aN_b , 66%; N_aN_c , 74%; N_aN_d , 69%; N_bN_c , 77%; and N_bN_d , 90%. The neutral stimuli composing the sixth test pair had both been associated with penalized target stimuli and it is perhaps not too surprising that the pigeons did not evince a clear preference for the neutral stimulus connected with the least penalty (N_cN_d , 57%). We will return to these test results in more detail later in the General Discussion when a more precise understanding of the value transfer process can be applied. Meanwhile, we note that regarding the above 5 first test pairs each pigeon exhibited a significant overall preference for the neutral stimulus associated with the best reinforced target stimulus (binomial tests, $ps < 0.05$).

The total pecking rates during the test trials corresponding to each of the neutral stimuli shown in Figure 2 were even more informative. The relevant rates varied in significant agreement with the reinforcement magnitudes associated with responses to the corresponding target stimuli ($N_a > N_b > N_c > N_d$; Page-Test, $L = 117$, $p < 0.01$).

Because of the design of the present experiment (successive discrimination procedure, separate target, and neutral stimuli) there is no way in which the indirect transfer mechanism mentioned in the Introduction can be made responsible for the present transfer results. Therefore, the transfer demonstrated must be caused by a different, more direct type of value transfer from the instrumental target stimuli to the non-instrumental, but spatially contiguous and temporally contingent neutral stimuli. Moreover, as the Fersen et al. (1991) theory demands, the value transferred to the neutral stimuli, as reflected in response rates that they later supported, is obviously commensurate with the reinforcement outcomes and magnitudes pertaining to the corresponding target stimuli.

diameter, centers 4 cm apart) with three small cups (1 cm diameter, 0.5 cm deep) behind them. Three overhead food dispensers could separately deliver millet into these cups (Fig. 3). Only the middle and left keys and respective feeders were used in this experiment; the right-hand key was covered with dark tape. A small lightbulb set in the roof of the alcove served as houselight. Four shapes of similar size and complexity served as stimuli. They were back-projected white on black onto the translucent keys with a slide projector (Kodak) via a 45° mirror and were enabled by separate solenoid shutters. The chamber and projector were controlled by a home computer (Commodore 64) through an interface (Dela; Xia, Wynne, Münchow-Pohl, & Delius, 1991).

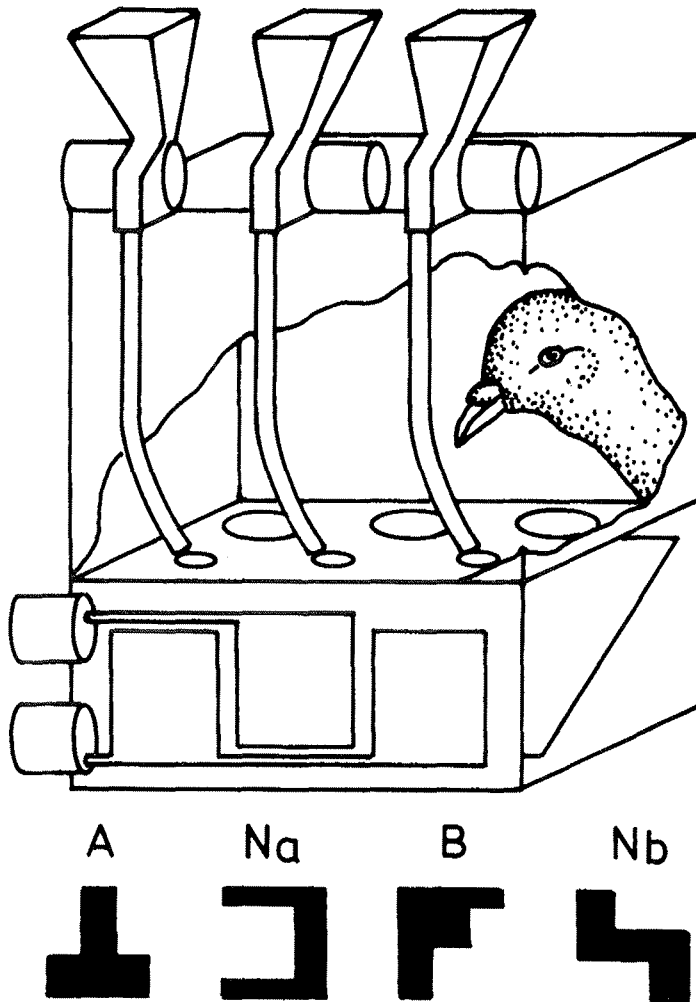


Figure 3. Conditioning alcove and stimuli used in Experiments 2 and 3.

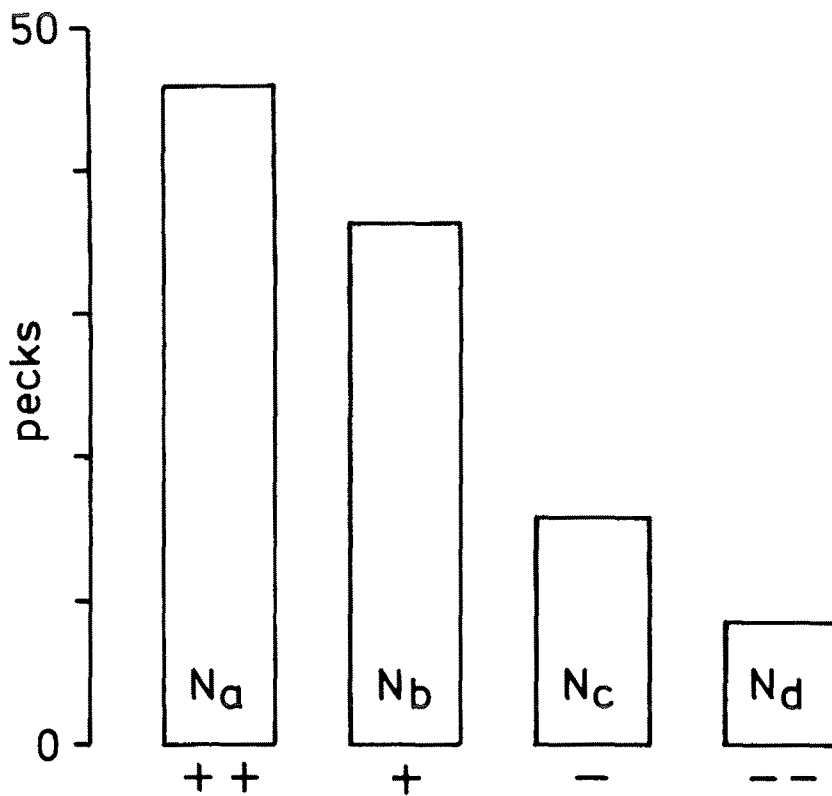


Figure 2. Mean number of pecks to the four neutral stimuli during test trials of Experiment 1. The symbols below the columns refer to the reinforcements grades pertaining to the target stimuli with which the different neutral stimuli were associated during training.

Experiment 2

This experiment, although modeled on the previous one, used only two target stimuli and two accompanying neutral stimuli. Its purpose was to ascertain the robustness of the direct value transfer effect using a discrimination reversal design which would allow to dismiss any explanations based upon spontaneous stimulus preferences. By paying attention to the responses issued to the neutral stimuli during training it also focused on the genesis of the transfer phenomenon.

Method

Six new, already trained to key peck, pigeons were used. They were housed and deprived as in Exp. 1, but were conditioned in a separate chamber. Opposite side-walls of this chamber were equipped with a one-way observation window and with a recessed alcove. The horizontal surface (20 x 5 cm) of the alcove bore three pecking keys (2.5 cm

During the initial training sessions, trials began with either Stimulus A+ or Stimulus B- being quasi-randomly (Gellermann, 1933) presented on one of the two keys. The other key remained dark. Positive trials lasted 15 s unless a peck towards A yielded two to four grains of millet delivered into the reward cup right next to it. A 2-s feeding-time followed and terminated the trial. Negative trials also lasted 15 s unless a peck to B led to an abortion of the trial followed by a 5-s time-out with the houselight turned off, followed by a repetition of the trial. Trials were separated by 2-s intervals. Within the first five sessions the response requirement for reward and trial termination was gradually increased from a FR1 to a FR10. The FR1 that led to penalty and trial termination was not altered. The daily sessions consisted of 64 trials.

When the pigeons yielded 60% correct trials ending in reward or no penalty (not counting the outcomes of repeat trials) in a given session, the satellite stimuli were introduced. They were displayed on the key not occupied by the target stimuli A+ or B-. Stimulus N_a was always paired with A ($A+N_a$), N_b was always paired with B ($B+N_b$). Pecks to these neutral stimuli were recorded, but they had no other programmed consequence. After 5 sessions, the next 10 sessions were run without repeat trials and additionally, a random eight of the trials were not reinforced either way as a preparation for later feedback-free test trials.

Two test sessions, separated by four normal training sessions as above, followed. During the test sessions four random training trials were replaced by as many test trials, involving the simultaneous presentation of N_a and N_b on the two keys for a maximum of 15 s. The right/left allocations of the two stimuli were balanced. Pecks during these test trials were separately counted. They had no other consequence except that 10 pecks delivered to any of the stimuli prematurely terminated the test trial.

In the second part of the experiment the reinforcement allocations of the target stimuli were reversed to B+A-. Stimulus A, that had been rewarded before, was now penalized with time-out, whereas Stimulus B had now to be pecked in order to obtain food reward. This was to control for any spontaneous stimulus preferences and to explore the reversibility of value transfer. During the initial reversal sessions, until the subjects achieved 60% trials correct, only the target stimuli were presented. The response requirement for reward with Stimulus B was gradually increased within the first five sessions from a FR1 to FR10 but to facilitate relearning the animals were always rewarded at the end of B+ trials regardless of schedule completion. Single pecks to A- yielded penalties and trial repetition as described earlier. As soon as the animals reached the aforementioned criterion the neutral stimuli were again introduced, N_b accompanying B+ and N_a accompanying A-. Reward was now only delivered after fulfillment of the FR10 requirement on B. After 10 sessions the repeat trials were discontinued and 1/8th of the trials were not reinforced. When five sessions had been completed, a first test session, four further training sessions, and a final test session were run. Within each test session the test stimulus pair $N_a N_b$ was presented four times.

Results and Discussion

During the initial training with target stimuli alone the pigeons reached the criterion of 60% trials correct within 6 to 10 sessions. Upon their introduction, the neutral stimuli were pecked quite frequently. During the first five sessions with them, the subjects directed an average 62 and 246 pecks per session to respectively N_a and N_b . The difference between these response rates suggests that at least some of the responses to N_b were caused by a response displacement away from B- to avoid the penalties associated with reactions directed at this latter stimulus. However, the responding to and the difference between the two neutral stimuli decreased markedly in the course of a few sessions. During the last five training sessions the average per-session pecks to N_a and N_b were only 22 and 35 respectively.

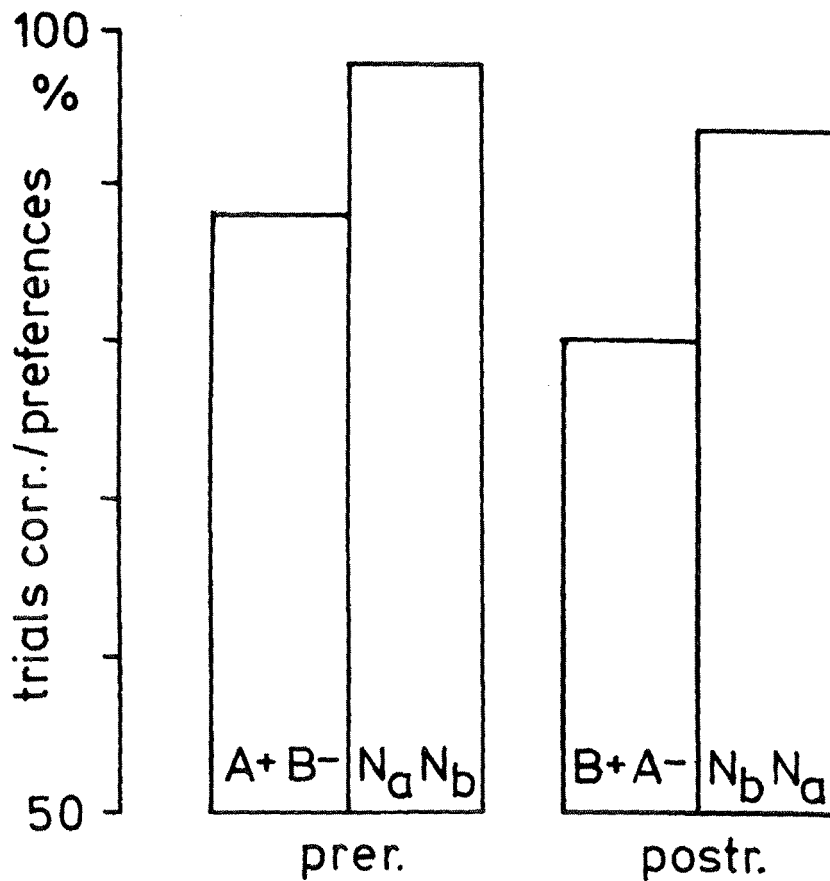


Figure 4. Mean percent correct trials with target stimuli and mean percent preference scores with neutral stimuli during the test sessions corresponding to the prereversal (prer.) and postreversal (postr.) phases of Experiment 2.

The prereversal test sessions revealed a clear-cut target stimulus discrimination with a mean 95% of the A+ trials terminating with reward and a mean 82% of the B- trials ending without penalty (Fig. 4). The relative preference for N_a in the $N_a N_b$ test trials amounted to a mean 98% pecks out of the average total of 55 pecks per pigeon to either stimulus. This result shows that at least N_a had benefitted from some kind of value transfer.

The reversal training to the 60% correct trials criterion with the target stimuli B+ and A- required between 9 and 22 sessions. One of the 6 pigeons ceased to peck altogether at this stage and had to be excluded. In the subsequent phase, when the target stimuli were again paired with the neutral stimuli N_b and N_a , the latter stimulus was initially pecked more often than the former but this difference reduced as training progressed. During the first five sessions there was an average of 40 and 121 pecks per session for N_b and N_a respectively but for the last five sessions these scores were respectively, 25 and 37 pecks per session.

The results of the postreversal test sessions are summarized in Figure 4. The target stimuli were again discriminated at a high level (81% B+ trials correct, 78% A- trials correct). Though the pecking rates during the test trials were somewhat lower (on average a total of 32 pecks per pigeon) than during the prereversal test trials, N_b was now clearly preferred (93% choices) over N_a . This reversal demonstrated that the neutral stimulus choices recorded were not caused by a preexisting spontaneous preference and that N_b could similarly benefit from value transfer.

The discrimination shown during both pre- and postreversal tests with the stimulus pair $N_a N_b$ seems stronger than that concurrently evinced with the training stimuli A and B (Fig. 4). However, not too much can be made of this because, whereas the training involved a successive discrimination, the testing involved a simultaneous discrimination. The corresponding performances had to be scored somewhat differently and it is also a general finding that simultaneous paradigms tend to yield a superior discriminative performance than successive paradigms (Mackintosh, 1974). The results, in any case, are in agreement with the assumption that some of the instrumental value of the target stimuli can transfer to contiguous and contingent satellite stimuli. Both after the original discrimination training and after its reversal the pigeons preferred the neutral stimulus that had been correspondingly paired with the rewarded target stimulus.

What mechanism might have brought about the direct value transfer demonstrated? In an aside, Davis (1992) pointed out that classical conditioning adventitious to the instrumental procedure might possibly generate value transfer. We take up this suggestion and consider whether the idea can be applied to the present experiments and its results. During the prereversal phase for instance, pecks at the target stimulus A+ (instrumental S^D) led to food rewards, pecks to the other target stimulus, B- (instrumental S^A) led to time-out penalties. The neutral stimuli shown on the alternative key can however be viewed,

respectively, as a CS+ (N_a) and as a CS- (N_b). Responses to these stimuli had no instrumental consequences, but N_a quite regularly preceded the occurrence of a certainly appetitive US (food) and N_b less regularly, the occurrence of a probably aversive US (time-out). Given these contingencies the incidence of classical conditioning with respect to at least N_a seems inescapable. The well known tendency of pigeons to develop CR pecking towards discrete CS+ (autoshaping: Brown & Jenkins, 1968) should become effective. However, given the A+ N_a combination, the tendency to peck N_a because of classical conditioning, would be overridden by the tendency to peck A+ because of the instrumental contingency. This means that the associative value increments accruing to N_a would be largely latent during the training phase and would only reveal themselves during the later tests that paired N_a with N_b . This latter stimulus might, in turn, be expected to enter these tests with a slight associative value decrement, due to the occasional classical inhibitory contingency it had been exposed to while being a member of the B- N_b training pair.

However, other mechanisms capable of effecting a value transfer must also be considered. A pseudooperant conditioning process could have intervened (Schlosberg, 1934). Pecks to N_a or N_b had no programmed consequences, but within a training trial the pigeon sometimes switched between pecking N_a and pecking A or alternatively between pecking N_b and pecking B, eventually being rewarded or penalized for pecking the target stimuli. Thus, in at least some training trials pecks to the satellite stimuli might have appeared as being eventually reinforced. Alternatively, there could have been a differential devaluation of the neutral stimuli during the training phase. At the beginning of training the pigeons pecked more at the neutral stimulus accompanying the negative target stimulus (the response displacement effect alluded to before) than at that accompanying the positive target stimulus. Towards the end of training the pecking at the neutral stimuli had diminished and was more evenly distributed between both neutral stimuli. That could possibly have been associated with a more pronounced extinction affecting the negative neutral stimulus than the positive neutral stimulus. To assess the role of both these alternative accounts, in the next experiment some subjects were simply prevented from pecking the neutral stimuli.

Experiment 3

This experiment included a condition intended to prevent the operation of both pseudooperant conditioning and extinction-occasioned devaluation, while still allowing the development of classical conditioning. One group of pigeons was trained much as in Experiment 2, but another group was prevented from pecking the neutral stimuli by transparent key enclosures. If transfer was caused by the pseudooperant process or the instrumental extinction process alone the latter group should not yield

any evidence of transfer. If, however, adventitious classical conditioning was important they should do so, as it has been amply demonstrated that classically conditioned pecking develops even when the response to the relevant CS is prevented during the training phase (Parisi & Matthews, 1975; Richardson & Hansen, 1980).

Method

Eight new adult pigeons were used. The same chamber with conditioning alcove as in Experiment 2 was used but with all three pecking keys in operation. Also, the same stimuli were employed but the allocation of the shapes serving as N_a and N_b was exchanged. This was meant to exclude the unlikely possibility that the transfer observed in Experiment 2 was caused by a simple stimulus generalization from target to neutral shapes. The procedure was equivalent to that of Experiment 1 but incorporated some modifications. The birds were randomly allocated to two equally sized groups. Group F subjects were free to reach all three keys of the conditioning alcove. Group P subjects had only free access to the middle key and although they could see the side keys, they were prevented from reaching them because the right and left sections of the alcove containing them were separated off by two transparent enclosures. The P group subjects were also extensively watched through the rear one-way window of the chamber. Otherwise the two groups were subjected to the same procedure. The target stimuli A+ or B- always appeared on the middle key while the corresponding neutral stimulus N_a or N_b was shown on one of the two side keys. The sequencing of the A+ N_a and B- N_b pairs, as well as the left or right key allocation of the neutral stimuli, was quasi-random (Gellermann, 1933). During test sessions no enclosures were used with either group so that P group pigeons could now also peck the side keys. In test trials the middle key remained dark while N_a and N_b were simultaneously presented on the side keys, with quasi-randomly varying side allocations.

Results and Discussion

The pigeons achieved the initial A+B- discrimination criterion of 60% trials correct within 3 to 12 sessions. There was no significant difference between the F and P groups. Two pigeons, one from each group, were later excluded as they did not reach a predetermined criterion of 70% trials correct within the preset 15 training sessions involving the A+ N_a and B- N_b presentations. Extended and repeated observations showed that the transparent side-key enclosures effectively suppressed any peck attempts towards the satellite stimuli by the group P subjects. This behavior may have been aided by the fact that the enclosures were present from the beginning of training onwards, that is even before any neutral stimuli were presented on the side keys, the animals thus learning early on that the corresponding areas of the alcove were totally out of reach.

The discriminatory performance during the test phase was closely

similar for both the F and P groups (Fig. 5). As might be expected according to the above observations the P group issued on average, fewer pecks during $N_a N_b$ tests than the F group (a total of 29 pecks against 38 pecks), but the difference was not significant. Each pigeon in both groups exhibited a significant preference for N_a over N_b (binomial tests, $ps < 0.05$). Because the average preference for N_a was if anything stronger in the P group pigeons (Fig. 5) that had effectively not pecked the neutral stimuli during training, the preference for N_a is highly unlikely to have been determined by the pseudooperant contingency or the displacement-extinction process expounded in connection with the discussion of the Experiment 2 results. The incidence of both these instrumental processes depends essentially on the occurrence of responses targeted at the relevant neutral stimuli. The F group results

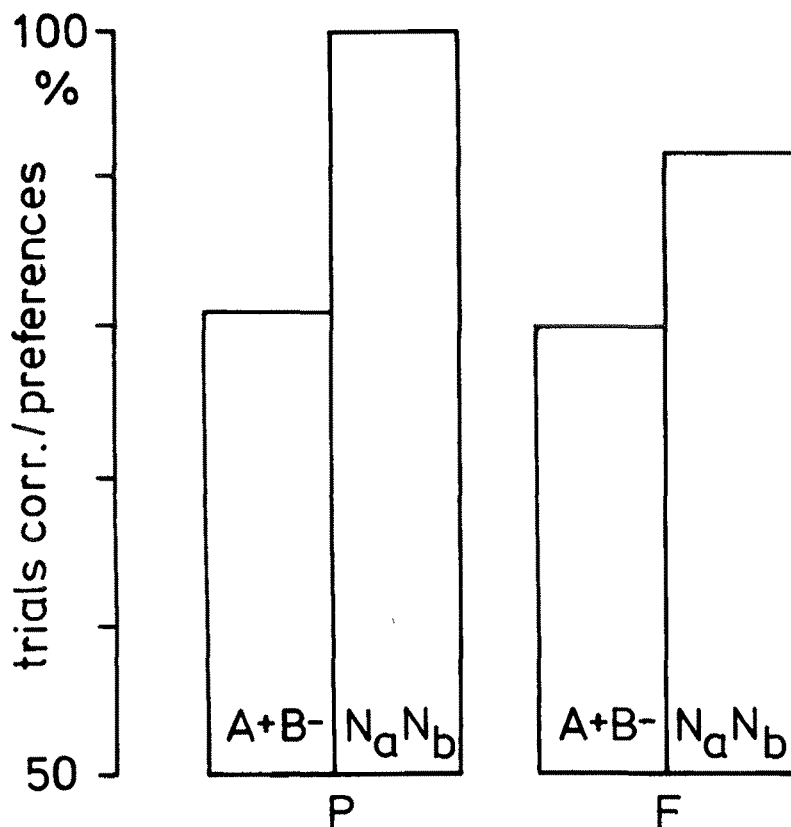


Figure 5. Mean percent correct trials with target stimuli and mean percent preference with neutral stimuli for the F group and P group of pigeons during the test sessions of Experiment 3. The F pigeons had been free to peck the neutral stimuli during training, the P pigeons had been prevented from doing so.

also contribute somewhat to this argument insofar as the relevant subjects issued far fewer pecks to these stimuli during training than their counterparts had done in Experiment 2 (an average of 3 and 14 pecks per session during the first five sessions and 10 and 13 pecks during the last five sessions to N_a and N_b respectively). In spite of this, the F birds still showed a good test discrimination between the neutral stimuli. Their reduced pecking at the neutral stimuli may have been because in the present experiment the target stimuli were consistently allocated to the central key and the neutral stimuli were consistently allocated to the side keys, with food rewards delivered only next to the central key. In Experiment 2 the equivalent allocations had freely alternated between the two relevant keys and that might have generated a certain amount of confusion about their instrumentalities. Note also that the combined results of Experiments 2 and 3 exclude an explanation in terms of stimulus generalization as the correspondences between the target and the neutral stimuli were exchanged between the two experiments.

General Discussion

Comments

The experiments described above investigated whether a direct value transfer can occur between stimuli that control an operant discrimination and stimuli that are contingent with them but do not intervene in the operant paradigm. All experiments demonstrated the existence of such a transfer. Experiment 1 showed that varying magnitudes of target stimulus reinforcement yielded graded value transfers to the corresponding neutral stimuli. Experiment 2 also showed that the preference for the neutral stimuli reversed when the reinforcements associated with the target stimuli were reversed and suggested a classical conditioning explanation. Experiment 3 allowed dismissing the alternative pseudooperant and response extinction accounts of direct value transfer. Pigeons prevented from responding to the neutral stimuli during training evinced no weaker preference than those allowed to do so.

As already mentioned in the introduction there are other experimental results that support the existence of a direct value transfer phenomenon. Zentall and Sherburne (1994) trained pigeons to discriminate simultaneously and concurrently two pairs of stimuli A++B- and C+D-. Responses to the strongly positive stimulus of the first pair were always rewarded, those to the weakly positive stimulus of the second pair were only partially rewarded. During test trials the negative stimuli of both pairs were shown paired B D. As expected according to the value transfer hypothesis of Fersen et al. (1991) the B stimulus was preferred. Steirn et al. (1995), among other tasks, trained pigeons with the stimulus pairs A+B-, A+E-, C+D-, and E+C-. Differently from the usual transitive responding design (A+B-, B+C-, C+D-, D+E-), the stimuli forming these training pairs did not fit into a linear sequence and could

probably not be subject to an indirect transfer caused by a choice/reinforcement biasing process as described in the Introduction. Nevertheless, the pigeons still showed a preference for B in test pair B D as might be expected from a direct value transfer point of view, B having been paired as a negative stimulus with a twice rewarded A, and D having been paired as a negative stimulus with a half rewarded, half penalized C. This suggests such transfer could indeed play a role in transitive responding designs. However, it is fair to say that the two studies, much as the Fersen et al. (1991) study, are not particularly specific about the precise mechanism that might have brought about the direct transfer they recorded. Higher order conditioning, within-event learning, and spatio-temporal generalization are very briefly mentioned as explanatory alternatives (Zentall & Sherburne, 1994).

We suggest that the classical conditioning account developed in connection with Experiment 2 and upheld by the results of Experiment 3 can also be applied to the results of Zentall and colleagues. When subjects learn to operantly discriminate simultaneously presented stimuli X+Y- and begin to respond nearly exclusively to Stimulus X, this stimulus functions as a S^D eliciting the instrumental response. Stimulus Y, on the contrary, functions as a S^A which does not do that. But Y can also be viewed as a CS+ insofar as its appearances are now quite regularly followed by food presentations, an undoubtedly effective appetitive US. Under these circumstances autoshaping (Brown & Jenkins, 1968) can be expected to take its course. However, the emerging pecking CR in response to Stimulus Y would be of a largely latent nature, because in most trials the subject would still be most likely to peck X, the stimulus that undisputably predicted reward best. From a similar vantage point, responses to Stimulus Y effect an instrumental US omission procedure because they effectively abort any presentation of food at the end of trial (Sheffield, 1965). In pigeons, this paradigm in its pure form is known to weaken the pecking CR due to a classical CS-US contingency, but to be incapable of suppressing it altogether (Moore, 1973; Williams & Williams, 1969). Thus, Stimulus Y can nevertheless be assumed to eventually accrue some CS+ value despite its formal S^A status. The circumstance that Y- is soon rarely chosen because of the aversive instrumental consequence that this has, ensures that X, which in principle is a CS-, can only lose little in value through this fact. Thus, this adventitious respondent mechanism could also bring about a direct value transfer in the simultaneous discrimination context considered by Fersen et al. (1991).

The claim that a direct value transfer due to classical conditioning exists does not of course deny that other mechanisms can also mediate value transfer. One of them, the gradual, differential reinforcement ratio biasing process outlined in the Introduction, is virtually inevitable and certainly very important in the transitive responding experiments that gave rise to the value transfer issue (Siemann, 1993, 1994; Wynne, 1995; Wynne et al., 1992). A pseudooperant mechanism, diagnosed as being insignificant in the present experiments, may nevertheless play a

role in other contexts (Mackintosh, 1974). A further mechanism, that of stimulus generalization similarly dismissed here, undoubtedly can mediate value transfer when physically similar stimuli are involved (Rilling, 1971). But the fact that the development of pecking at a key that repeatedly displays a stimulus which is mostly followed by food access is an exceedingly robust phenomenon with pigeons (Schwartz & Gamzu, 1977) indicates that the incidence of classical conditioning should not be ignored.

Modeling

The classical conditioning idea can be embodied in a more specific algebraic model than the very general value transfer model proposed by Fersen et al. (1991). As already mentioned a simple model based on Luce's (1959) β -operator has proved to yield reasonable emulations of several transitive responding results (Siemann, 1993; Siemann & Delius, 1993, 1996a; Werner et al., 1992). We now propose a modification that incorporates a value transfer component caused by classical conditioning and that accounts well for the present findings.

In Luce's model the associative (response eliciting) value of a stimulus is increased by a certain amount if the choice of that stimulus is followed by reward and decreased by a certain amount if it is followed by penalties. Thus, given that the stimulus pair X+Y- is presented and the subject chooses X, being accordingly rewarded, the value of this stimulus is updated according to $V_x \leftarrow V_x + (V_x * \beta_+)$ where β_+ is a learning parameter corresponding to reward. If, however, the subject chooses Y then its value is updated according to $V_y \leftarrow V_y - (V_y * \beta_-)$ where β_- is the parameter corresponding to penalty. The current values V_x and V_y in turn determine the probability with which X will be chosen in preference to Y in a given trial according to $p_{xy} = V_x / (V_x + V_y)$. For a larger population of subjects participating in the same experiment, the average value changes in a given X+Y- trial are thus approximated by $V_x \leftarrow V_x + (V_x * \beta_+ * p_{xy})$ and $V_y \leftarrow V_y - (V_y * \beta_- * (1 - p_{xy}))$.

Although Luce (1959) left it open whether his β -operator referred to classical or instrumental learning, as applied above it only enforces the latter type of learning. A model modification proposed by Siemann and Delius (1996b), however, incorporates an additional classical conditioning mechanism that operates much as that described earlier in connection with simultaneous stimulus discriminations. This model is now adapted to the successive discrimination procedure used in the present experiments. The adaptation assumes that the probability of a correct X+ trial is given by $p_{xt} = V_x / (V_x + T)$ and that of an incorrect Y- trial is given by $p_{yt} = V_y / (V_y + T)$, where T represents the value of some constant tendency not to respond to either kind of target stimulus. Conversely, of course, the probability of an incorrect X+ trial is $p_{tx} = 1 - p_{xt}$ and the probability of a correct Y- trial is $p_{ty} = 1 - p_{yt}$. Given the stimulus pair X+N_x and given that a rewarded response to Stimulus X+ occurs, it is assumed that value V_{nx} of Stimulus N_x benefits by a fraction τ from the increment $V_x * \beta_+ * p_{xt}$ accruing to V_x because it is effectively a

conditioned stimulus with respect to an unconditioned stimulus (food) that is $V_{nx} \leftarrow V_{nx} + (\tau \cdot V_x \cdot \beta_+ \cdot p_{xt})$. A no-response, no-outcome incorrect trial with Stimulus X+ does not alter V_x , V_{nx} , or T. Given the stimulus pair Y-N_y and a penalized response to Stimulus Y- (with its value being updated $V_y \leftarrow V_y - (V_y \cdot \beta_- \cdot p_{yt})$), the value N_y is decreased according to $V_{ny} \leftarrow V_{ny} - (\tau \cdot V_y \cdot \beta_- \cdot p_{yt})$. A no-response, no-outcome, correct trial with Stimulus Y- does not alter V_y , V_{ny} , or T. Note that the parameter τ corresponds roughly to the parameter α of the Fersen et al. (1991) model. The percentage correct responses for target stimulus X is computed with $p_{xt} = 100 \cdot V_x / (V_x + T)$ and for target stimulus Y with $p_{ty} = 100 \cdot T / (V_y + T)$ and for test pairs with $P_{nx} = 100 \cdot V_{nx} / (V_{nx} + V_{ny})$.

To illustrate the capabilities of this instrumental-classical conditioning model an emulation of the relatively complex transfer results of Experiment 1 was carried out. The constant T and the variables V_x , V_{nx} , V_y , and V_{ny} were all set to the same small, arbitrary initial value. To take into account the various graded reinforcements used in that experiment there had to be four different learning rate parameters. The values $\beta_+ = 0.10$, $\beta_- = 0.08$, $\beta_+ = 0.52$, $\beta_- = 0.74$, for these parameters and $\tau = 0.30$ for the transfer factor with 28 stimulation trials were found to yield an acceptable emulation of the empirical results (Table 1). Notice that because of the threat of extinction to unreinforced test pairs, the observed data concerning them is based on relatively few replicate trials (see Exp. 1, Methods). The fact that this adventitious transfer model could still produce approximately correct predictions for the neutral stimulus test choices of Experiment 1 and, although not detailed here, also the simpler ones of Experiments 2 and 3, suffices to demonstrate that a minor modification of a conventional and elementary conditioning model is capable of yielding transfer based on classical conditioning in an instrumental conditioning setting (see also Siemann & Delius, 1996b).

Table 1

Predicted and Observed Percent Correct Trials during Test Phase
for Target Stimuli and Test Pairs of Experiment 1

	Target Stimuli				Test Pairs					
	A+ +	B+	C-	D- -	N _a N _b	N _a N _c	N _a N _d	N _b N _c	N _b N _d	N _c N _d
Predicted	86	80	93	95	58	78	77	72	71	49
Observed	81	84	96	92	66	74	69	77	90	57

Conclusion

We believe that the results reported and the literature reviewed make it plausible that an adventitious classical conditioning can take effect in the transitive responding paradigm based on instrumental discriminations and described in the introduction. The net effect of the classical conditioning account is in substantial agreement with Fersen et al.'s (1991) sketchy theory insofar as it generates a direct, within-trial value transfer across stimuli. Elsewhere we have described a variant of

the theoretical model described here that expressly fits the transitive responding design (Siemann & Delius, 1996b) and that is capable of emulating the contribution of the classical process to the transitive responding phenomenon.

It will nevertheless remain difficult to assess any quantitative contribution of direct value transfer/adventitious classical conditioning to empirical transitive responding results. That is because in the A+B-, B+C-, C+D-, D+E- design it acts synergistically with the already repeatedly mentioned and unquestionably potent instrumental choice/reinforcement biasing process (Siemann & Delius, 1996a, Wynne, 1995). There are no obvious experimental procedures capable of cleanly separating the effect of the two mechanisms. Even Steirn et al.'s (1995) aforementioned design can not quite get at the heart of this insidious problem.

Simpler instrumental discrimination designs of the X+Y- type may however offer alternative and interesting prospects as Zentall and Sherbourne (1994) have already intimated. Indeed, concerning such discriminations there is older evidence suggesting that the response inhibiting effect that the S^A is supposed to acquire through the instrumental contingencies can be weakened when the S^D exerts overwhelming and persistent control over responding in such tasks. This should be so because the CS+ quality of the nominal S^A then becomes dominant. When pigeons are overtrained on such discrimination, rather than just trained to criterion, Stimulus Y progressively loses inhibitory stimulus control as assessed by generalization gradient measurements (Yarczower & Curto, 1972). More extremely, when pigeons are trained on an X+Y- task using an error-minimizing technique, B can end up devoid of any detectable inhibitory effect (Terrace, 1966). Overtraining again an X+Y- discrimination in pigeons facilitates the acquisition of a subsequent reversed discrimination X-Y+, indicating that the stimulus Y has a less inhibitory effect after overtraining than before (Williams, 1967). It is true that these last two effects have not replicated too readily (Mackintosh, 1974; Rilling, 1977), but unsuccessful attempts might have incorporated procedural details that were unfavorable to the classical conditioning process that we believe is essential for their emergence. The fact that pigeons trained to discriminate an infrequent stimulus pair C+D- and a frequent and thus overtrained A+B- stimulus pair paradoxically prefer B to D in corresponding tests (Biederman, Poulos, & Heighington, 1976) appears not to have been similarly challenged. As far as we are aware nobody has explicitly considered the role of adventitious classical conditioning in connection with these effects. The classically mediated, direct transfer mechanism suggests a new vista on these old phenomena, an issue that we pursue elsewhere (Siemann & Delius, 1996c).

There is of course a large body of experimentation and even larger body of argumentation concerned with the potential incidence of classical conditioning upon the stimulus control of instrumental responding (Domjan, 1992; Mackintosh, 1983). Nevertheless, none of this literature seems to squarely address the special question that was

the theme of the present experiments. These involved two types of stimuli, target stimuli (instrumental stimuli), responses directed to which resulted in differential reinforcement outcomes, and neutral stimuli (conditioned stimuli) that accompanied these stimuli but responses to which had no reinforcement outcomes. Could the neutral stimuli acquire some response eliciting, or respectively, inhibiting properties because they were respondently contingent with the operantly mediated reinforcements? The aforementioned literature, applied to the present context, seems to focus instead on the more central but obviously much more difficult question whether the target stimuli, and not just only the neutral stimuli, might have acquired some of their response eliciting potency through the analogous adventitious classical contingency that they were naturally also exposed to. We have not been concerned with this latter question because its solution, whatever it may be, is largely immaterial for the value transfer phenomenon that started this study. The Fersen et al. (1991) direct value transfer idea, in any case, has certainly proven to be less preposterous than once thought and to be heuristically fruitful. It has pointed to a factual phenomenon that still needs further elucidation.

References

- BIEDERMAN, G. B., POULOS, C. X., & HEIGHINTON, G. A. (1976). Paradoxical preference for more frequently occurring negative stimuli and for less frequently occurring positive stimuli as a function of amount of training in simultaneous discrimination teaming. *Learning and Motivation*, 7, 603-613.
- BROWN, P. L., & JENKINS, H. M. (1968). Auto-shaping of the pigeons key peck. *The Journal of Experimental Analysis of Behavior*, 11, 1-8.
- COUVILLON, P., & BITTERMAN, M. E. (1992). A conventional conditioning analysis of "transitive inference" in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 18, 308-310.
- DAVIS, H. (1992). Logical transitivity in animals. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 405-429). Hillsdale, NJ: Erlbaum.
- DOMJAN, M. (1992). *The principles of learning and behavior* (3rd ed). Brooks-Cole: Belmont.
- FERSEN, L. VON, WYNNE, C. L. D., DELIUS, J. D., & STADDON, J. E. R. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 334-341.
- GELLERMANN, C. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *Journal of Genetic Psychology*, 42, 206-208.
- LUCE, D. R. (1959). *Individual choice behavior*. New York: Wiley.
- MACKINTOSH, N. J. (1974). *The psychology of animal learning*. London: Academic Press.
- MACKINTOSH, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.

- MOORE, B. R. (1973). The role of directed Pavlovian reactions in simple instrumental learning in the pigeon. In R. A. Hinde & J. Stevenson-Hinde (Eds.), *Constraints on learning* (pp. 159-186). London: Academic Press.
- PARISI, T., & MATTHEWS, T. J. (1975). Pavlovian determinants of the autoshaped keypeck response. *Bulletin of the Psychonomic Society*, 6, 527-529.
- RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- RICHARDSON, W. K., & HANSEN, S. (1980). Autopecking with gross movement physically restrained. *The Psychological Record*, 30, 39-46.
- RILLING, M. (1977). Stimulus control and inhibitory processes. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 432-480). Englewood Cliffs: Prentice Hall.
- SCHLOSBERG, H. (1934). Conditioned responses in the white rat. *Journal of Genetic Psychology*, 45, 303-305.
- SCHWARTZ, B., & GAMZU, E. (1977). Pavlovian control of operant conditioning. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 53-97). Englewood Cliffs: Prentice Hall.
- SHEFFIELD, F. D. (1965). Relation between classical conditioning and instrumental learning. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium* (pp. 302-322). New York: Appleton-Century-Crofts.
- SIEMANN, M. (1993). *Transitive Inferenz: Experimentelle Untersuchung einer kognitiven Leistung* (Transitive inference, investigation of a cognitive performance). Universität Konstanz: Dissertation.
- SIEMANN, M. (1994). Überprüfung einfacher Modelle zum transitiven Schlußfolgern bei nonverbaler Aufgabenstellung (Test of a simple model of transitive inference using a nonverbal task). *Zeitschrift für experimentelle und angewandte Psychologie*, 41, 584-616.
- SIEMANN, M., DANIEL, D., DOMBROWSKI, D., & DELIUS, J. D. (1993). Value transfer in pigeon conditioning. In N. Elsner & M. Heisenberg (Eds.), *Gene, brain, behavior* (p. 856). Stuttgart: Thieme.
- SIEMANN, M., & DELIUS, J. D. (1993). Implicit deductive responding in humans. *Naturwissenschaften*, 80, 363-366.
- SIEMANN, M., & DELIUS, J. D. (1994). Processing of hierarchic stimulus structures has advantages in humans and animals. *Biological Cybernetics*, 71, 531-536.
- SIEMANN, M., & DELIUS, J. D. (1996a). Influence of task concreteness upon transitive responding in humans. *Psychological Research* (in press).
- SIEMANN, M., & DELIUS, J. D. (1996b). *Algebraic learning and neural network models for transitive and nontransitive responding in humans and animals*. Manuscript submitted for publication.
- SIEMANN, M., & DELIUS, J. D. (1996c). Overlearning and company revisited. Manuscript submitted for publication.
- SIEMANN, M., DELIUS, J. D., & WRIGHT, A. A. (in press). Transitive responding in pigeons: influence of stimulus frequency and reinforcement history. *Behavioral Processes*.
- STEIRN, J. N., WEAVER, J. E., & ZENTALL, T. R. (1995). Transitive inference in pigeons and a test of value transfer theory. *Animal Learning & Behavior*, 23, 76-82.

- TERRACE, H. S. (1966). Stimulus control. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 271-344). New York: Appleton-Century-Crofts.
- WERNER, U., KÖPPL, U., & DELIUS, J. D. (1992). Transitive Inferenz bei nichtverbaler Aufgabendarbietung (Transitive inference with nonverbal task presentation). *Zeitschrift für experimentelle and angewandte Psychologie*, 39, 662-683.
- WILLIAMS, D. I. (1967). The overtraining reversal effect in the pigeon. *Psychomic Science*, 7, 261-262.
- WILLIAMS, D. R., & WILLIAMS, H. (1969). Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. *The Journal of the Experimental Analysis of Behavior*, 12, 511-520.
- WYNNE, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning & Behavior*, 23, 207-217.
- WYNNE, C. D. L., FERSEN, L. VON, & STADDON, J. E. R. (1992). Pigeons' inferences are transitive and the outcome of elementary conditioning principles: A response. *Journal of Experimental Psychology: Animal Behavior Processes*, 18, 313-315.
- XIA, L., DELIUS, J. D., & SIEMANN, M. (1996). A multistimulus, portable and programmable conditioning panel for pigeons. *Behavior Research Methods, Instruments, & Computers*, 28, 49-54.
- XIA, L., WYNNE, C. D. L., MÜNCHOW-POHL, F. VON, & DELIUS, J. D. (1991). Psychobasic: A basic dialect for the control of psychological experiments with the Commodore 64 and Dela interfacing. *Behavior Research Methods, Instruments & Computers*, 23, 72-76.
- YARCZOWER, M., & CURTO, K. (1972). Stimulus control in pigeons after extended discriminative training. *Journal of Comparative and Physiological Psychology*, 80, 484-489.
- ZENTALL, T. R., & SHERBURNE, L. M. (1994). Transfer of value from S+ to S- in a simultaneous discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 1-8.